
Electron paramagnetic resonance spectroscopic analyses of membrane transport proteins

Dissertation
der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

vorgelegt von
Daniel Hilger
aus Georgsmarienhütte

München
26. Januar 2010

GUTACHTER:

- 1. Prof. Dr. Heinrich Jung**
- 2. Prof. Dr. Dirk Schüler**

TAG DER MÜNDLICHEN PRÜFUNG: 04. Mai 2010



THE TRANSMEMBRANE DOMAIN BY MELANIE HILGER 2009

Publications and manuscripts originating from this thesis

CHAPTER 2

Hilger, D., Böhm, M., Hackmann, A., and Jung, H. (2008) Role of Ser340 and Thr341 in transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli* in ligand binding and transport. *J. Biol. Chem.* 283: 4921-4929.

CHAPTER 3

Raba, M., Baumgartner, T., Hilger, D., Klempahn, K., Härtel, T., Jung, K., and Jung, H. (2008) Function of transmembrane domain IX in the Na⁺/proline transporter PutP. *J. Mol. Biol.* 382: 884-893.

CHAPTER 4

Hilger, D., Polyhach, Y., Jung, H., and Jeschke, G. (2009) Backbone structure of transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli*. *Biophys. J.* 96: 217-225.

CHAPTER 5

Hilger, D., Polyhach, Y., Jung, H., and Jeschke, G. (2009) Modeling of the helix bundle of the Na⁺/proline transporter PutP of *Escherichia coli* based on EPR distance restraints and structural homology. *Manuscript*

CHAPTER 6

Hilger, D., Jung, H., Padan, E., Wegener, C., Vogel, K.-P., Steinhoff, H.-J., and Jeschke, G. (2005) Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na⁺/H⁺ antiporter of *E. coli*. *Biophys. J.* 89: 1328-1338.

CHAPTER 7

Hilger, D., Polyhach, Y., Padan, E., Jung, H., and Jeschke, G. (2007) High-resolution structure of a Na⁺/H⁺ antiporter dimer obtained by pulsed electron paramagnetic resonance distance measurements. *Biophys. J.* 93: 3675-3683.

CHAPTER 8

Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H. (2006) DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Reson.* 30: 473-498.

Reviews and book chapters (not presented in this thesis)

Jung, H., Pirch, T., and Hilger, D. (2006) Secondary transport of amino acids in prokaryotes. *J. Membr. Biol.* 213: 119-133.

Hilger, D., and Jung, H. (2009) Protein chemical and EPR spectroscopic approaches to monitor membrane protein structure and dynamics. In *Bacterial Signaling* (Krämer, R., and Jung, K., eds.) pp. 247-263, Wiley-VCH, Weinheim.

Contributions to publications and manuscripts presented in this thesis

CHAPTER 2

M. Böhm constructed the mutants, performed the transport analyses, and determined the kinetic parameters of the PutP derivatives in intact cells. D. Hilger analyzed the expression of the mutants, generated the double-Cys mutants, and established the *in situ* cross-linking approach. A. Hackmann carried out the accessibility analyses. H. Jung performed the analyses of the ion dependence of proline uptake and determined the kinetic parameters of the PutP derivatives in proteoliposomes. He wrote the manuscript and discussed the results with D. Hilger.

CHAPTER 3

M. Raba, T. Baumgartner, and K. Klempahn constructed the mutants. M. Raba, T. Baumgartner, K. Klempahn, and T. Härtel performed the transport measurements. M. Raba and T. Baumgartner analyzed the expression of the mutants, determined the kinetic parameters of the PutP derivatives, and performed the accessibility analyses. D. Hilger analyzed the influence of ligands on MTSET labeling. He conducted the experiments performed by M. Raba, T. Baumgartner, T. Härtel, and K. Klempahn. H. Jung wrote the manuscript and discussed the results with D. Hilger and M. Raba.

CHAPTER 4

D. Hilger constructed the mutants. He prepared the samples and performed the functional analyses and DEER measurements. Y. Polyhach analyzed the data, established the modeling approach, and generated the rotamer libraries together with G. Jeschke. H. Jung developed the concept of the study together with G. Jeschke, Y. Polyhach, and D. Hilger. G. Jeschke wrote the manuscript together with D. Hilger, Y. Polyhach, and H. Jung.

CHAPTER 5

D. Hilger constructed the mutants. He prepared the samples and performed the functional analyses and EPR measurements. Y. Polyhach analyzed the data and established the modeling approach together with G. Jeschke. H. Jung developed the concept of the study together with D. Hilger, G. Jeschke, and Y. Polyhach. D. Hilger wrote the manuscript with G. Jeschke.

CHAPTER 6

E. Padan provided the mutants. D. Hilger prepared the samples and performed the DEER measurements. He carried out the CW EPR power saturation measurements and analyzed the data together with K.-P. Vogel. C. Wegener performed the room- and low-temperature

CW EPR measurements. H. Jung developed the concept of the study together with H.-J. Steinhoff, G. Jeschke, and E. Padan. G. Jeschke wrote the manuscript with D. Hilger, H.-J. Steinhoff, and H. Jung.

CHAPTER 7

E. Padan provided the mutants. D. Hilger prepared the samples and performed the DEER measurements. Y. Polyhach analyzed the data, established the approach, and generated the rotamer libraries together with G. Jeschke. H. Jung developed the concept of the study together with G. Jeschke and E. Padan. G. Jeschke wrote the manuscript with D. Hilger, Y. Polyhach, and H. Jung.

CHAPTER 8

G. Jeschke designed research together with V. Chechik, A. Godt, C. Timmel, and H. Jung. He analyzed the data, developed the software and wrote the manuscript. P. Ionita synthesized spin-labeled gold nanoparticles. H. Zimmermann synthesized deuterated shape-persistent biradical model compounds. A. Godt synthesized shape-persistent biradical model compounds. J. Banham analyzed data and performed DEER measurements. D. Hilger prepared spin-labeled proteins and performed DEER measurements.

I hereby confirm the above statements:

Daniel Hilger

Prof. Dr. Heinrich Jung

INDEX

ABBREVIATIONS	XI
SUMMARY	1
ZUSAMMENFASSUNG	3
CHAPTER 1	
Introduction	6
1.1. Diversity of transport systems.....	6
1.2. Diversity and function of secondary transporters.....	7
1.3. Coupling ions of secondary transporter systems.....	8
1.4. Mechanism of secondary transport.....	9
1.5. Structure of secondary transporters.....	12
1.5.1. <i>Structure of transmembrane domains</i>	12
1.5.2. <i>Transmembrane helix packing and structural symmetry</i>	13
1.5.3. <i>Quaternary structure</i>	15
1.6. Site-directed spin labeling and EPR spectroscopy on membrane proteins.....	17
1.7. The Na ⁺ /proline symporter PutP of <i>E. coli</i>	20
1.8. The Na ⁺ /H ⁺ antiporter NhaA of <i>E. coli</i>	24
1.9. Scope of this work.....	30
1.10. References.....	31
CHAPTER 2	
Role of Ser340 and Thr341 in transmembrane domain IX of the Na⁺/proline transporter PutP of Escherichia coli in ligand binding and transport	44
2.1. Abstract.....	44
2.2. Introduction	44
2.3. Materials and Methods	45
2.3.1. <i>Bacterial strains and plasmids</i>	45
2.3.2. <i>Site-directed mutagenesis</i>	46
2.3.3. <i>Proline transport in intact cells</i>	46
2.3.4. <i>Immunological analysis</i>	46
2.3.5. <i>Cys accessibility analyses</i>	47
2.3.6. <i>In situ site-directed cross-linking</i>	47
2.3.7. <i>Purification and reconstitution of PutP</i>	48
2.3.8. <i>Transport assay with proteoliposomes</i>	48
2.3.9. <i>Determination of Na⁺</i>	48
2.3.10. <i>Protein determination</i>	49
2.4. Results.....	49
2.4.1. <i>Conserved amino acids of TM IX</i>	49
2.4.2. <i>Effect of the substitution of Ser340 and Thr341 on proline uptake and putP expression in E. coli WG170</i>	49
2.4.3. <i>Influence of the substitutions on the ion selectivity of PutP</i>	50
2.4.4. <i>Kinetic analyses</i>	52
2.4.5. <i>Accessibility to sulfhydryl reagents</i>	54
2.4.6. <i>Proximity between TM II and TM IX</i>	55
2.5. Discussion.....	57
2.6. Acknowledgments	61
2.7. References.....	61
2.8. Supplementary Material.....	65
2.8.1. <i>Sequence alignment of members of the SSS family</i>	65
2.8.2. <i>Establishment of protease cleaving sites in PutP</i>	65
2.8.3. <i>Effect of ligand on Cys cross-linking</i>	67
2.9. References.....	68

CHAPTER 3

Function of transmembrane domain IX in the Na⁺/proline transporter PutP..... 69

3.1. Abstract.....	69
3.2. Introduction	69
3.3. Materials and Methods	71
3.3.1. <i>Bacterial strains and plasmids</i>	71
3.3.2. <i>Site-directed mutagenesis</i>	71
3.3.3. <i>Purification and reconstitution of PutP</i>	72
3.3.4. <i>Proline transport assays</i>	72
3.3.5. <i>Cys accessibility to fluorescein-5-maleimide</i>	73
3.3.6. <i>Western Blot analysis</i>	73
3.3.7. <i>Protein determination</i>	73
3.4. Results.....	73
3.4.1. <i>Cys scanning mutagenesis of TM IX</i>	73
3.4.2. <i>Effect of Cys replacements in TM IX of PutP(wild type) on transport kinetics</i>	74
3.4.3. <i>Influence of Cys modification on transport activity</i>	76
3.4.4. <i>Accessibility of Cys to fluorecein-5-maleimide</i>	78
3.4.5. <i>Influence of ligand on Cys accessibility</i>	80
3.5. Discussion.....	81
3.6. Acknowledgments	84
3.7. References.....	85

CHAPTER 4

Backbone structure of transmembrane domain IX of the Na⁺/proline transporter PutP of Escherichia coli..... 88

4.1. Abstract.....	88
4.2. Introduction	88
4.3. Materials and Methods	90
4.3.1. <i>Sample preparation</i>	90
4.3.2. <i>DEER measurements</i>	91
4.3.3. <i>Rotamer library</i>	92
4.3.4. <i>Structure determination</i>	92
4.4. Results and Discussion	93
4.4.1. <i>Test of the helix-loop-helix model for TM domains in transporters</i>	93
4.4.2. <i>TM domain IX of PutP is kinked</i>	96
4.4.3. <i>Functional implications of the structural model</i>	102
4.5. Conclusions	103
4.6. Acknowledgments	104
4.7. References.....	104
4.8. Supplementary Material.....	108
4.9. References.....	112

CHAPTER 5

Modeling of the helix bundle of the Na⁺/proline transporter PutP of Escherichia coli based on EPR distance restraints and structural homology 113

5.1. Abstract.....	113
5.2. Introduction	113
5.3. Materials and Methods	116
5.3.1. <i>Sample preparation</i>	116
5.3.2. <i>DEER measurements</i>	118
5.3.3. <i>Determination of number of spins per molecules</i>	118
5.3.4. <i>Generation of template restraints</i>	119
5.3.5. <i>Structure determination</i>	119
5.3.6. <i>Structure refinement</i>	120
5.4. Results and Discussion	121
5.4.1. <i>Derivation of internal restraints</i>	121
5.4.2. <i>Internal constraints on interhelical distances</i>	123
5.4.3. <i>Test of the approach on LeuT_{Aa}</i>	124
5.4.4. <i>Helix bundle model for PutP</i>	128

5.4.5. <i>Functional implications of the helix bundle models</i>	132
5.5. Conclusions	136
5.6. Acknowledgments	137
5.7. References.....	137
5.8. Supplementary Material.....	141
5.9. References.....	145
CHAPTER 6	
<i>Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na⁺/H⁺ antiporter of E. coli</i>	146
6.1. Abstract.....	146
6.2. Introduction	146
6.3. Materials and Methods	148
6.3.1. <i>Preparation and labeling of NhaA</i>	148
6.3.2. <i>Four-pulse DEER measurements</i>	149
6.3.3. <i>Four-pulse DEER data analysis</i>	149
6.3.4. <i>CW EPR measurements</i>	150
6.3.5. <i>Simulation of CW EPR spectra</i>	151
6.4. Results.....	151
6.4.1. <i>Local mobility-CW EPR spectra at room temperature</i>	151
6.4.2. <i>Polarity- and proximity-CW EPR spectra at low temperature</i>	153
6.4.3. <i>Accessibility measurements by CW EPR</i>	153
6.4.4. <i>Calibration of spin counting by DEER measurements</i>	154
6.4.5. <i>Evidence for specific oligomerization by four-pulse DEER</i>	155
6.4.6. <i>Dependence of the average degree of oligomerization on pH</i>	158
6.5. Discussion.....	160
6.5.1. <i>Comparison with structural models for NhaA dimers</i>	160
6.5.2. <i>Mechanistic implications of the observed pH dependence</i>	162
6.6. Acknowledgments	164
6.7. References.....	164
CHAPTER 7	
<i>High-resolution structure of a Na⁺/H⁺ antiporter dimer obtained by pulsed EPR distance measurements</i>	168
7.1. Abstract.....	168
7.2. Introduction	168
7.3. Materials and Methods	170
7.3.1. <i>Preparation and labeling of NhaA</i>	170
7.3.2. <i>EPR measurements</i>	171
7.3.3. <i>Estimates of mean distances</i>	171
7.3.4. <i>Grid search for determination of the initial structure</i>	172
7.3.5. <i>Structure refinement by direct fit to the primary experimental data</i>	172
7.3.6. <i>Estimate of the resolution of the structure</i>	172
7.3.7. <i>Superposition of the EPR and cryo-electron microscopy structures</i>	173
7.4. Results and Discussion	173
7.4.1. <i>Structure determination and refinement</i>	173
7.4.2. <i>Contacts in the NhaA dimer</i>	178
7.4.3. <i>Comparison with the low-resolution structure</i>	179
7.4.4. <i>Functional implications of the EPR structure</i>	181
7.5. Acknowledgments	183
7.6. References.....	184
CHAPTER 8	
<i>DeerAnalysis2006 - A comprehensive software package for analyzing pulsed ELDOR data</i> ...	187
8.1. Abstract.....	187
8.2. Introduction	187
8.3. Materials and Methods	189
8.4. Theory	190
8.4.1. <i>Separation into form factor and background factor</i>	190

8.4.2. Tikhonov regularization	192
8.4.3. Fitting of the form factor by a model.....	195
8.4.4. Spin counting	196
8.4.5. Excitation bandwidth correction	197
8.5. Algorithms	199
8.5.1. Preprocessing	199
8.5.2. Separation into form factor and background factor.....	200
8.5.3. Determining and using experimental background functions.....	201
8.5.4. Concentration measurements.....	202
8.5.5. Tikhonov regularization with L curve computation	203
8.5.6. Fitting of models for the distance distribution	203
8.5.7. Suppression test for artefact peaks in the distance distribution	204
8.5.8. Dual display-scaling procedures for direct comparison of two data sets	206
8.6. Examples	206
8.6.1. Asymmetric distance distribution in shape-persistent biradicals.....	206
8.6.2. Broad distance distribution in spin-labelled gold nanoparticles	207
8.6.3. Derivation and test of an experimental background function for a membrane protein.....	210
8.7. Appendix	211
8.8. Acknowledgments	211
8.9. References.....	212
CHAPTER 9	
Discussion.....	216
9.1. Functional role and structural dynamics of transmembrane domain IX of PutP.....	216
9.2. Backbone structure of transmembrane domain IX of PutP	220
9.3. Helix packing of PutP	224
9.4. Structural dynamics of the NhaA dimer.....	229
9.5. High-resolution structure of the NhaA dimer.....	232
9.6. DeerAnalysis2006 - A comprehensive software package for analyzing of pulsed ELDOR data.....	237
9.7. References.....	241
ACKNOWLEDGMENTS	247
CURRICULUM VITAE	249
EHRENWÖRTLICHE VERSICHERUNG.....	252

ABBREVIATIONS

APT	approximate Pake transformation
AU	arbitrary units
BMH	1,6-bis-(maleimido) hexane
BSA	bovine serum albumin
CROX	chromium oxalate
cryo-EM	cryoelectron microscopy
CuPh	copper(1,10-phenanthroline) ₃
CW	continuous-wave
DEER	double electron-electron resonance
ELDOR	electron-electron double resonance
EPR	electron paramagnetic resonance
FM	fluorescein-5-maleimide
HRP	horseradish peroxidase
Imf	lithium motive force
Mes	2-[<i>N</i> -morpholino]ethanesulfonic acid
MMTS	methyl methanethiosulfonate
MTSEA	methanethiosulfonate ethylammonium
MTSES	methanethiosulfonate ethylsulfonate
MTSET	methanethiosulfonate ethyltrimethylammonium
MTSSL	(1-oxyl-2,2,5,5,-tetramethylpyrroline-3-methyl)-methanethiosulfonate
m.w.	microwave
NEM	<i>N</i> -ethylmaleimide
Ni ²⁺ -NTA	Ni ²⁺ -nitrilotriacetic acid
NMR	nuclear magnetic resonance
NSS	neurotransmitter/sodium symporter
PELDOR	pulsed electron-electron double resonance
pmf	proton motive force
<i>p</i> -PDM	<i>N,N'</i> - <i>p</i> -phenylenedimaleimide
PutP(Δ Cys)	engineered PutP devoid of all five native cysteine residues
RMSD	root mean-square deviation
SD	standard deviations
SDSL	site-directed spin labeling
smf	sodium motive force
SSS	Na ⁺ /solute symporter
TEMPOL	4-hydroxy-2,2,6,6-tetramethylpiperidine- <i>N</i> -oxyl
TM	transmembrane domain
TRIS	Tris(hydroxymethyl)aminomethane

SUMMARY

Secondary transporters use the free energy stored in ion and/or solute gradients to drive the selective transport of a variety of substrates across the membrane of pro- and eukaryotic cells. To understand the molecular mechanism underlying solute transport, knowledge of the structure and structure-function relationships of secondary transporters and their complexes is a basic prerequisite. However, due to their size, hydrophobicity, and conformational flexibility, the generation of high-resolution structural data by X-ray crystallography and NMR is still a challenge. The main purpose of this thesis was to develop novel approaches for structure elucidation of membrane proteins and their complexes based on site-directed spin labeling (SDSL) and pulsed electron paramagnetic resonance (pulsed EPR) spectroscopy. By using the Na⁺/proline symporter PutP and the Na⁺/H⁺ antiporter NhaA of *Escherichia coli* as model systems, new insights into the structure and mechanism of these ion-coupled secondary transporters were obtained.

In the first part of this work, the functional role and structure of transmembrane domain IX of PutP were analyzed. Individual site-directed replacement of all amino acids of TM IX and determination of the kinetic properties and accessibility of the derivatives indicate that the domain participates in the formation of a ligand-translocation pathway in PutP. Six residues were identified as being important for ligand binding and/or translocation of which Thr341 is suggested to be directly involved in the coordination of Na⁺. Furthermore, cross-linking of paired Cys residues placed in TMs II and IX supports the idea of a ligand-binding site located at the intersection of both domains. Moreover, investigation of ligand-dependent changes in the accessibility of positions in TM IX implicates that this domain is involved in conformational alterations possibly associated with the alternating access mode of ligand transport. The backbone structure of TM IX was determined by SDSL and double electron-electron resonance (DEER) distance measurements. For this purpose, a new approach was developed based on helix-loop-helix models and explicit modeling of spin label conformations. The resulting ensemble of backbone structures shows a pronounced kink near residue Thr341. This kink may serve as a hinge that alters its conformation upon substrate binding.

In the second part of this thesis, the helix bundle of PutP was modeled based on DEER distance measurements between spin labels attached to helix ends and template restraints derived from the ten-helix core of the vSGLT X-ray structure. By applying a novel approach that relies on distance geometry, the fold of the core of PutP was reliably recognized and the location of the three non-core helices of PutP with respect to the core was established. Comparison of the obtained core structure of PutP with known structures of structurally related transporters leads to a model of putative ligand-binding sites for PutP.

The pH-regulated Na⁺/H⁺ antiporter NhaA is known to function as a dimer. In the third part of this work, continuous-wave (CW) EPR techniques and DEER distance measurements were used to investigate pH-induced changes in the structure of the NhaA monomer and dimer. Analyses of the mobility, accessibility, and polarity of the environment of the spin label attached to positions 225 and 254 at different pH values suggest that residues His225 and Val254 do not participate in the formation of the pH sensor or ion-translocation site. Rather they are located in domains that convey the pH signal from the pH sensor to the translocation site. Furthermore, intermolecular DEER distance measurements indicate that NhaA exists as a dimer in the membrane. Quantification of pH-dependent alterations of the modulation depth of the DEER signals suggests a dimerization equilibrium, which depends moderately on pH. To understand the role of dimerization for ion transport and/or pH-regulation of NhaA, the first high-resolution structure of the NhaA dimer was determined. For this purpose, a new approach was developed for structure elucidation of protein complexes in the native environment. This approach is based on intermolecular DEER distance measurements and explicit modeling of spin label conformations. The EPR structure shows two contact sites between the monomers and suggests an ion-translocation mechanism that depends on relative motion of the translocation domain with respect to the interfacial domain.

Finally, the software package DeerAnalysis2006 is presented for analyzing pulsed electron-electron double resonance data. Its graphical user interface provides access to a comprehensive set of tools for extracting distance distributions from dipolar evolution functions. Features include approximate Pake transformation, Tikhonov regularization with L curve computation, model-based fitting, excitation bandwidth correction, spin counting, concentration measurements, and background correction based on experimental data sets of singly labeled samples. Furthermore, the interface allows for several tests of the reliability and precision of the extracted information.

ZUSAMMENFASSUNG

Sekundäre Transporter nutzen die freie Energie von Ionen- und/oder anderen Stoffgradienten, um selektiv eine Vielzahl von Substraten über die Zytoplasmamembran von pro- und eukaryotischen Zellen zu transportieren. Um den Transportmechanismus auf molekularer Ebene verstehen zu können, werden Informationen über die Struktur und Struktur-Funktionsbeziehungen von sekundären Transportern und deren Komplexen benötigt. Aufgrund ihrer Größe, Hydrophobizität und konformationellen Flexibilität stellt die Gewinnung von hochaufgelösten strukturellen Daten mittels Röntgenkristallographie und Kernspinresonanzspektroskopie jedoch nach wie vor eine große Herausforderung dar. Das primäre Ziel dieser Arbeit war daher die Entwicklung neuer Methoden zur Strukturaufklärung von Membranproteinen und deren Komplexen mittels ortsspezifischer Spinmarkierung und gepulster Elektronen-paramagnetischer Resonanz(EPR)-Spektroskopie. Unter der Verwendung des Na⁺/Prolin-Symporters PutP und des Na⁺/H⁺-Antiporters NhaA von *Escherichia coli* als Modellsysteme konnten neue Erkenntnisse über Struktur und Mechanismus dieser Ionen-gekoppelten sekundären Transporter gewonnen werden.

Im ersten Teil dieser Arbeit wurde die funktionelle Bedeutung und Struktur der Transmembrandomäne (TM) IX von PutP untersucht. Durch den individuellen ortsspezifischen Austausch aller Aminosäurereste in TM IX und Ermittlung der kinetischen Eigenschaften und Zugänglichkeit der jeweiligen Derivate konnte eine Beteiligung der Domäne am Liganden-Translokationsweg in PutP gezeigt werden. Sechs Reste wurden identifiziert, die eine wichtige Rolle für die Bindung und/oder Translokation der Liganden spielen. Von diesen Aminosäureresten konnte für Thr341 eine direkte Beteiligung an der Koordination des Na⁺ nachgewiesen werden. Weiterhin zeigen Quervernetzungen von Cysteinresten, die in TM II und TM IX eingefügt wurden, dass die Ligandenbindestelle zwischen diesen Domänen lokalisiert ist. Darüber hinaus wurde über den Nachweis einer Liganden-abhängigen Veränderung der Zugänglichkeit einzelner Aminosäurereste in TM IX gezeigt, dass TM IX an konformationellen Änderungen beteiligt ist, die mit dem *alternating access*-Mechanismus des Transportes von Liganden assoziiert sind. Die Struktur des Proteinrückgrades von TM IX wurde mittels ortsgerichteter Spinmarkierung und Doppel-Elektronen-Elektronen-Resonanz (DEER) Distanzmessungen ermittelt. Zu diesem Zweck wurde eine neue Methode entwickelt, die auf Helix-Loop-Helix Modellen und Modellierung der Konformationen des Spinmarkers basiert. Das resultierende Strukturensemble zeigt eine deutliche Krümmung in TM IX in der Nähe des Restes Thr341. Diese Krümmung fungiert vermutlich als eine Art von Scharnier, dessen Konformation durch Bindung der Substrate verändert wird.

Im zweiten Teil der Arbeit wurde das Helixbündel von PutP mit Hilfe von DEER Distanzmessungen zwischen spinmarkierten Helixenden und Strukturbedingungen, die von der 10-Helix-Kernstruktur der vSGLT Kristallstruktur abgeleitet wurden, modelliert. Unter Anwendung einer neu entwickelten, auf Distanzgeometrie basierenden Methode konnte die Faltung der 10-Helix-Kernstruktur von PutP verlässlich erkannt und die Lokalisierung der drei weiteren transmembranen Helices relativ zur Kernstruktur bestimmt werden. Ein Vergleich der ermittelten Kernstrukturen von PutP mit denen bereits bekannter Kristallstrukturen von strukturell verwandten Transportern ermöglichte die Entwicklung eines Modells für vermutliche Ligandenbindestellen für PutP.

Der pH-regulierte Na^+/H^+ Antiporter NhaA liegt in der Membran als funktionelles Dimer vor. Im dritten Teil dieser Arbeit wurden *continuous-wave* (CW) EPR Techniken und DEER Distanzmessungen verwendet, um pH-induzierte Strukturveränderungen innerhalb der Monomer- und Dimerstruktur von NhaA zu untersuchen. Durch eine Analyse der Dynamik, Zugänglichkeit und Umgebungspolarität der Spinmarkerseitenkette an den Positionen 225 und 254 bei verschiedenen pH Werten konnte gezeigt werden, dass His225 und Val254 nicht an der Ausbildung des pH Sensors oder an der Ionentranslokation beteiligt sind. Vielmehr wurde nachgewiesen, dass sie in Domänen lokalisiert sind, die das pH Signal von dem pH Sensor zum Translokationsort weiterleiten. Weiterhin wurde mittels intermolekulare DEER Distanzmessungen gezeigt, dass NhaA in der Membran von Proteoliposomen als Dimer vorliegt. Durch Quantifizierung von pH-abhängigen Änderungen in der Modulationstiefe des DEER Signals konnte ein Dimersierungs-Gleichgewicht nachgewiesen werden, welches moderat vom pH beeinflusst wird. Um die Rolle der Dimerisierung für den Ionentransport und/oder pH Regulation zu verstehen, wurde erstmalig eine hochaufgelöste Struktur des Dimers von NhaA ermittelt. Zu diesem Zweck wurde ein neues Verfahren entwickelt, das die Strukturbestimmung von Proteinkomplexen in deren natürlichen Umgebung erlaubt. Die Methode basiert auf intermolekularen DEER Distanzmessungen und Modellierung der Spinmarker-Konformationen. Mit Hilfe der EPR Struktur konnten zwei Interaktionsstellen zwischen den Monomeren identifiziert und Hinweise für einen Ionen-Translokationsmechanismus gewonnen werden, der eine Bewegung der Translokationsdomäne relativ zur Grenzflächendomäne beinhaltet.

Im letzten Teil der vorliegenden Arbeit wird das Softwarepaket DeerAnalysis2006 präsentiert, welches die Analyse von Puls-Elektron-Elektron-Doppel-Resonanz-Daten erlaubt. Die graphische Benutzeroberfläche ermöglicht den Zugriff auf verschiedene Anwendungsfunktionen für die Extraktion von Abstandsverteilungen aus dipolaren Evolutionsfunktionen. Desweiteren beinhaltet DeerAnalysis2006 folgende Funktionen: Pake Transformation, Tikhonov Regularization mit Berechnung von L-Kurven, Modell-basiertes Fitten, Korrektur von Anregungsbandbreiten, Spinzählung, Konzentrationsmessungen und

Hintergrundkorrektur mit Hilfe von experimentellen Datensätzen von einfach markierten Proben. Auch ermöglicht das Programm, die Verlässlichkeit und Präzision der extrahierten Informationen auf verschiedene Art und Weise zu testen.

CHAPTER 1

Introduction

1.1. Diversity of transport systems

Biological membranes are a basic prerequisite for life. They separate the cell's and organelle's interior from the environment and determine the exchange of compounds between the compartments. Unlike small polar and apolar molecules, solutes such as organic acids, peptides, sugars, and ions cannot pass the permeability barrier of the lipid bilayer by free diffusion. Therefore, membranes contain passive or active transport systems that allow the selective uptake and excretion of these compounds. The active transport of a solute against its concentration gradient is an endergonic reaction that requires an input of energy. According to the source of energy utilized, three types of transport systems can be discriminated (Mitchell, 1991) (Fig. 1.1.): (1) Primary transport systems use light, redox, or chemical energy to generate electrochemical gradients of ions or other solutes. Examples are the light-driven ion pumps bacterio- and halorhodopsin (Lanyi, 2006; Essen, 2002), Na⁺-translocating decarboxylases (Dimroth, 1997), and ATP-driven transport systems (Davidson and Maloney, 2007). (2) Group translocation systems couple the transport of a solute to the chemical modification of the solute. The energy for the transport process is provided by phosphoenolpyruvate that also functions as a phosphoryl donor for the phosphorylation of the transported solute (Reizer and Saier, 1997). The only known examples are the phosphoenolpyruvate:sugar phosphotransferase systems (PTS) of bacteria. (3) Secondary transport systems utilize the energy of an (electro)chemical gradient of a solute or ion to drive the uphill transport of another solute. Depending on the directions of transport of both solutes, symport and antiport processes are discriminated. Coupled transport of two or more solutes in the same direction is indicated as symport. The most extensively studied symporter is the H⁺/β-galactoside transporter LacY of *Escherichia coli* (Guan and Kaback, 2006). Further examples are sodium-coupled neurotransmitter transporters of *Homo sapiens* (e.g. SERT) (Rudnick, 1998; Sonders *et al.*, 2005) and the Na⁺/proline transporter PutP of *E. coli* (Jung, 1998). Antiport is defined as coupled transport of solutes in opposite directions and is exemplified by the Na⁺/H⁺ antiporter NhaA of *E. coli* (Padan *et al.*, 2004) and the mitochondrial ADP/ATP transporter of *Bos taurus* (Pebay-Peyroula *et al.*, 2003). Because of the similarity in transport kinetics and mechanism, a third category of secondary transport is defined as uniport. It comprises the unidirectional translocation of a single solute merely driven by the (electro)chemical gradient of the transported solute itself. A well known example is the glucose transporter Glf of *Zymomonas mobilis* (DiMarco and Romano, 1985; Barnell *et al.*, 1990).

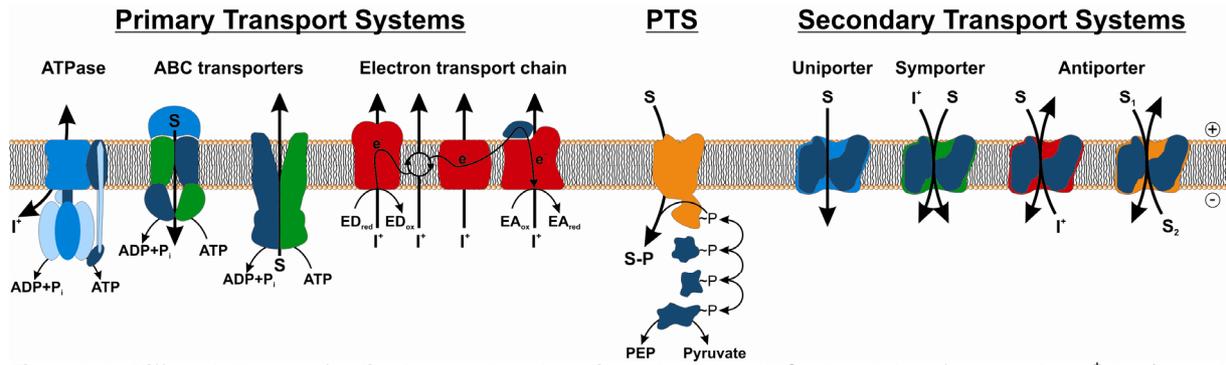


Figure 1.1.: Different classes of active transport systems in prokaryotes. PTS: phosphotransferase system, I^- : ion (in most cases H^+ or Na^+), S: solute, ED: electron donor, EA: electron acceptor, PEP: phosphoenolpyruvate.

1.2. Diversity and function of secondary transporters

Secondary transporters are found ubiquitously in all three kingdoms of living organism. In prokaryotic and eukaryotic systems their percentage among the total transporters varies between 17% and 80% (Ren and Paulsen, 2005). Thereby, genes encoding for secondary transporters are most divergent in prokaryotes and unicellular eukaryotes (Ren and Paulsen, 2005; Nagata *et al.*, 2008). Based on phylogeny and function, including mode of transport, energy coupling mechanism, and substrate specificity, approximately 200 families of secondary transporters are currently distinguished in the transporter classification database (TCDB) (Saier, 1998; Saier, 1999). The largest group of secondary active membrane transporters is represented by the major facilitator superfamily (MFS, TC 2.A.1) with over ten thousand identified members of pro- and eukaryotic origin in 69 different families (Saier *et al.*, 1999). This diversity of secondary transporters is reflected by the broad range of substrates that can be translocated, e. g., sugars, amino acids, ions, nucleosides, organic acids, compatible solutes, vitamins, neurotransmitters, peptides, urea, heavy metals, autoinducers, and tRNAs (Saier, 1999). Based on this substrate diversity, secondary transport systems play vital roles in cellular processes like catabolism, anabolism, homeostasis, detoxification, cell-cell communication, and energy metabolism. Examples for transporters, which facilitate these cellular processes are the glucose uniporter GLUT1 of *H. sapiens*, the lactose permease LacY, and the proline transporter PutP of *E. coli* that feed catabolic and anabolic pathways (Mueckler, 1994; Guan and Kaback, 2006; Jung, 2001). The symporters ProP of *E. coli*, BetP from *Corynebacterium glutamicum*, and OpuE of *Bacillus subtilis* are osmoregulated systems, which play a role in the osmoadaptation of the respective organisms (Wood, 2007; Krämer and Morbach, 2004; von Blohn *et al.*, 1997). Other adaptive cellular processes such as pH and Na^+ homeostasis involve Na^+/H^+ antiporters like the bacterial NhaA and the mammalian NHE1 (Padan *et al.*, 2005; Orłowski and Grinstein, 2004). The detoxification of cytotoxic compounds is another important cellular process. In this context, multidrug transporters like EmrD and EmrE of *E. coli* were found to

catalyze the efflux of a variety of toxic substances (e.g. antibiotics and uncouplers) (Yin *et al.*, 2006; Schuldiner *et al.*, 1997). Cell-cell communication in bacteria (quorum sensing) involves small hormone-like molecules, termed auto-inducers (AI), whose extracellular concentration increases as a function of increasing cell-population density (Camilli and Basler, 2006). In *E. coli*, TqsA was identified as an exporter for autoinducer-2 (AI-2) that regulates quorum sensing dependent processes, like biofilm formation (Herzberg *et al.*, 2006). Finally, some secondary transporters in anaerobic bacteria are used to generate metabolic energy. Examples are the oxalate/formiate antiporter OxIT of *Oxalobacter formigenes* (Ruan *et al.*, 1992), the citrate/lactate antiporter CitP from *Leuconostoc mesenteroides*, and the malate/lactate antiporter MleP of *Lactococcus lactis* (Bandell *et al.*, 1997). These antiporters couple the import of a divalent precursor (e.g. oxalate, citrate, or malate) to the export of its monovalent decarboxylation product (e.g. formiate or lactate). In addition to the cytosolic consumption of a proton by decarboxylation of the precursor, this leads to the formation of a membrane potential and a pH gradient. Energy stored in this way can then be used to drive other processes like ATP synthesis, flagellar rotation, or solute transport (Lolkema *et al.*, 1996).

1.3. Coupling ions of secondary transporter systems

The most important energy source of secondary transporters in prokaryotes is the electrochemical proton gradient ($\Delta\tilde{\mu}_{H^+}$) across the cytoplasmic membrane. This gradient is usually built up by primary proton pumps, such as respiratory and photosynthetic redox chains, pyrophosphatases, or in some anaerobic bacteria by ATPases or secondary ion-translocating processes (Skulachev, 1991; Konings *et al.*, 1995; Serrano *et al.*, 2004). As a result of the vectorial H^+ translocation, a proton motive force (*pmf* or Δp) is generated

$$pmf = \frac{\Delta\tilde{\mu}_{H^+}}{F} = -\frac{2.3RT}{F}\Delta pH + \Delta\psi, \quad (1.1.)$$

which is composed of a pH gradient (ΔpH), inside alkaline versus outside, and an electrical potential ($\Delta\psi$) across the membrane (Mitchell, 1966). According to the chemiosmotic hypothesis of Peter Mitchell (1961), the *pmf* provides a source of stored energy that can be used to drive energy-requiring membrane-bound processes. One example is the translocation of solutes by secondary transporters like LacY (H^+ /lactose symporter) and NhaA (H^+ / Na^+ antiporter) of *E. coli* that use energy released from the influx of H^+ in response

¹ *R*: gas constant ($R = 8.3145 \text{ JK}^{-1}\text{mol}^{-1}$); *T*: absolute temperature, [K]; *F*: Faraday constant ($F = 9.6485 \cdot 10^4 \text{ Cmol}^{-1}$)

to the *pmf* to catalyze transport of substrate against a concentration gradient (Kaback, 2005; Padan *et al.*, 2004). Beside this *pmf*-coupled osmotic work, also chemical or mechanical work in terms of ATP synthesis by the F_0F_1 -ATPase (Junge *et al.*, 1997) and the rotation of the flagellar motor can be *pmf*-driven (Blair, 2003).

In addition to H^+ -dependent transport systems, more and more secondary transporters such as PutP (Na^+ /proline symporter) of *E. coli* and LeuT_{Aa} (Na^+ /leucine symporter) of *Aquifex aeolicus* were identified that use Na^+ as coupling ion (Jung, 1998; Yamashita *et al.*, 2005). Na^+ -dependent secondary transport processes imply the existence of specific Na^+ pumps for the generation of a sodium motive force (*smf*) that depends on ΔpNa^+ rather than ΔpH (cf. Eq (1.1.)). The first example of a primary Na^+ pump, the oxaloacetate decarboxylase of *Klebsiella pneumoniae*, was discovered in 1980 (Dimroth, 1980). Today, several types of primary Na^+ transporters are known, including Na^+ -translocating methyl-transferases (Gottschalk and Thauer, 2001), NADH-quinone reductases (Hayashi *et al.*, 2001), oxidoreductases (Tokuda, 1989), ATPases (Speelmans *et al.*, 1993), pyrophosphatases (Malinen *et al.*, 2007), and various decarboxylases with different substrate specificities (Dimroth, 1997). The resulting *smf* is not only used to drive Na^+ -dependent solute uptake, but also ATP-synthesis or flagellar rotation (Dimroth, 1994; Sato and Homma, 2000). In addition to these Na^+ pumps, Na^+/H^+ antiporters are used to convert a *pmf* generated by primary H^+ transport systems into a *smf* (Schuldiner and Padan, 1993). By this means, some prokaryotes, e.g. *E. coli*, are able to functionally complement the $\Delta\tilde{\mu}_{H^+}$ by the sodium potential difference ($\Delta\tilde{\mu}_{Na^+}$) and thereby possess both a *pmf* and *smf* across the membrane (Skulachev, 1985; Hase *et al.*, 2001). This feature enables the utilization of H^+ - and Na^+ -dependent secondary transporters simultaneously in the same cell or the usage of transporters, like the melibiose permease (MelB) of *E. coli*, which can use both, *pmf* and *smf*, to couple the uptake of substrates (Pourcher *et al.*, 1990). However, in some anaerobic prokaryotes, the *smf* substitute the *pmf* completely. Examples are *Propionigenium modestum*, *Malonomonas rubra*, and *Caloramator fervidus*, which exhibit an energy metabolism that relies exclusively on Na^+ (Hilpert *et al.*, 1984, Speelmans *et al.*, 1993; Dimroth *et al.*, 1997).

1.4. Mechanism of secondary transport

Transporter-mediated movement of polar solutes and ions across the apolar core of a lipid bilayer is thought to occur through a hydrophilic translocation pathway inside the transporter protein (Mitchell, 1991; West, 1997). This polar pathway connects the substrate- and ion-binding sites with both sides of the membrane. However, in contrast to channel

pores, the translocation pathway of transporters is only open to one surface of the lipid bilayer at once. Therefore, it is suggested that secondary transporters operate via an alternating access mechanism (West, 1997; Maloney, 1994; Jardetzky, 1966). In this mechanism, the transporter can possess two major alternating conformations in which the opening of the translocation pathway is inward (C_i) or outward facing (C_o) (Fig. 1.2.). Transitions between these two conformational states expose the substrate-binding site alternately to the *cis* and *trans* side of the lipid bilayer, and thereby allow the translocation of substrates across the membrane. On the basis of structural, biochemical, and biophysical analyses of secondary transporters, the alternating access mechanism can be explained by two different models of conformational changes: the gated-pore and the rocker-switch mode (Yernool *et al.*, 2004; Yamashita *et al.*, 2005; Huang *et al.*, 2003; Abramson *et al.*, 2003; Smirnova *et al.*, 2007; Kaback *et al.*, 2007) (Fig. 1.2.).

The gated-pore model stands for transporters in which the substrate- and ion-binding sites are flanked by two gates that can alternately allow access to either the inner or outer side of the membrane bilayer (Fig. 1.2.a). This conformational mode is suggested to apply to the Na^+ /leucine symporter LeuT_{Aa} of the neurotransmitter:sodium symporter (NSS) family (TC 2.A.22) and to the Na^+ /aspartate symporter Glt_{Ph} of the dicarboxylate/amino acid:cation symporter (DAACS) family (TC 2.A.23) (Yamashita *et al.*, 2005; Yernool *et al.*, 2004; Boudker *et al.*, 2007). In the former case the gates are comprised of hydrophobic residues and salt bridges that close the passage to either the extra- or intracellular solution, whereas in the case of the aspartate transporter Glt_{Ph} , this function is assumed by two re-entrant helical hairpins.

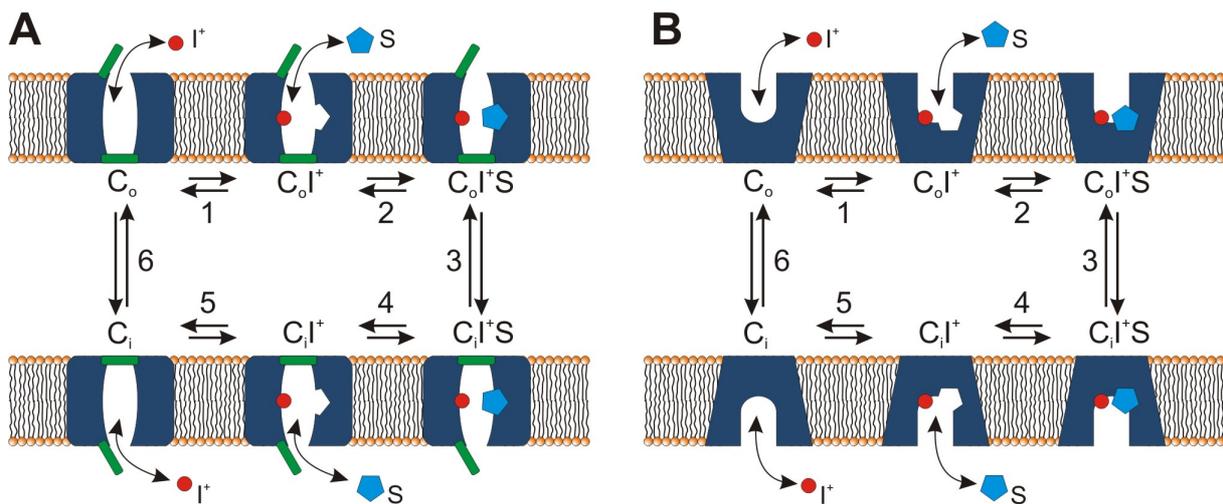


Figure 1.2.: Schematic diagram of alternating access mechanisms with a gated-pore (A) and a rocker-switch type of movement (B). The diagram describes the kinetic steps of ion-driven symport of solutes and the proposed conformational alterations of the transporter during the reaction cycle. I^+ , coupling ion (mostly H^+ or Na^+); S , substrate; C_o , outward facing conformation of the transporter; $C_o I^+$, binary complex of the transporter and coupling ion; $C_o I^+ S$, ternary complex of the transporter with bound coupling ion and substrate; C_i , inward facing conformation of the transporter; $C_i I^+$, binary complex of the transporter with bound coupling ion; numbers 1-6, steps of the reaction cycle (see text for further information).

The rocker-switch mode describes a domain movement of the N- and C-terminal portions of the transporter against each other (Fig. 1.2.b). This movement is thought to involve a rotation of the protein domains towards one another resulting in an interconversion between the outward and inward facing form of the transporter conformation (Smirnova *et al.*, 2007; Law *et al.*, 2008a; Law *et al.*, 2008b). As compared to the gated-pore mode, the rocker-switch mode requires larger conformational alterations to change between the outward and inward facing form of the transporter (Fig. 1.2.). Examples for this mechanistic principle are LacY (H⁺/lactose symporter), GlpT (glycerol-phosphate/P_i antiporter), EmrD (H⁺/multidrug antiporter) of *E. coli*, and OxlT (oxalate/formiate antiporter) from *Oxalobacter formigenes* (Abramson, 2003; Huang *et al.*, 2003; Yin *et al.*, 2006; Hirai *et al.*, 2002). All of these transporters belong to the MFS, suggesting the idea that the rocker-switch mode of movement is a common feature of this superfamily.

The reaction cycle of secondary transporters, working in an alternate accessibility mode, can be depicted by a simple kinetic scheme which is mostly broken into six discrete stages. Kinetic analyses of ion-dependent symporters, like LacY of *E. coli*, thereby suggest an ordered binding mechanism with sequential binding and release of the coupling ion and translocated substrate (Guan and Kaback, 2006) (Fig. 1.2.). Initially, in the absence of ligand, the symporters exist in an unloaded, outward-facing conformation (C_o). The first stage of the reaction cycle involves the binding of the coupling ion (I⁺) from the periplasmic side of the membrane. This binary complex (C_o-I⁺) induces rearrangements in the transporter structure accompanied by an increase of the affinity for the cotransported solute. After binding of the solute (S) in the second stage, the ternary complex (C_o-I⁺S) passes through conformational changes (third stage) that result in an inward orientation of the transporter (C_i-I⁺S). Upon successive release of the solute and the coupling ion to the inner side of the membrane in the fourth and fifth stage, the transporter reorientates to the outside resuming the initial state (C_o).

Whereas all known symporters may be presumed to operate via an ordered binding mechanism that requires two binding sites and a ternary complex, several antiporters are thought to possess either an alternating access mechanism with a single-binding site or a ternary complex (West, 1997, Law *et al.*, 2008a). Antiporters of the MFS, like GlpT, are presumed to operate via the single-binding site alternating access mechanism with a six step reaction cycle (Law *et al.*, 2008b). The first stage of the transport represents the binding of cosubstrate (S₁) to the unloaded outward facing carrier to yield the C_oS₁ complex that induces a flip of the conformation via the rocker-switch mechanism into the C_iS₁ state (second stage). In the third stage, the translocated bound cosubstrate is released to the inner side of the membrane resulting in an unloaded transporter in the inward facing conformation (C_i). The half reaction (step1-3) of the translocation is then simply reversed for the transport

of the countersubstrate (S_2) from the inner to the outer side of the membrane. With this transport mechanism there is only one substrate-binding site and no formation of a ternary complex. Thus, it is called single-binding site alternating access or ping-pong mechanism (West, 1997). In contrast, kinetics studies of the ADP/ATP antiporter (ADP/ATP carrier, AAC) of the mitochondrial carrier (MC) family (TC 2.A.29) lead to the conclusion of a transport cycle that involves a ternary complex, formed by simultaneous binding of the co- and countersubstrate to two substrate-binding sites located on different sides of the membrane. The formation of the ternary complex then triggers conformational changes to facilitate the translocation of both substrates in opposite directions (Duyckaerts *et al.*, 1980). However, it is still not known if the functional unit of the transporter is represented by a monomer or a dimer. Structural data of AAC favored an arrangement of two adjacent monomers in two different conformations, one binding a nucleotide from the outside and the other one from the inside. Transport then occurs by a conformational switchover between the two transitions in a cooperative fashion (Nury *et al.*, 2006).

1.5. Structure of secondary transporters

Knowledge of the structure of secondary transporters is a basic prerequisite for understanding the molecular mechanism of translocation. Today, three-dimensional (3D) X-ray structures of ten secondary transporters (AcrB, AAC, CLC, EmrD, EmrE, GlpT, Glt_{Ph}, LacY, LeuT_{Aa}, and NhaA) are available (Murakami *et al.*, 2002; Pebay-Peyroula *et al.*, 2003; Dutzler *et al.*, 2002; Yin *et al.*, 2006; Chen *et al.*, 2007; Huang *et al.*, 2003; Yernool *et al.*, 2004; Abramson *et al.*, 2003 and Guan *et al.*, 2007; Yamashita *et al.*, 2005; Hunte *et al.*, 2005). The structures revealed a surprisingly high diversity, but also showed some similarities on different levels of structural organization.

1.5.1. Structure of transmembrane domains

Common structural feature of secondary transporters are hydrophobic, α -helical domains that traverse the membrane in zig-zag fashion connected by hydrophilic loops. These transmembrane domains (TMs) are oriented to each other in a specific manner to build up the helix packing of the tertiary structure. In the X-ray structures of AcrB, AAC, EmrD, EmrE, GlpT, and LacY, the transmembrane domains display continuous α -helices that span the lipid bilayer and exhibit a shape that can be bended or kinked. The most pronounced helix kinks are thereby observed in odd-numbered TMs of the mitochondrial ATP/ADP carrier with bends up to 35° (Nury *et al.*, 2006) (Fig. 1.3.). Moreover, most TM helices of these six secondary transporters are tilted with respect to the membrane normal. Crystal structures of CLC, Glt_{Ph}, LeuT_{Aa}, and NhaA additionally revealed irregular structural

motifs in form of discontinuous helices (Screpanti and Hunte, 2007). These structures are integrated in the membrane and can be classified into two different types, the transmembrane- and the non-transmembrane discontinuous helices. Examples for the former type were found in LeuT_{Aa} and NhaA in form of two antiparallel transmembrane domains that possess breaks in their α -helical structure at positions approximately in the middle of the lipid bilayer (Hunte *et al.*, 2005; Yamashita *et al.*, 2005). In contrast, the two non-transmembrane discontinuous helices found in CLC as well as in Glt_{Ph} do not span the membrane completely but enter and exit the lipid bilayer on one side of the lipid bilayer shaping hairpin like, helix-turn-helix structures (Dutzler *et al.*, 2002; Yernool *et al.*, 2004). Irrespective of the ability of a discontinuous helix to traverse the membrane or not, it introduces partially charged helix termini and unwound regions in the core of the lipid bilayer, elements that seem to be tightly connected to transporter function and dynamics. Thereby, the partial charges of the helix ends and polar backbone groups of the unwound regions are not fully saturated with internal hydrogen-bonds. These elements can be exploited for attraction and binding of ions and substrates, since they provide acceptor and donor groups for hydrogen bonds as well as an appropriate electrostatic environment. In addition, discontinuous helices are more flexible than continuous α -helices and may be associated with a hinge to allow flexibility of these structural elements during conformational changes of the transporters (Screpanti and Hunte, 2007).

1.5.2. Transmembrane helix packing and structural symmetry

Within the ten crystallized transporters, three shared almost similar 3D structures, whereas the remaining seven revealed clearly different folds. The former three transporters, namely LacY, GltP, and EmrD of *E. coli*, show structural similarities despite of differences in the primary sequence. Since all of them belong to the MFS family this suggests that all members of this superfamily possess a common fold. The high-resolution structures of LacY, GltP, and EmrD reveal a topology of 12 transmembrane domains subdivided into two domains of six helices (Fig. 1.3.). The N- and C-terminal domains show a similar topology and are related by a pseudo two-fold symmetry axis perpendicular to the membrane plane. This structural symmetry between the domains is in agreement with the suggestion that many transporters arose by gene duplication events, although there is low sequence homology between them (Saier, 2003). In accordance with the rocker-type alternating access mechanism, the two six-helix domains of LacY and GltP form an internal hydrophilic cavity that is open to the intracellular side of the membrane and allows the access to the substrate-binding site from the cytoplasm (Abramson *et al.*, 2003; Guan *et al.*, 2007; Huang *et al.*, 2003). Differing from this inward-facing conformation (C_i) of LacY and GltP, the structure of

the multidrug transporter EmrD does not show this cytoplasmic cleft and, therefore, probably represents an occluded state between the C_o and the C_i conformation (Yin *et al.*, 2006). However, a 3D structure of an MFS transporter in the C_o conformation is not available yet. Further structures of secondary transporters with parallel orientated repeat domains were obtained for the multidrug transporter AcrB of the resistance-nodulation-cell division (RND) superfamily (TC 2.A.6.) and AAC of the mitochondrial carrier (MC) family (TC 2.A.29.). Similar to MFS transporters, the transmembrane part of AcrB is comprised of 12 helices with a two-fold symmetry between the N- and C-terminal six-helix bundles. However, the helix packing is considerably different with respect to helix arrangement, tilt, and length, indicating a different genetic origin (Murakami *et al.*, 2002). Characteristics of AcrB are the two periplasmic loops, one per domain, that are significantly larger than those of other transporters crystallized so far (Fig. 1.3.). In contrast to the 12-transmembrane proteins described above, AAC possesses a structure of six TMs consisting of three repeats containing each two TMs around a pseudo-threefold axis (Pebay-Peyroula *et al.*, 2003) (Fig. 1.3.). These helices enclose a hydrophilic cavity that is open towards the outer side of the membrane corresponding to a C_o conformational state of the crystallized antiporter.

Internal repeats with opposite orientations in the membrane were found in the X-ray structures of CLC, LeuT_{Aa}, NhaA, and Glt_{Ph} (Dutzler *et al.*, 2002; Yamashita *et al.*, 2005, Hunte *et al.*, 2005; Yernool *et al.*, 2004) (Fig. 1.3.). For the H⁺/Cl⁻ exchange transporter CLC of the chloride carrier/channel (CLC) family (TC 2.A.49.) an inverted internal repeat was identified that comprises the whole molecule. The N- and C-terminal halves have similar structure and are related by a pseudo two-fold axis in the plane of the membrane. The overall structure of CLC exhibits a complex topology of 18 α -helices that encompass two aqueous vestibules on either side of the lipid bilayer. In the X-ray structures of LeuT_{Aa}, NhaA, and Glt_{Ph}, the structural symmetry involves only part of the transporter molecules. LeuT_{Aa} is organized in two unanticipated inverted repeat domains that comprise only the first ten TMs of the total 12 TMs containing protein (Fig. 1.3.). This core region forms a hydrophilic cavity, which is orientated towards the outer side of the membrane as it is supposed for a transporter in C_o conformation. Similar to LeuT_{Aa}, the crystal structure of NhaA revealed an internal structural repeat that includes only part of the protein (Fig. 1.3.). The 12 TMs of NhaA are organized in two domains with different structures. The first domain is formed by an elongated assembly of six helices without internal symmetry, whereas the second domain contains two structurally related bundles of three TMs with opposite orientation. At the domain interface two funnels with openings either to the inner or outer side of the membrane exist that are both blocked in the middle of the lipid bilayer (Hunte *et al.*, 2005). The smallest inverted repeats were found in the X-ray structure of Glt_{Ph} (Yernool *et al.*, 2004). Here, only the two hairpins (HP1 and HP2) are structurally related and orientated by a pseudo two-fold

symmetry axis parallel to the membrane plane (Fig. 1.3.). For the residual eight transmembrane domains of Glt_{Ph} no internal symmetry could be observed. The only crystallized secondary transporter that does not show any structural homologous domains in the protomer structure is the multidrug transporter EmrE of the drug/metabolite transporter (DMT) superfamily (TC 2.A.7.) (Chen *et al.*, 2007). It is one of the smallest transporter proteins with only 110 amino acids and four TMs (structure not shown).

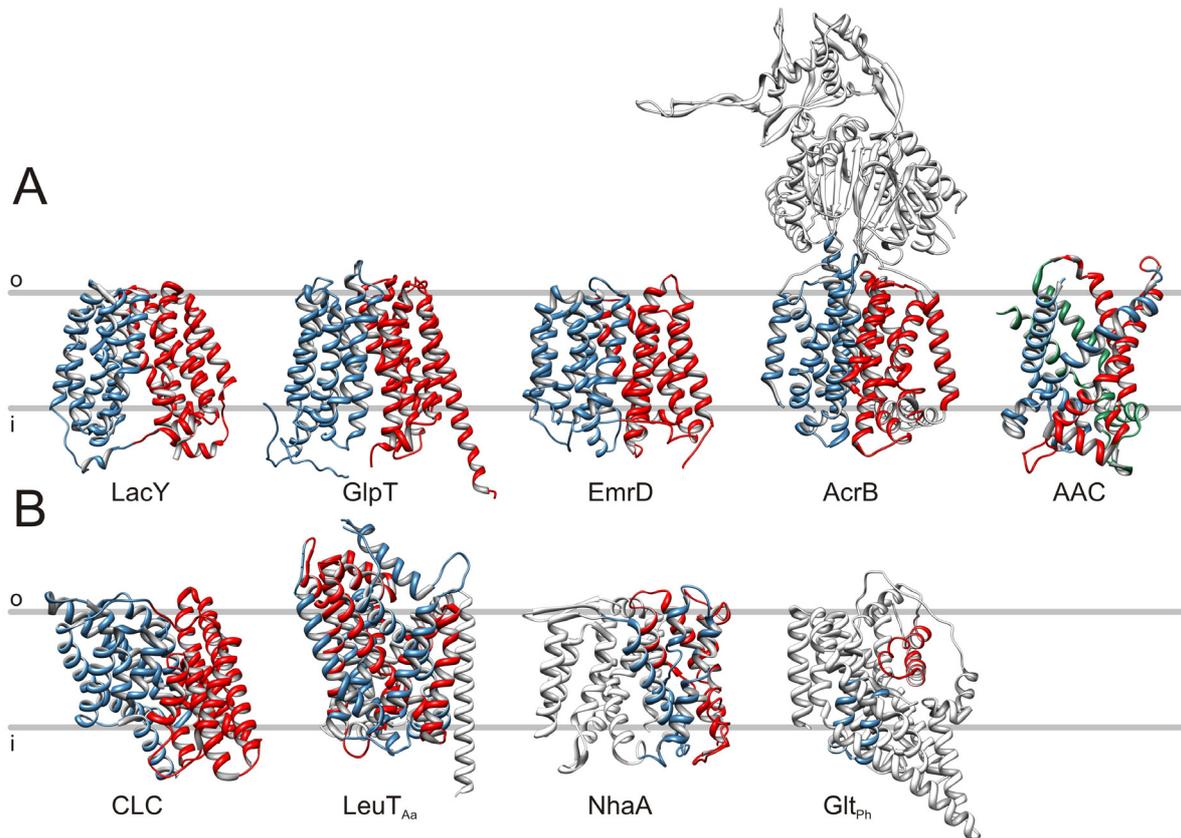


Figure 1.3.: Ribbon representation of known X-ray structures for secondary transporter containing internal structural repeats. The structures are viewed parallel to the membrane. In case of oligomeric transporters, only the structure of one monomer is presented. The repeat domains are colored in blue, red, and green, while unpaired transmembrane and extramembrane segments are colored in white. (A) Secondary transporters with parallel orientated repeat domains. (B) Secondary transporters with inverted repeats. o: outer side of the membrane; i: inner side of the membrane side. Following PDB accession codes (Berman *et al.*, 2007) were used for the generation of the ribbon models: 1PV7 (LacY), 1PW4 (GlpT), 2GFP (EmrD), 1IWG (AcrB), 1OKC (AAC), 1KPL (CLC), 2A65 (LeuT_{Aa}), 1ZCD (NhaA), and 1XFH (Glt_{Ph}). Structural representations were created using UCSF chimera (Pettersen *et al.*, 2004).

1.5.3. Quaternary structure

Protein oligomerization is a common structural feature of many secondary transporters and may fulfill several functions, including modulation of transporter activity, formation of one functional unit, stabilization of distinct conformations, and formation of aqueous “waiting areas” for transported solutes (Hillar *et al.*, 2005; Chen *et al.*, 2007; Kavanaugh, 2004, Veenhoff *et al.*, 2002). The ten X-ray structures of secondary transporters obtained to date revealed four different quaternary states, including monomers, dimers, and trimers. For the MFS transporters EmrD, GlpT, and LacY no physiological relevant

association of protein molecules was observed in the crystals, suggesting that these transporters function as monomers (Yin *et al.*, 2006; Huang *et al.*, 2003; Abramson *et al.*, 2003). The monomeric structure of LacY and GlpT was additionally confirmed by different complementary analyses that address the quaternary structure of these proteins in a non-crystalline environment (Veenhoff *et al.*, 2002; Auer *et al.*, 2001).

Other secondary transporters appear to operate as dimers, like AAC, CLC, EmrE, LeuT_{Aa}, and NhaA. However, with the exception of EmrE, each protomer contains a ligand-binding site and is most likely capable to form a translocation pathway, implicating that each monomer constitutes a functional unit. CLC and LeuT_{Aa} form homodimers in the crystal with a two-fold axis perpendicular to the membrane (Dutzler *et al.*, 2002; Yamashita *et al.*, 2005). The dimer interface between two parallel oriented protomers of CLC is very extensive, suggesting that the protein exists and functions only as a dimer (Dutzler, 2006). Thereby, the intermolecular contact comprises several intermolecular interactions between a number of transmembrane helices. In the case of LeuT_{Aa}, the dimer interface area between the subunits is composed of one α -helix-containing loop that is also involved in the formation of the hydrophilic cleft and two helices of the C-terminal part of this transporter. Oligomerization was also reported for eukaryotic relatives of LeuT_{Aa}. Here, a dimer of dimers may be formed under physiological conditions (Sitte *et al.*, 2004; Just *et al.*, 2004). Although AAC and NhaA were proposed to be organized as dimers in the native membrane (Huang *et al.*, 2001; Gerchman *et al.*, 2001), both of them were not associated in a physiological dimeric state in the crystals (Pebay-Peyroula *et al.*, 2003; Hunte *et al.*, 2005). This is thought to be a result of the harsh conditions during purification and crystallization of the transport proteins that may lead to excessive delipidation and, hence, to dimer dissociation if the interactions between the protomers are only weak (Pebay-Peyroula *et al.*, 2003). In a second crystal structure of AAC, parallel homodimers were discerned in which cardiolipins are tightly bound at the putative dimerization interface (Nury *et al.*, 2005). The lipids strongly expand the interacting surface of the dimer and are thus presumably essential for mediating the monomer contact. In contrast to the parallel dimers described above, EmrE crystallized as an antiparallel, dual topology homodimer with an internal pseudo-twofold symmetry axis in the plane of the membrane (Chen *et al.*, 2007). The three N-terminal helices from each monomer form together one hydrophobic substrate-binding chamber, suggesting that EmrE is an obligate dimer that displays one functional unit. The dimer interface consists of the two latter TMs of EmrE that compose, together with their symmetry partners, a four-helix bundle. However, even the dual topology is also supported by a genetic approach (Rapp *et al.*, 2007), biochemical data suggested a parallel orientation of the EmrE monomers around a two-fold symmetry axis perpendicular to the membrane (Schuldiner, 2007). Therefore, the topology is

still discussed controversially and further experiments are needed to exclude one of the models with reasonable certainty.

In the crystal structures of AcrB and Glt_{Ph}, both transporters were found to exist in a trimeric organization with a triangular shape and a three-fold symmetry axis perpendicular to the membrane (Murakami *et al.*, 2002; Yernool *et al.*, 2004). Thereby, the three protomers of Glt_{Ph} form a large extracellular basin that protrudes from the membrane up to the middle of the bilayer to allow the access to the substrate-binding site in each monomer. The interface between individual Glt_{Ph} molecules in the trimer is composed of three TMs of the N-terminal helix bundle. The physiological significance of the trimer formation and the intersubunit contacts was proven by cross-linking analysis of Glt_{Ph} in the native membrane (Yernool *et al.*, 2004). Among the crystallized secondary transporters, AcrB exhibits the most complex quaternary structure. The AcrB homotrimer consists of two major parts, a membrane integral domain and a periplasmic domain that are formed by the 12 TMs and the large periplasmic loops of each protomer, respectively. In the transmembrane part of the trimer the three symmetrically arranged 12-TMs helix bundles encompass a central cavity, which is proposed to be filled with phospholipids (Murakami *et al.*, 2002). The respective contact sites between the individual subunits are mediated by two transmembrane α -helices. The periplasmic part of AcrB is more densely packed than the membrane integral domain and comprises stronger interactions between the adjacent molecules. It exhibits a central pore that connects a cavity near the membrane plane with a funnel-like structure at the top of the complex. This central pore is proposed to form the translocation pathway for substrates that first enters the hydrophobic binding sites in each protomer via three vestibules at the monomer interfaces of the extra-membrane domain. In other crystals of AcrB, asymmetric trimers were obtained in which each monomer had a different conformation corresponding to the consecutive states of a three-step transport cycle (Seeger *et al.*, 2006; Murakami *et al.*, 2006).

1.6. Site-directed spin labeling and EPR spectroscopy on membrane proteins

The 3D X-ray structures of secondary transporters provided new insights into the architecture and function of these proteins. However, despite the high diversity and physiological importance of secondary transporters, the number of atomic structures determined to date is very small. This is largely due to the experimental challenges that were encountered with membrane proteins using classical techniques, like high-resolution NMR and X-ray crystallography (Lacapère *et al.*, 2007). Most polytopic, α -helical membrane proteins are too large for current nuclear magnetic resonance (NMR) approaches and, based on their hydrophobicity and conformational flexibility, often very difficult to crystallize. New techniques for their structural characterization are thus required. The electron paramagnetic

resonance (EPR) spectroscopy displays a powerful method for exploring the structure and dynamics of membrane proteins and their complexes without crystallization under conditions relevant to function. It is applicable to proteins in solution, detergent micelles, and liposomes with no significant limitation to the size of the protein (Fanucci and Cafiso, 2006; Bordignon and Steinhoff, 2007; Jeschke and Polyhach; 2007). EPR is a spectroscopic technique that measures the resonant absorption of microwave radiation by unpaired electrons in a strong magnetic field. However, with the exception of radical enzymes (e.g., metalloproteases), the vast majority of proteins do not possess paramagnetic centers and are not accessible to EPR analyses without introduction of stable radicals by site-directed spin labeling (Hubbell *et al.*, 2000). This procedure includes a cysteine (Cys) substitution mutagenesis of native amino acids at desired positions, followed by modification of the sulfhydryl group via a specific paramagnetic nitroxide spin label, which contains an unpaired electron located in the nitroxide N-O bond (Fig. 1.4.). The continuous-wave (CW) EPR spectrum of protein immobilized spin label yields structural information based on nitroxide side chain mobility (Columbus and Hubbell, 2002), its accessibility to collisions with polar or apolar paramagnetic quenchers (Altenbach *et al.*, 1994), the polarity of its immediate surrounding, or distances between the nitroxide and a second paramagnetic center in the protein (Altenbach *et al.* 2001). Sequential analysis of a series of spin-labeled membrane protein variants with respect to these different SDSL techniques allows the detection of secondary structure elements (Hubbell *et al.*, 1998), the investigation of protein topology (Wegener *et al.*, 2000) and topography (Steinhoff *et al.*, 2000), and characterization of tertiary and quaternary structures at the backbone level (Perozo *et al.*, 1998; Wegener *et al.*, 2001; Banham *et al.*, 2006). Furthermore, time-resolved measurements of the parameters, described above, can follow conformational changes with up to 100 μ s time resolution (Rink *et al.*, 2000).

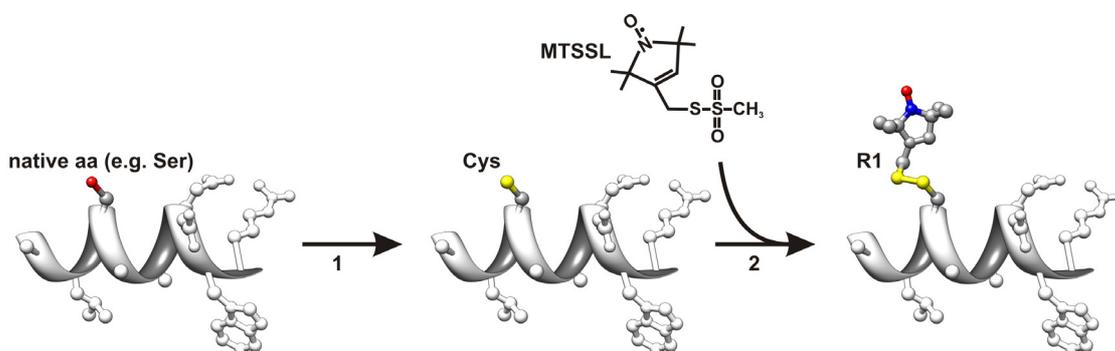


Fig. 1.4.: The general strategy of site-directed spin labeling (SDSL) of proteins. Native amino acids (aa) at desired sites are replaced by cysteine (Cys) residues via site-directed mutagenesis (1) followed by modification of the sulfhydryl group with a specific methanethiosulfonate spin label (MTSSL) (2) to generate the nitroxide side chain, designated R1. The protein backbone is depicted in ribbon representation and amino acid residues are shown in ball-and-stick representation. Hydrogen atoms are not shown for clearness. Only the side chains of the substituted and modified amino acids are colored by atom type (grey: carbon, red: oxygen, yellow: sulfur, blue: nitrogen). The figure was prepared with the program UCSF chimera (Pettersen *et al.*, 2004).

The determination of interspin distances in proteins or protein complexes with two or more attached spin labels is the most powerful application for structure elucidation. The EPR distance measurements are based on the dipole-dipole coupling between spins that is proportional to the inverse cube of the distance r^{-3} (Jeschke, 2002). In the conventional CW EPR spectrum, the dipolar interaction leads to considerable line broadening that can be analyzed to extract distances in the range between 0.8 nm and 2.0 nm (Rabenstein and Shin, 1995; Steinhoff *et al.*, 1997). Above interspin distances of 2 nm CW EPR is not applicable anymore to separate the dipolar broadening contribution from the natural line width in the EPR spectrum. Pulsed EPR techniques such as double electron-electron resonance (DEER) (Milov *et al.*, 1984; Pannier *et al.*, 2000) and double quantum coherence (DQC) (Borbat and Freed, 1999; Borbat *et al.*, 2004) enable the determination of small dipolar interactions at distances up to 6 nm or, in favorable cases, up to 8 nm and thus considerably enhance the distance sensitivity of EPR spectroscopy to a range that matches the dimensions of proteins and their complexes. In particular, the four-pulse DEER experiment was shown to be applicable to membrane proteins like the light-harvesting chlorophyll a/b complex (LHCIIb) (Jeschke *et al.*, 2005), PutP (Jeschke *et al.*, 2004a), LacY (Smirnova *et al.*, 2007), the ABC transporter MsbA (Borbat *et al.*, 2007), rhodopsin (Altenbach *et al.*, 2008), and the voltage-dependent potassium channel KvAP (Vamvouka *et al.*, 2008). Beside the determination of accurate mean spin-to-spin distances, the four-pulse DEER experiment is also capable of providing information on the distance distribution between spins (Pannier *et al.*, 2000; Jeschke *et al.*, 2004b), which is important for characterization of structures with high inherent flexibility, as it is expected for many transporters whose function demands large-amplitude conformational motions (Smirnova *et al.*, 2007; Borbat *et al.*, 2007; Jeschke *et al.*, 2004a). However, interpretation of distance distributions in terms of the structure of the analyzed protein is complicated, since the distance distribution is measured between the unpaired electrons in the N-O bond of the nitroxide group rather than between the backbone carbons to which the labels are attached. For this reason, the size and the conformational dynamics of the spin label side chain have to be considered to avoid errors of up to 1.2 nm if the interspin distance is interpreted as a C $^{\alpha}$ -C $^{\alpha}$ distance (Borbat *et al.*, 2002). The recently established semidynamic rotamer library approach has been proven as an efficient means for predicting the tether length and conformational distribution of the label within the specific protein environment (Jeschke and Polyhach, 2007). In this approach, the conformational dynamics of the isolated labeled residue are represented by a rotamer library that contains a discrete set of the possible conformations of this modified side chain. The energies of these rotamers are then computed by introducing them into a static potential field, which is considered to be created by a rigid model of the entire protein or at least of the immediate vicinity of the labeled residue. Finally,

the energy of each particular rotamer is used to obtain its probability to occur at the selected position in the protein structure (Jeschke and Polyhach, 2007). In comparison to other molecular modeling approaches such as Monte Carlo simulations (Sale *et al.*, 2005) and molecular dynamics (Borovykh *et al.*, 2006), this strategy is significantly faster and needs less computational power. It is therefore an appropriate method for optimizing a protein structural model constrained by experimental distance distributions.

In the present work, CW EPR techniques and the four-pulse DEER experiment were used for characterization of two ion-dependent secondary transporters, PutP and NhaA, at different levels of structural organization. In cases where reliable interpretation of label-to-label distance distributions was important for structural modeling of the labeled protein, the rotamer library approach was used to compute the conformational distribution of the spin labels for overcoming the intrinsic imprecision of SDSL EPR.

1.7. The Na⁺/proline symporter PutP of *E. coli*

PutP belongs to the solute:sodium symporter (SSS) family (TC 2.A.21), which contains several hundred members of prokaryotic and eukaryotic origin (Jung, 2002; Wright and Turk, 2004; Reizer *et al.*, 1994). Proteins of this family utilize a *smf* to drive uphill transport of substrates such as sugars, amino acids, vitamins, iodide, *myo*-inositol, phenyl acetate, and urea (Dohán *et al.* 2006; Jung 2001; Wright *et al.* 2007). Most of the functionally characterized transporters feed catabolic pathways (e.g. PutP, SGLT, and Ppa) or are involved in cell adaptation to osmotic stress (e.g. OpuE) (Jung, 2001; Spiegelhalter and Bremer, 1998). Among the eukaryotic members of the SSS family the Na⁺/glucose transporter (SGLT1) and Na⁺/iodide symporter (NIS) are implicated in the human diseases glucose-galactose malabsorption and iodide transport defect (Wright *et al.*, 2007; Reed-Tsur *et al.*, 2008). Both proteins also play an important role in medical therapy (Wright *et al.*, 2007; Dohán *et al.*, 2006). Furthermore, bacterial transporters such as PutP of *Helicobacter pylori* and *Staphylococcus aureus* contribute to bacterial virulence and thus represent putative targets for the development of new drugs against pathogens (Kavermann *et al.*, 2003; Schwan *et al.*, 2006).

Topology and functional properties of PutP - PutP of *E. coli* is one of the best characterized prokaryotic members of the SSS family. It is composed of 502 amino acid residues and has a molecular mass of 54.3 kDa. Gene fusion analyses, Cys accessibility studies, site-specific proteolysis, and SDSL EPR measurements revealed a secondary structure model according to which PutP contains 13 TMs with the N terminus located on the periplasmic side of the membrane and the C terminus facing the cytoplasm (Jung *et al.*, 1998; Wegener *et al.*, 2000) (Fig. 1.5.). The transporter catalyzes the coupled translocation

of Na⁺ and proline with a stoichiometry of 1:1 (Yamato and Anraku, 1993). The apparent K_m values for Na⁺ and proline uptake were determined with 30 and 2 μM, respectively (Yamato and Anraku, 1993). Li⁺ can also be used as coupling ion for proline transport, although the affinity of the transporter for Li⁺ ($K_m \sim 125$ μM) is significantly lower than for Na⁺ (Chen *et al.*, 1985). Kinetic analyses of Na⁺/proline transport catalyzed by PutP suggested that transport occurs according to an ordered binding mechanism (Yamato and Anraku, 1990; Yamato, 1992). Thereby, the initial binding of Na⁺ is proposed to induce a conformational alteration that increases the affinity of the transporter for proline. The ternary complex assumedly reorientates in the membrane, leading to release of Na⁺ and proline on the other side of the lipid bilayer. At high substrate concentrations, proline can also bind to the protein in the absence of Na⁺, as it was shown by electrophysiological measurements using the rapid solution exchange technique combined with a solid-supported membrane (SSM) (Zhou *et al.*, 2004).

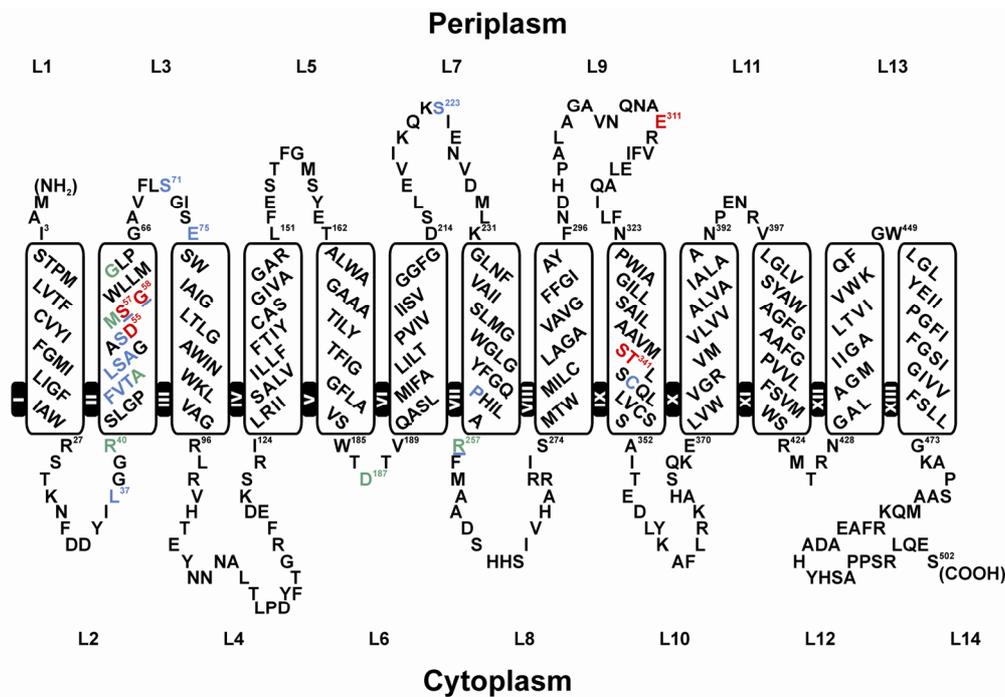


Fig. 1.5.: Secondary structure model of PutP of *E. coli* highlighting important residues. The model is based on a gene fusion approach, Cys accessibility analyses, SDSL EPR, and site-specific proteolysis (Jung *et al.*, 1998; Wegener *et al.*, 2000). Putative TMs are represented as rounded rectangles and numbered with Roman numerals; loops are numbered with Arabic numerals starting from the N terminus. The single-letter amino acid code is used and the location in the amino acid sequence is indicated by a number. Residues proposed to be directly involved in ligand binding are shown in red. Amino acids proposed to be involved in ligand-induced conformational alterations based on Cys accessibility analyses, fluorescence, and EPR measurements are highlighted or underlined in blue. Other residues of structural and/or functional importance are represented in green.

Functional important residues of PutP - Site-directed mutagenesis studies identified several amino acids in PutP that are crucial for transport activity (Fig. 1.5.). Most of the found residues are located in TM II of PutP, suggesting that this segment is of particular functional importance. Thus, the carboxylate of Asp55 in TM II was proved to be essential for transport, and significant, albeit highly reduced activity was detected only with Glu at this position (Quick and Jung, 1997). The latter substitution caused an about 50-fold decrease of the

apparent affinity of the protein for Na^+ compared to the wild-type transporter. On the other hand, only a relatively small alteration of the apparent affinity for proline was observed, suggesting that Asp55 is at or close to the cation-binding site (Quick and Jung, 1997). Met56 was also identified as a residue which is crucial for high-affinity ion binding (Pirch *et al.*, 2002). However, in contrast to Asp55 it is not essential for transport and is assumed to be not involved in direct binding of the coupling ion (Pirch *et al.*, 2002). In addition to Asp55 and Met56, individual replacements of Ser57 and Gly58 in TM II led to significantly altered transport kinetics. Substitutions of Ser57 by Ala, Cys, Gly, or Thr caused a reduction of the apparent Na^+ and proline affinities by up to two orders of magnitude with little influence on V_{max} values (Quick *et al.*, 1996). Similarly, Cys substitution of Gly58 decreased the apparent affinity of PutP for Na^+ and proline although the effect was less dramatic as in case of Ser57 (Pirch *et al.*, 2002). The influence of these replacements on Na^+ and proline affinities supports the idea of a close cooperativity between the cation- and substrate-binding sites in PutP. Furthermore, Ala48 and Gly63 of TM II were found to be crucial for transport function, since replacement of these conserved residues by Cys significantly impaired proline uptake (Pirch *et al.*, 2003). Analysis of the accessibility of single Cys individually placed at different positions in TM II to *N*-ethylmaleimide (NEM) and fluorescein-5-maleimide indicated a participation of the TM in the formation of a hydrophilic cleft open to the cytoplasmic side of the membrane (Pirch *et al.*, 2003). Together with the transport analyses, the data support the conclusion that residues of TM II of PutP contribute to the formation of an ion- and/or substrate-translocation pathway.

Besides amino acids of TM II, residues of TM IX were shown to be of functional importance. In particular, experiments on Ser340 and Thr341 of TM IX demonstrated that these residues are crucial for proline transport (Böhm and Jung, unpublished information). This is supported by the fact that Ser340 as well as Thr341 are conserved within the SSS family. Replacement of the corresponding amino acids in NIS (Ser353 and Thr354) revealed that the hydroxyl groups are involved in Na^+ binding and/or translocation (De la Vieja *et al.*, 2007). Furthermore, a similar amino acid arrangement was found in the X-ray structure of the leucine transporter LeuT_{Aa} of the NSS family in which two adjacent polar residues (Thr354, Ser355) were shown to participate in Na^+ binding (Yamashita *et al.*, 2005). Combined with the observed accessibility of the native Cys residue at position 344 in TM IX of PutP to NEM (Yamato and Anraku, 1988; Hanada *et al.*, 1992), this suggests that TM IX forms part of a ligand-translocation pathway similar to TM II of PutP.

In addition to the functionally important residues identified in transmembrane domains of PutP, four conserved charged residues (Arg40, Asp187, Arg257, and Glu311) in loop regions of PutP were identified as being relevant for the transport process (Quick *et al.*, 1999; Quick and Jung, 1998; Ohsawa *et al.*, 1988; Böhm and Jung unpublished information)

(Fig. 1.5.). Removal of the charged side chain at the individual positions led to a reduced Na^+ dependence of proline binding. In the case of Arg40 this was accompanied by decreased apparent affinities of the transporter for Na^+ while the apparent affinity for proline was only slightly altered. Therefore, it is proposed that Arg40 is located close to the site of ion binding and is important for the coupling of ion and proline transport (Quick *et al.*, 1999). Replacement of Asp187 with Cys led to high-affinity proline binding even at very low Na^+ concentrations, which was attributed to an enhanced Na^+ affinity of the transporter (Quick and Jung, 1998). The same was expected for Arg257. However, the effect of the Cys substitution of this residue on proline uptake was significantly lower than that for the alteration of Asp187 (Ohsawa *et al.*, 1988). These results indicate that Asp187 as well as Arg257 are located close to the pathway of the coupling ion through the membrane and are probably involved in the release of Na^+ to the cytoplasmic side of the membrane. Finally, the PutP derivative PutP-E311C showed biphasic kinetics with one component manifesting the kinetic parameters of active transport (i.e. high apparent affinity for proline comparable to the wild-type) and the other exhibiting the characteristics of facilitated diffusion (i.e. low apparent affinity for proline) as it is typically for an uncoupled transport process (Böhm and Jung, unpublished information). Thereby, the binding of the proline was shown to be independent of the presence of Na^+ , suggesting that this residue is important for Na^+ binding and/or coupling of ion and proline transport.

Conformational dynamics of PutP - A series of protein chemical and spectroscopic studies was employed to explore Na^+ - and/or proline-induced conformational alterations at specific sites in PutP (Wegener *et al.*, 2000; Pirch *et al.*, 2002; Pirch *et al.*, 2003; Pirch and Jung, unpublished information). SDSL of Cys introduced at position of Leu37 or Phe45 in loop 2 and TM II of the transporter and analysis of the corresponding EPR line shapes revealed ligand-induced mobility changes of the attached nitroxide (Wegener *et al.*, 2000) (Fig. 1.5.). In the case of PutP-L37R1 (R1 designates the modified nitroxide side chain, see above), binding of Na^+ and/or proline to the transporter led to an immobilization of the spin label side chain while the nitroxide at position 45 becomes more mobile upon addition of proline, and Na^+ alone had no effect. From these findings it was concluded that proline binding induces a conformational alteration of PutP that involves at least parts of TM II and the preceding cytoplasmic loop. Na^+ could only be shown to affect the structure of loop 2 (Wegener *et al.*, 2000). These results are further supported by DEER distance measurements between spin labels attached to positions 37 (loop 2) and 187 (loop 6). The measurements revealed an increase of the mean interspin distance between the nitroxides at these positions upon Na^+ binding, suggesting that one or both positions are involved in Na^+ -induced structural alterations (Jeschke *et al.*, 2004a). The EPR studies were confirmed and extended by Cys accessibility analyses in the presence and absence of ligands. Thus, Na^+

was shown to increase the accessibility of Cys placed at the position of Ser57, Gly58 in TM II and Ser71 or Glu75 in loop 3 (Pirch *et al.*, 2003; Pirch and Jung, unpublished information) (Fig. 1.5.). In contrast to Ser57 and Gly58, proline did not affect the accessibility of Cys at positions 71 and 75. Furthermore, proline inhibited the reaction of sulfhydryl-specific reagents with Cys at positions in TM II (Phe45, Val46, Thr47, Leu49, Ser50, Ala51, Ser54, Gly63, and Leu64), TM VII (Pro252), Loop 8 (Arg257), and TM IX (Cys344) (Pirch *et al.*, 2003; Pirch and Jung, unpublished information). Since substitution of the native amino acids at most of these positions has no or only very little effect on PutP transport kinetics, the altered Cys accessibility can be attributed to conformational alteration of the protein upon ligand binding and not to direct steric hindering as it is discussed for Ser57 and Gly58 (Pirch *et al.*, 2002; Pirch *et al.*, 2003). In addition to these results, fluorescence of PutP site-specifically labeled with a fluorescence group at position 223 (loop7) was altered by addition of proline if Na⁺ was present (Pirch and Jung, unpublished information). In conclusion, the studies suggest that at least parts of TM II, TM VII, and TM IX and loops 2, 3, 7, and 8 are involved in structural alterations induced by Na⁺ and/or proline, which correlate well with the ordered binding model of Na⁺/proline transport.

Tertiary structural information - To understand the transport mechanism of PutP at the molecular level, knowledge of the structure is indispensable. However, the only information on the tertiary structure of PutP available so far was gained by DEER distance measurements between spin labels attached to positions in loops 2, 4, 6, and 7 (Jeschke *et al.*, 2004a). The results revealed that the proposed cytoplasmic loops 2 and 6 and the loops 2 and 4 are in close proximity (~2 nm) to each other, whereas the cytoplasmic loop 4 and the periplasmic loop 7 are separated by a distance of 4.8 nm. This DEER data strongly supports the idea of loops 2, 4, and 6 being located at the same side and loops 4 and 7 at opposite sides of the membrane, which is in good agreement with the proposed secondary structure model of PutP (Jung *et al.*, 1998; Wegener *et al.*, 2000) (Fig. 1.5.).

1.8. The Na⁺/H⁺ antiporter NhaA of *E. coli*

Na⁺/H⁺ antiporters are found ubiquitously in all three kingdoms of life. They are involved in cell energetics and play primary roles in the homeostasis of intracellular pH, cellular Na⁺ content, and cell volume, which are processes crucial for cell viability (Padan *et al.*, 2005). In humans, the antiporters participate in many pathological conditions of heart, kidney, and stomach (Engelhardt *et al.*, 2002; Myers *et al.*, 2003). For example, NHE1 of the monovalent cation/proton antiporter-1 (CPA1) family (2.A.36) has a role in heart hypertrophy and in the damage that occurs during ischemia and reperfusion (Orlowski and Grinstein, 2004). In plants, deletion of genes encoding the vacuolar and cytoplasmic membrane Na⁺/H⁺

antiporters generates salt-sensitive plants (Apse and Blumwald, 2007), whereas overexpression of these genes is exploited to engineer salt-resistant plants (Yamaguchi and Blumwald, 2005).

Functional properties of NhaA - *E. coli* has four cation/proton antiporters (NhaA, NhaB, ChaA, and MdfA) of which NhaA is the only one that is indispensable for the adaptation to high salinity, resistance to Li⁺ ions, and growth at alkaline pH in the presence of Na⁺ (Padan *et al.*, 2005). NhaA of *E. coli* is the prototype of the NhaA Na⁺:H⁺ antiporter (NhaA) family (2.A.33) that contains homologs from numerous bacteria and archaea including important pathogenic bacteria such as *H. pylori*, *Vibrio cholera*, *V. parahaemolyticus*, and *V. alginolyticus* (Padan *et al.*, 2001). NhaA is an electrogenic antiporter with a stoichiometry of 2H⁺/Na⁺ (Taglicht *et al.*, 1993) and apparent affinities for Na⁺ and Li⁺ of 0.2 and 0.02 mM, respectively (Galili *et al.*, 2004). The activity of the transporter is strictly regulated by pH, a property it shares with many other prokaryotic and eukaryotic antiporters and which is essential for cytoplasmic pH regulation (Padan *et al.*, 2004; Taglicht *et al.*, 1991). It is downregulated below pH 6.5 and exhibits an activity increase by over three-orders of magnitude upon shift to alkaline pH with maximum activity at pH 8.5 (Taglicht *et al.*, 1991).

Structure of NhaA - NhaA consists of 388 amino acids and has a molecular mass of 41.4 kDa. Gene fusion analyses, antibody epitope mapping, accessibility studies, and site-specific proteolysis revealed a secondary structure model according to which NhaA is composed of 12 TMs with N and C termini exposed to the cytoplasm (Olami *et al.*, 1997; Rothman *et al.*, 1996) (Fig. 1.6.a). This topology model was confirmed by the low-resolution structure obtained by cryo-electron microscopy (cryo-EM) on two-dimensional (2D) crystals (Williams *et al.*, 1999; Williams, 2000). In the crystal, NhaA formed homodimers with each monomer containing 12 tilted transmembrane helices. Cross-linking analyses, copurification experiments, and functional complementation studies showed that NhaA also exists and functions as a dimer in the native membrane (Gerchman *et al.*, 2001). Intermolecular cross-linking between Cys residues introduced at positions 254 at the cytoplasmic membrane boundary of TM IX caused a dramatic change in the pH response for a short and rigid linker and no effect for a long and flexible linker (Gerchman *et al.*, 2001). Furthermore, the functional complementation was obtained only between conditional lethal NhaA derivatives, which are either inactive (NhaA-H225R) or overactive (NhaA-G338S) under the nonpermissive conditions (alkaline pH in the presence of Na⁺), but not with fully inactive derivatives (NhaA-D163C and -D164C) (Gerchman *et al.*, 2001). This suggests that the functional interaction between NhaA molecules is important for the pH response of NhaA rather than for its transport activity.

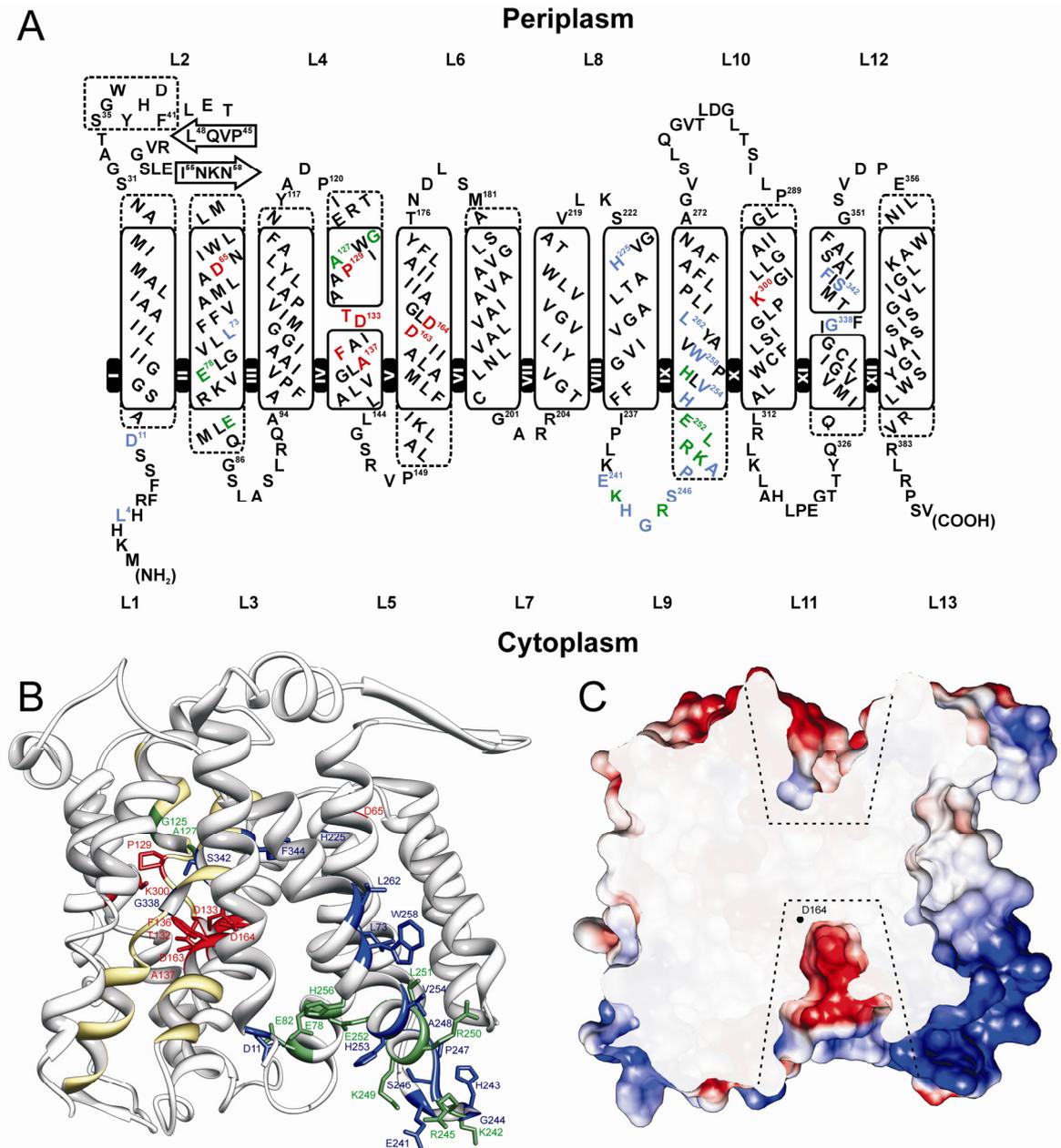


Fig. 1.6.: General architecture of NhaA of *E. coli*. (A) Secondary structure model of NhaA highlighting important residues. The model is based on the 3D structure (Hunte *et al.*, 2005). TMs are represented as rounded rectangles and numbered with Roman numerals; loops are numbered with Arabic numerals starting from the N terminus. The membrane-embedded regions of the TMs are designated by a continuous line and the regions of the TMs exposed to the cytoplasm or periplasm are designated by a broken line. The arrow-like boxes show the antiparallel β -sheet element. The single-letter amino acid code is used and the location in the amino acid sequence is indicated by a number. Sites of mutations that affect the K_m for Na^+ and Li^+ without influencing the pH profile are highlighted in red. Sites where mutations affect both the pH profile and the K_m for Na^+ and Li^+ are shown in green. Sites where mutations affect only the pH dependence of NhaA are colored in blue. (B) Ribbon representation of the crystal structure of NhaA (PDB code: 1ZCD) viewed parallel to the membrane. Functionally important residues are shown as stick models and are colored according to the color scheme in (A). The TMs IV/XI assembly is highlighted in yellow and numbered with Roman numerals. The figure was prepared with the program UCSF chimera (Pettersen *et al.*, 2004). (C) Cross-section through the electrostatic potential surface of NhaA with the front part removed along the membrane normal. Surface model and electrostatic potential were calculated with the program UCSF chimera (Pettersen *et al.*, 2004). The polar surfaces are colored blue (positively charged) and red (negatively charged). The cytoplasmic and periplasmic funnels are marked by broken lines. The location of cation binding D164 is indicated by a black dot.

As aforementioned, the 3D X-ray structure of NhaA was determined at 3.45 Å resolution in its acidic-pH downregulated state (pH 4) (Hunte *et al.*, 2005). In line with the previously established topology model, the architecture of the NhaA crystal structure is comprised of 12 TMs connected by extramembrane loops with the N and C termini residing

in the cytoplasm. The longest loop (periplasmic loop 2) contains a short helix (residues 35-41) followed by a two-stranded antiparallel β -sheet (β 1: 45-48; β 2: 55-58) (Fig. 1.6.a, b). Together with the other periplasmic loops it forms a flat and rigid face at the outer boundary of the lipid bilayer. At the intracellular side of the membrane more flexible loops and several TMs protrude into the cytoplasm forming a rough surface. The transmembrane segments are organized in two densely packed domains. As mentioned above, the first domain comprises a roughly linear arrangement of six helices (TMs I, II, VI, VII, VIII, and IX), whereas the second domain contains two internal structural repeats of three TMs each (TMs III, IV and VI; X, XI, and XII) in opposite orientation with respect to the membrane (Fig. 1.3.). All helices of the 3D X-ray structure were shown to superimpose well with the electron densities from cryo-electron microscopy of the 2D crystals, revealing that the helix packing of NhaA in the detergent-solubilized state in the 3D crystals is similar to that in the lipid bilayer of the 2D crystals (Screpanti *et al.*, 2006). However, no physiologically feasible association of two monomers could be observed in the 3D crystals (Hunte *et al.*, 2005).

Sites of ion binding - A characteristic feature of the structural fold of NhaA is the assembly of the two discontinuous transmembrane helices IV and XI, as described in section 1.5.1. (Fig. 1.6.b). These helices are in opposite direction in the membrane and are each composed of two short helices, one located in the cytoplasmic half and one in the periplasmic half of the lipid bilayer (IVc, IVp, and XIc, XIp), connected by extended polypeptide chains. The extended polypeptide chains cross each other so that the short helices of TMs IV and XI combine in an intertwined manner with their partially positively charged N termini and their partially negatively charged C termini facing each other in the middle of the membrane (Hunte *et al.*, 2005; Screpanti and Hunte, 2007). For stabilization of this assembly, the partially positive dipole ends of the short helices are charge-compensated by Asp133 of helix IV, whereas the partially negative dipole ends are compensated by Lys300 of helix X. Removal of the charge at position 133 caused a dramatic reduction of the apparent affinity of NhaA for both Na^+ and Li^+ (Galili *et al.*, 2002). The NhaA derivative NhaA-K300C had no Na^+/H^+ antiport activity, whereas it catalyzed Li^+/H^+ antiport, albeit with very low apparent affinity for Li^+ in the millimolar range (Kozachkov *et al.*, 2007). These results suggest that the delicately electrostatically balanced TMs IV/XI assembly is crucial for ion binding and/or conformational alterations of the transporter. Indeed, in addition to Asp133 and Lys300 a number of residues (Pro129, Thr132, Phe136, Ala137 (TM IV), Asp163, and Asp164 (TM V)) are found to be important for cation translocation that cluster near the intersection of TMs IV and XI (Galili *et al.*, 2002; Galili *et al.*, 2004, Inoue *et al.*, 1995) (Fig. 1.6.b). Although neither Na^+ nor Li^+ could be observed in the structure, four of these residues (Asp163, Asp164 (TM V) and Asp133, Thr132 (TM IV)) were designated to form the putative Na^+ -binding site (Hunte *et al.*, 2005) (Fig. 1.6.b). Similar to Asp133, substitution of Thr132 by

Cys resulted in a dramatic reduction of the apparent affinity for both Na⁺ and Li⁺ (Galili *et al.*, 2002), while replacement of Asp163 and Asp164 by Cys or Asn completely abolished the transport activity of NhaA (Galili *et al.*, 2002; Inoue *et al.* 1995). The functional importance of this cluster for ion binding is further supported by the fact that Asp164 is exposed to a cytoplasmic cavity (Fig. 1.6.c). This cavity, formed by TMs II, IV, V, and IX, is lined by negatively charged residues and is thus well suited for the attraction of cations. It opens to the cytoplasm and continues to the middle of the membrane, where it is blocked at the crossing of the extended chains of the TMs IV/XI assembly (Fig. 1.6.c). Likewise, a shallow negatively charged funnel that is bordered by TMs II, VIII, and XI opens to the periplasm (Fig. 1.6.c). Both hydrophilic cavities points towards each other but are separated by densely packed, non-polar residues forming a periplasmic barrier (Fig. 1.6.c).

pH-sensor of NhaA - Site-directed and random mutagenesis studies identified several amino acid residues that participate in the pH response of NhaA (Galili *et al.*, 2002; Galili *et al.*, 2004; Kozachkov *et al.*, 2007; Padan *et al.*, 2004; Tzuber *et al.*, 2004; Tzuber *et al.*, 2008; Rimon *et al.*, 1998). Two groups of residues were found to constitute the putative pH sensor. The first group includes residues in which alterations influence only the pH-dependence of NhaA activity, while in the second group replacement of the residues alters both the pH response and the apparent affinity of the transporter (Fig. 1.6.a, b). Most of the residues affecting the pH response cluster on the N-terminal half of TMs IX and the adjoining loop 9, as well as on TM XI and the extended chain of TM XI (Hunte *et al.*, 2005) (Fig. 1.6.a, b). Loop 9 and the N terminus of TM IX are thought to constitute part of the pH sensor at the orifice of the cytoplasmic cavity (Hunte *et al.*, 2005). One remarkable residue found in this region is Glu252. Replacement of Glu252 with Cys causes a drastic alkaline shift of the pH profile of the NhaA activity by one pH unit and decreases the apparent affinity of the antiporter for Na⁺ by factor 50 (Tzuber *et al.*, 2004). In addition to this residue a number of charged amino acid residues were identified that cluster, together with Glu252, at the cavity entrance and are able to change their protonation states, which may lead to conformational alterations and activation of NhaA (Hunte *et al.*, 2005; Olkhova *et al.*, 2006). Using the multiconformation continuum electrostatics (MCCE) method, the effect of pH values on the protonation state of ionisable residues in NhaA was investigated *in silico* (Olkhova *et al.*, 2006). The results revealed four clusters of strongly electrostatically interacting residues, which were spread along the cross-membrane axis of the protein. The first two clusters are formed by residues of the pH sensor, the second one mainly by the residues within the ion-binding site, including Asp163 and Asp164, and the last one by amino acids at the rim of the periplasmic funnel. Interestingly, the clusters of the pH sensor are connected electrostatically with the one of the ion-binding site by the functional important residue Asp133. Based on these results, it is suggested that the electrostatic interactions between the clusters are

essential for pH-dependent NhaA activation (Olkhova *et al.*, 2006). The calculations also predicted abnormally high pK_a values for Asp163 and Asp164. This indicates that these residues remain protonated up to a pH value of 15 or more under the assumption that the structure of NhaA at acidic pH is identical to that at alkaline pH. Thus, structural changes, which lower their pK_a values, are required for activation of NhaA.

Mechanism of pH regulation and translocation of NhaA - Molecular dynamics (MD) simulations at pH 4 and pH 8 revealed that structural changes must occur in the binding site, at helix X and in the loop regions, whereas no conformational alterations were observed for the other α -helices throughout the pH shift (Olkhova *et al.*, 2007). These predictions were validated and complemented by different experimental approaches. pH-induced conformational changes of the N terminus (loop 1) and loop 9 were identified by site-specific immunodecoration with a monoclonal antibody (Venturi *et al.*, 2000) and accessibility analysis to trypsin (Gerchman *et al.*, 1999). The predicted pH-dependent conformational alterations of TM X (at acidic pH, helix X is slightly curved; at alkaline pH, it is kinked around residue Lys300) were verified by cross-linking between residues of TM IVp and TM X, and TM II and TM X (Kozachkov *et al.*, 2007). Accessibility analyses of different residues that lines the cytoplasmic cavity to various reagents as a function of pH implied that the cavity deepens at pH 8 (Tzuberly *et al.*, 2008). Furthermore, analysis of the accessibility of the Cys at position 252 to 2-(4'-maleimidylanilino)-naphthalene-6-sulfonic acid (MIANS) showed that the labeling efficiency of this residue increased with increasing pH, suggesting that TM IX is also subjected to conformational changes (Tzuberly *et al.*, 2004). Finally, accessibility analysis with two conformational monoclonal antibodies that inhibit NhaA function revealed that TM XI also plays a role in the structural dynamics of NhaA (Rimon *et al.*, 2008).

Based on these results a putative mechanism of pH regulation and translocation of NhaA was proposed (Hunte *et al.*, 2005; Padan *et al.*, 2005). Below pH 6.5, NhaA exists in its downregulated conformation in which only a part of the cation-binding site (Asp164) faces the cytoplasmic cavity whereas the periplasmic cation passage is blocked. After an alkaline pH shift, it is suggested that in response to a signal from the pH sensor, TMs IX and X undergo conformational alterations that trigger a reorientation of the active site. This process would expose the full Na^+/Li^+ -binding site (Asp164, Asp163 and most probably also Asp133 and Thr132) to the cytoplasmic cavity but leaving the periplasmic sealing intact. This conformation is thought to allow the Na^+/H^+ exchange according to the alternating access mechanism. Binding of Na^+/Li^+ to the substrate-binding site from the cytoplasm is proposed to result in a charge imbalance at the intersection of TMs IV and XI that then would trigger small movements of XIp and IVc and their extended chains. As a result, the Na^+/Li^+ -loaded binding site is suggested to be exposed to the periplasm and sealed off from the cytoplasm. Upon Na^+/Li^+ release, Asp164 and Asp163 would become protonated, thus inducing

conformational alterations that would expose both residues back to the cytoplasm. The subsequent deprotonation of the aspartates will release the two H⁺ to the periplasm to complete the transport cycle. By contrast, under the assumption that the NhaA structure at acidic pH is identical to that at alkaline pH, recent extended MD simulations led to a proposed mechanism for NhaA catalyzed antiport that implied different roles for Asp164 and Asp163 (Arkin *et al.*, 2007). Asp164 was suggested to be the Na⁺-binding site and Asp163 was assigned to control the alternating accessibility of the binding site to the cytoplasm or periplasm in dependence of its protonation state. The role of NhaA dimerization in the pH-regulation and/or activity of the antiporter was not implicated in both models, since no physiologically significant high-resolution structure of the dimer is available yet.

1.9. Scope of this work

The main objective of this thesis was the development of new approaches for structure determination of membrane proteins in their native environment based on DEER distance measurements. Further aims were the characterization of the functional role and dynamics of TM IX of the Na⁺/proline symporter PutP of *E. coli* and the improvement of the DEER data analysis.

Previous site-directed mutagenesis studies revealed that Ser340 and Thr341 in TM IX are crucial for transport activity of PutP. In the *first part* of this thesis, the functional relevance of these residues was further elucidated by genetic and biochemical approaches. In addition, intramolecular Cys cross-linking studies were performed to test for participation of Ser340 and Thr341 in the formation of a ligand-binding site together with residues of the functional important TM II. For identification of additional important amino acids, all residues of TM IX (positions 324 to 351) were individually replaced by Cys and the impact on transporter function was analyzed. Furthermore, accessibilities of the generated Cys substitutions to sulfhydryl-specific reagents of various sizes and polarities were determined to examine if residues of TM IX line a cytoplasmic or periplasmic cavity. The influence of ligands (Na⁺ and proline) on the accessibility was investigated in order to identify conformational alterations. Since the available X-ray structures of ion-dependent secondary transporters revealed that TMs are often discontinuous or kinked helices (Screpanti and Hunte, 2007), the backbone structure of TM IX of PutP was determined. For this purpose, a novel approach for structure elucidation of TMs was established based on SDSL, DEER distance measurements, and explicit modeling of spin label conformations.

Since no high-resolution X-ray structure for PutP is available yet, the *second part* of this thesis was devoted to model the helix bundle of PutP. For this aim, a new approach was

established that relies on DEER distance measurements between spin labels attached to helix ends, structural homology, and distance geometry.

Several experiments indicated that the pH-regulated Na⁺/H⁺ antiporter NhaA forms homodimers in the native membrane (Padan *et al.*, 2004). Although the interaction between NhaA molecules was proposed to be involved in the pH response (Gerchman *et al.*, 2001), the explicit role of the dimeric state for NhaA activity and/or pH regulation has remained elusive. Thus, in the *third part* of this thesis, pH dependent structural dynamics of the NhaA monomer and dimer were studied by CW EPR techniques and DEER distance measurements. Furthermore, since the available X-ray structure of NhaA displays only a nonphysiological association of the monomers (Hunte *et al.*, 2005), the native structural arrangement of the NhaA dimer was determined based on SDSL and DEER distance measurements. For this purpose a novel approach was introduced that allows for structure elucidation of protein dimers with high resolution if the X-ray structure of the monomer is available.

Finally, a comprehensive software package for analyzing DEER or other pulsed electron-electron double resonance (PELDOR) data was developed. Algorithms for Tikhonov regularization with L curve computation, model-based fitting, excitation bandwidth correction, determination and use of experimental background functions, spin counting, and concentration measurements were implemented into the program. Furthermore, these tools were tested on three model systems: a shape-persistent biradical, spin-labeled gold nanoparticles, and singly spin-labeled PutP.

1.10. References

- Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R., and Iwata, S.** (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* 301: 610-615.
- Altenbach, C., Greenhalgh, D. A., Khorana, H. G., and Hubbell, W. L.** (1994) A collision gradient method to determine the immersion depth of nitroxides in lipid bilayers: application to spin-labeled mutants of bacteriorhodopsin. *Proc. Natl. Acad. Sci. U. S. A.* 91: 1667-1671.
- Altenbach, C., Oh, K. J., Trabanino, R. J., Hideg, K., and Hubbell, W. L.** (2001) Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: experimental strategies and practical limitations. *Biochemistry* 40: 15471-15482.
- Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P., and Hubbell, W. L.** (2008) High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl. Acad. Sci. U. S. A.* 105: 7439-7444.
- Apse, M. P., and Blumwald, E.** (2007) Na⁺ transport in plants. *FEBS Lett.* 581: 2247-2254.

- Arkin, I. T., Xu, H., Jensen, M. Ø., Arbely, E., Bennett, E. R., Bowers, K. J., Chow, E., Dror, R. O., Eastwood, M. P., Flitman-Tene, R., Gregersen, B. A., Klepeis, J. L., Kolossváry, I., Shan, Y., and Shaw, D. E.** (2007) Mechanism of Na⁺/H⁺ antiporting. *Science* 317: 799-803.
- Auer, M., Kim, M. J., Lemieux, M. J., Villa, A., Song, J., Li, X. D., and Wang, D. N.** (2001) High-yield expression and functional analysis of *Escherichia coli* glycerol-3-phosphate transporter. *Biochemistry* 40: 6628-6635.
- Bandell, M., Ansanay, V., Rachidi, N., Dequin, S., and Lolkema, J. S.** (1997) Membrane potential-generating malate (MleP) and citrate (CitP) transporters of lactic acid bacteria are homologous proteins. Substrate specificity of the 2-hydroxycarboxylate transporter family. *J. Biol. Chem.* 272: 18140-18146.
- Banham, J. E., Timmel, C. R., Abbott, R. J., Lea, S. M., and Jeschke, G.** (2006) The characterization of weak protein-protein interactions: evidence from DEER for the trimerization of a von Willebrand Factor A domain in solution. *Angew. Chem. Int. Ed. Engl.* 45: 1058-1061.
- Barnell, W. O., Yi, K. C., and Conway, T.** (1990) Sequence and genetic organization of a *Zymomonas mobilis* gene cluster that encodes several enzymes of glucose metabolism. *J. Bacteriol.* 172: 7227-7240.
- Berman, H., Henrick, K., Nakamura, H., and Markley, J. L.** (2007) The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.* 35: D301-D303.
- Blair, D. F.** (2003) Flagellar movement driven by proton translocation. *FEBS Lett.* 545: 86-95.
- Borbat, P. P., and Freed, J. H.** (1999) Multiple-quantum ESR and distance measurements. *Chem. Phys. Lett.* 313: 145-154.
- Borbat, P. P., McHaourab, H. S., and Freed, J. H.** (2002) Protein structure determination using long-distance constraints from double-quantum coherence ESR: study of T4 lysozyme. *J. Am. Chem. Soc.* 124: 5304-5314.
- Borbat, P. P., Davis J. H., Butcher, S. E., and Freed, J. H.** (2004) Measurement of large distances in biomolecules using double-quantum filtered refocused electron spin-echoes. *J. Am. Chem. Soc.* 126: 7746-7747.
- Borbat, P. P., Surendhran, K., Bortolus, M., Zou, P., Freed, J. H., and Mchaourab, H. S.** (2007) Conformational motion of the ABC transporter MsbA induced by ATP hydrolysis. *PLoS Biol.* 5: 2211-2219.
- Bordignon, E., and Steinhoff, H.-J.** (2007) Membrane protein structure and dynamics studied by site-directed spin labeling ESR. In *ESR spectroscopy in membrane biophysics*. (Hemminga, M.A., and Berliner, L.J., eds.) pp. 129-164, Springer Science and Business Media, New York.
- Borovykh, I. V., Ceola, S., Gajula, P., Gast, P., Steinhoff, H.-J., and Huber, M.** (2006) Distance between a native cofactor and a spin label in the reaction centre of *Rhodobacter sphaeroides* by a two-frequency pulsed electron paramagnetic

resonance method and molecular dynamics simulations. *J. Magn. Reson.* 180: 178-185.

- Boudker, O., Ryan, R. M., Yernool, D., Shimamoto, K., and Gouaux, E.** (2007) Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature* 445: 387-393.
- Camilli, A. and Bassler, B. L.** (2006) Bacterial small-molecule signaling pathways. *Science* 311: 1113-1116.
- Chen, C. C., Tsuchiya, T., Yamane, Y., Wood, J. M., and Wilson, T. H.** (1985) Na⁺ (Li⁺)-proline cotransport in *Escherichia coli*. *J. Membr. Biol.* 84: 157-164.
- Chen, Y. J., Pornillos, O., Lieu, S., Ma, C., Chen, A. P., and Chang, G.** (2007) X-ray structure of EmrE supports dual topology model. *Proc. Natl. Acad. Sci. U. S. A.* 104: 18999-19004.
- Columbus, L., and Hubbell, W. L.** (2002) A new spin on protein dynamics. *Trends Biochem. Sci.* 27: 288-295.
- Davidson, A. L., and Maloney, P. C.** (2007) ABC transporters: how small machines do a big job. *Trends Microbiol.* 15: 448-455.
- De la Vieja, A., Reed, M. D., Ginter, C. S., and Carrasco, N.** (2007) Amino acid residues in transmembrane segment IX of the Na⁺/I⁻ symporter play a role in its Na⁺ dependence and are critical for transport activity. *J. Biol. Chem.* 282: 25290-25298.
- DiMarco, A. A., and Romano, A. H.** (1985) D-Glucose transport system of *Zymomonas mobilis*. *Appl. Environ. Microbiol.* 49: 151-157.
- Dimroth, P.** (1980) A new sodium-transport system energized by the decarboxylation of oxaloacetate. *FEBS Lett.* 122: 234-236.
- Dimroth, P.** (1994) Bacterial sodium ion-coupled energetics. *Antonie Van Leeuwenhoek.* 65: 381-395.
- Dimroth, P.** (1997) Primary sodium ion translocating enzymes. *Biochim. Biophys. Acta* 1318: 11-51.
- Dohán, O., De la Vieja, A., and Carrasco, N.** (2006) Hydrocortisone and purinergic signaling stimulate sodium/iodide symporter (NIS)-mediated iodide transport in breast cancer cells. *Mol. Endocrinol.* 20: 1121-1137.
- Dutzler, R.** (2006) The CIC family of chloride channels and transporters. *Curr. Opin. Struct. Biol.* 16: 439-446.
- Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T., and MacKinnon R.** (2002) X-ray structure of a CIC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature* 415: 287-294.

- Duyckaerts, C., Sluse-Goffart, C. M., Fux, J. P., Sluse, F. E., and Liebecq, C.** (1980) Kinetic mechanism of the exchanges catalysed by the adenine-nucleotide carrier. *Eur. J. Biochem.* 106: 1-6.
- Engelhardt, S., Hein, L., Keller, U., Klämbt, K., and Lohse, M. J.** (2002) Inhibition of Na⁺-H⁺ exchange prevents hypertrophy, fibrosis, and heart failure in β₁-adrenergic receptor transgenic mice. *Circ. Res.* 90: 814-819.
- Essen, L. O.** (2002) Halorhodopsin: light-driven ion pumping made simple? *Curr. Opin. Struct. Biol.* 12: 516-522.
- Fanucci, G. E., and Cafiso, D. S.** (2006) Recent advances and applications of site-directed spin labeling. *Curr. Opin. Struct. Biol.* 16: 644-653.
- Galili, L., Rothman, A., Kozachkov, L., Rimon, A., and Padan, E.** (2002) Transmembrane domain IV is involved in ion transport activity and pH regulation of the NhaA-Na⁺/H⁺ antiporter of *Escherichia coli*. *Biochemistry* 41: 609-617.
- Galili, L., Herz, K., Dym, O., and Padan, E.** (2004) Unraveling functional and structural interactions between transmembrane domains IV and XI of NhaA Na⁺/H⁺ antiporter of *Escherichia coli*. *J. Biol. Chem.* 279: 23104-23113.
- Gerchman, Y., Rimon, A., and Padan, E.** (1999) A pH-dependent conformational change of NhaA Na⁺/H⁺ antiporter of *Escherichia coli* involves loop VIII-IX, plays a role in the pH response of the protein, and is maintained by the pure protein in dodecyl maltoside. *J. Biol. Chem.* 274: 24617-24624.
- Gerchman, Y., Rimon, A., Venturi, M., and Padan, E.** (2001) Oligomerization of NhaA, the Na⁺/H⁺ antiporter of *Escherichia coli* in the membrane and its functional and structural consequences. *Biochemistry* 40: 3403-3412.
- Gottschalk G., and Thauer, R. K.** (2001) The Na⁺-translocating methyltransferase complex from methanogenic archaea. *Biochim. Biophys. Acta* 1505: 28-36.
- Guan, L., and Kaback, H. R.** (2006) Lessons from lactose permease. *Annu. Rev. Biophys. Biomol. Struct.* 35: 67-91.
- Guan, L., Mirza, O., Verner, G., Iwata, S., and Kaback, H. R.** (2007) Structural determination of wild-type lactose permease. *Proc. Natl. Acad. Sci. U. S. A.* 104: 15294-15298.
- Hanada, K., Yoshida, T., Yamato, I., and Anraku, Y.** (1992) Sodium ion and proline binding sites in the Na⁺/proline symport carrier of *Escherichia coli*. *Biochim. Biophys. Acta* 1105: 61-66.
- Hase, C. C., Fedorova, N. D., Galperin, M. Y., and Dibrov, P. A.** (2001) Sodium ion cycle in bacterial pathogens: evidence from cross-genome comparisons. *Microbiol. Mol. Biol. Rev.* 65: 353-370.
- Hayashi, M., Nakayama, Y., and Unemoto, T.** (2001) Recent progress in the Na⁺-translocating NADH-quinone reductase from the marine *Vibrio alginolyticus*. *Biochim. Biophys. Acta* 1505: 37-44.

- Herzberg, M., Kaye, I. K., Peti, W., and Wood, T. K.** (2006) YdgG (TqsA) controls biofilm formation in *Escherichia coli* K-12 through autoinducer 2 transport. *J. Bacteriol.* 188: 587-598.
- Hillar, A., Culham, D. E., Vernikovska, Y. I., Wood, J. M., and Boggs, J. M.** (2005) Formation of an antiparallel, intermolecular coiled coil is associated with *in vivo* dimerization of osmosensor and osmoprotectant transporter ProP in *Escherichia coli*. *Biochemistry* 44: 10170-10180.
- Hilpert W., Schink B., and Dimroth P.** (1984) Life by a new decarboxylation-dependent energy conservation mechanism with Na⁺ as coupling ion. *EMBO J.* 3: 1665-1670.
- Hirai, T., Heymann, J. A., Shi, D., Sarker, R., Maloney, P. C., and Subramaniam, S.** (2002). Three-dimensional structure of a bacterial oxalate transporter. *Nat. Struct. Biol.* 9: 597-600.
- Huang, S. G., Odoy, S., and Klingenberg, M.** (2001) Chimers of two fused ADP/ATP carrier monomers indicate a single channel for ADP/ATP transport. *Arch. Biochem. Biophys.* 394: 67-75.
- Huang, Y., Lemieux, M. J., Song, J., Auer, M., and Wang, D. N.** (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* 301: 616-620.
- Hubbell, W. L., Gross, A., Langen, R., and Lietzow, M. A.** (1998) Recent advances in site-directed spin labeling of proteins. *Curr. Opin. Struct. Biol.* 8: 649-656.
- Hubbell, W. L., Cafiso, D. S., and Altenbach, C.** (2000) Identifying conformational changes with site-directed spin labeling. *Nat. Struct. Biol.* 7: 735-739.
- Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E., and Michel, H.** (2005) Structure of a Na⁺/H⁺ antiporter and insights into mechanism of action and regulation by pH. *Nature* 435: 1197-1202.
- Inoue, H., Noumi, T., Tsuchiya, T., and Kanazawa, H.** (1995) Essential aspartic acid residues, Asp-133, Asp-163 and Asp-164, in the transmembrane helices of a Na⁺/H⁺ antiporter (NhaA) from *Escherichia coli*. *FEBS Lett.* 363: 264-268.
- Jardetzky, O.** (1966) Simple allosteric model for membrane pumps. *Nature* 211: 969-970.
- Jeschke G.** (2002) Distance measurements in the nanometer range by pulse EPR. *ChemPhysChem* 3: 927-932.
- Jeschke, G., Wegener, C., Nietschke, M., Jung, H., and Steinhoff, H.-J.** (2004a) Interresidual distance determination by four-pulse double electron-electron resonance in an integral membrane proteins: the Na⁺/proline transporter PutP of *Escherichia coli*. *Biophys J.* 86: 2551-2557.
- Jeschke, G., Panek, G., Godt, A., Bender, A., and Paulsen, H.** (2004b) Data analysis procedures for pulse ELDOR measurements of broad distance distributions. *Appl. Magn. Reson.* 26: 223-244.

- Jeschke, G., Bender, A., Schweikardt T., Panek, G., Decker, H., and Paulsen H.** (2005) Localization of the N-terminal domain in light-harvesting chlorophyll *a/b* protein by EPR measurements. *J. Biol. Chem.* 280: 18623-18630.
- Jeschke, G., and Polyhach, Y.** (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* 9: 1895-1910.
- Jung, H.** (1998) Topology and function of the Na⁺/proline transporter of *Escherichia coli*, a member of the Na⁺/solute cotransporter family. *Biochim. Biophys. Acta* 1365: 60-64.
- Jung, H., Rübenhagen, R., Tebbe, S., Leifker, K., Tholema, N., Quick, M., and Schmid, R.** (1998) Topology of the Na⁺/proline transporter of *Escherichia coli*. *J. Biol. Chem.* 273: 26400-26407.
- Jung, H.** (2001) Towards the molecular mechanism of Na⁺/solute symport in prokaryotes. *Biochim. Biophys. Acta* 1505: 131-143.
- Jung, H.** (2002) The sodium/substrate symporter family: structural and functional features. *FEBS Lett.* 529: 73-77.
- Junge, W., Lill, H., and Engelbrecht, S.** (1997) ATP synthase: an electrochemical transducer with rotatory mechanics. *Trends. Biochem. Sci.* 22: 420-423.
- Just, H., Sitte, H. H., Schmid, J. A., Freissmuth, M., and Kudlacek, O.** (2004) Identification of an additional interaction domain in transmembrane domains 11 and 12 that supports oligomer formation in the human serotonin transporter. *J. Biol. Chem.* 279: 6650-6657.
- Kaback, H. R.** (2005) Structure and mechanism of the lactose permease. *C. R. Biol.* 328: 557-567.
- Kaback, H. R., Dunten, R., Frillingos, S., Venkatesan, P., Kwaw, I., Zhang, W., and Ermolova, N.** (2007) Site-directed alkylation and the alternating access model for LacY. *Proc. Natl. Acad. Sci. U. S. A.* 104: 491-494.
- Kavanaugh, M. P.** (2004) Accessing a transporter structure. *Nature* 431: 752-753.
- Kavermann, H., Burns, B. P., Angermuller, K., Odenbreit, S., Fischer, W., Melchers, K., and Haas, R.** (2003) Identification and characterization of *Helicobacter pylori* genes essential for gastric colonization. *J. Exp. Med.* 197: 813-822.
- Konings, W. N., Lolkema, J. S., and Poolman B.** (1995) The generation of metabolic energy by solute transport. *Arch. Microbiol.* 164: 235-242.
- Kozachkov, L., Herz, K., and Padan, E.** (2007) Functional and structural interactions of the transmembrane domain X of NhaA, Na⁺/H⁺ antiporter of *Escherichia coli*, at physiological pH. *Biochemistry* 46: 2419-2430.
- Krämer, R., and Morbach, S.** (2004) BetP of *Corynebacterium glutamicum*, a transporter with three different functions: betaine transport, osmosensing, and osmoregulation. *Biochim. Biophys. Acta* 1658: 31-36.

- Lacapère, J. J., Pebay-Peyroula, E., Neumann, J. M., and Etchebest, C.** (2007) Determining membrane protein structures: still a challenge! *Trends Biochem. Sci.* 32: 259-70.
- Lanyi, J. K.** (2006) Proton transfers in the bacteriorhodopsin photocycle. *Biochim. Biophys. Acta* 1757: 1012-1018.
- Law, C. J., Almqvist, J., Bernstein, A., Goetz, R. M., Huang, Y., Soudant, C., Laaksonen, A., Hovmöller, S., and Wang, D.-N.** (2008a) Salt-bridge dynamics control substrate-induced conformational change in the membrane transporter GlpT. *J. Mol. Biol.* 378: 826-837.
- Law, C. J., Maloney, P. C., and Wang, D.-N.** (2008b) Ins and outs of major facilitator superfamily antiporters. *Annu. Rev. Microbiol.* 62: 289-305.
- Lolkema, J. S., Poolman, B., and Konings, W. N.** (1996) Secondary transporters and metabolic energy generation in bacteria. In *Transport processes in eukaryotic and prokaryotic organisms (Handbook of biological physics)* pp. 229-260, Elsevier, Amsterdam.
- Malinen, A. M., Belogurov, G. A., Baykov, A. A., and Lahti, R.** (2007) Na⁺-pyrophosphatase: a novel primary sodium pump. *Biochemistry* 46: 8872-8878.
- Maloney, P. C.** (1994) Bacterial transporters. *Curr. Opin. Struct. Biol.* 6: 571-582.
- Milov, A. D., Ponomarev, A. B., and Tsvetkov, Yu. D.** (1984) Electron-electron double resonance in electron spin echo: Model biradical systems and the sensitized photolysis of decalin. *Chem. Phys. Lett.* 110: 67-72.
- Mitchell, P.** (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemiosmotic type of mechanism. *Nature* 191: 144-148.
- Mitchell, P.** (1966) Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. *Physiol. Rev.* 41: 445-502.
- Mitchell, P.** (1991) Foundations of vectorial metabolism and osmochemistry. *Biosci. Rep.* 11: 297-344.
- Mueckler, M.** (1994) Facilitative glucose transporters. *Eur. J. Biochem.* 219: 713-725.
- Murakami, S., Nakashima, R., Yamashita, E., and Yamaguchi, A.** (2002) Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* 419: 587-593.
- Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T., and Yamaguchi, A.** (2006) Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* 443: 173-179.
- Myers, M. L.** (2003) NHE-1 inhibitors: potential application in cardiac surgery. In *The sodium-hydrogen exchange, from molecule to its role in disease.* (Karmazyn, M., Avkiran, M., and Fliegel, L., eds.) pp. 279-290, Editor Kluwer Academic, Boston, Massachusetts

- Nagata, T., Iizumi, S., Satoh, K., and Kikuchi, S.** (2008) Comparative molecular biological analysis of membrane transport genes in organisms. *Plant Mol. Biol.* 66: 565-585.
- Nury, H., Dahout-Gonzalez, C., Trézéguet, V., Lauquin, G., Brandolin, G., and Pebay-Peyroula, E.** (2005) Structural basis for lipid-mediated interactions between mitochondrial ADP/ATP carrier monomers. *FEBS Lett.* 579: 6031-6036.
- Nury, H., Dahout-Gonzalez, C., Trézéguet, V., Lauquin, G. J., Brandolin, G., and Pebay-Peyroula, E.** (2006) Relations between structure and function of the mitochondrial ADP/ATP carrier. *Annu. Rev. Biochem.* 75: 713-741.
- Ohsawa, M., Mogi, T., Yamamoto, H., Yamato, I., and Anraku, Y.** (1988) Proline carrier mutant of *Escherichia coli* K-12 with altered cation sensitivity of substrate-binding activity: cloning biochemical characterization, and identification of the mutation. *J. Bacteriol.* 170: 5185-5191.
- Olami, Y., Rimon, A., Gerchman, Y., Rothman, A., and Padan, E.** (1997) Histidine 225, a residue of the NhaA-Na⁺/H⁺ antiporter of *Escherichia coli* is exposed and faces the cell exterior. *J. Biol. Chem.* 272: 1761-1768.
- Olkhova, E., Hunte, C., Screpanti, E., Padan, E., and Michel, H.** (2006) Multiconformation continuum electrostatics analysis of the NhaA Na⁺/H⁺ antiporter of *Escherichia coli* with functional implications. *Proc. Natl. Acad. Sci. U. S. A.* 103: 2629-2634.
- Olkhova, E., Padan, E., and Michel, H.** (2007) The influence of protonation states on the dynamics of the NhaA antiporter from *Escherichia coli*. *Biophys. J.* 92: 3784-3791.
- Orlowski, J., and Grinstein, S.** (2004) Diversity of the mammalian sodium/proton exchanger SLC9 gene family. *Pflugers Arch.* 447: 549-565.
- Padan, E., Venturi, M., Gerchman, Y., and Dover, N.** (2001) Na⁺/H⁺ antiporters. *Biochim. Biophys. Acta* 1505: 144-157.
- Padan, E., Tzuberly, T., Herz, K., Kozachkov, L., Rimon, A., and Galili L.** (2004) NhaA of *Escherichia coli*, as a model of a pH-regulated Na⁺/H⁺ antiporter. *Biochim. Biophys. Acta* 1658: 2-13.
- Padan, E., Bibi, E., Ito, M., and Krulwich, T. A.** (2005) Alkaline pH homeostasis in bacteria: new insights. *Biochim. Biophys. Acta* 1717: 67-88.
- Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W.** (2000) Dead-time free measurement of dipole-dipole interactions between electron spins. *J. Magn. Reson.* 142: 331-340.
- Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trézéguet, V., Lauquin, G. J., and Brandolin, G.** (2003) Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* 426: 39-44.
- Perozo, E., Cortes, D. M., and Cuello, L. G.** (1998) Three-dimensional architecture and gating mechanism of a K⁺ channel studied by EPR spectroscopy. *Nat. Struct. Biol.* 5: 459-469.

- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E.** (2004) UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25: 1605-1612.
- Pirch, T., Quick, M., Nietschke, M., Langkamp, M., and Jung, H.** (2002) Sites important for Na⁺ and substrate binding in the Na⁺/proline transporter of *Escherichia coli*, a member of the Na⁺/solute symporter family. *J. Biol. Chem.* 277: 8790-8796.
- Pirch, T., Landmeier, S., and Jung, H.** (2003) Transmembrane domain II of the Na⁺/proline transporter PutP of *Escherichia coli* forms part of a conformationally flexible, cytoplasmic exposed aqueous cavity within the membrane. *J. Biol. Chem.* 278: 42942-42949.
- Pourcher, T., Bassilana, M., Sarkar, H. K., Kaback, H. R., and Leblanc, G.** (1990) The melibiose/Na⁺ symporter of *Escherichia coli*: kinetic and molecular properties. *Philos. Trans. R. Soc. Lond., B Biol. Sci.* 326: 411-423.
- Quick, M., Tebbe, S., and Jung, H.** (1996) Ser57 in the Na⁺/proline permease of *Escherichia coli* is critical for high-affinity proline uptake. *Eur. J. Biochem.* 239: 732-736.
- Quick, M., and Jung, H.** (1997) Aspartate 55 in the Na⁺/proline permease of *Escherichia coli* is essential for Na⁺-coupled proline uptake. *Biochemistry* 36: 4631-4636.
- Quick M., and Jung, H.** (1998) A conserved aspartate residue, Asp187, is important for Na⁺-dependent proline binding and transport by the Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 37: 13800-13806.
- Quick, M., Stölting, S., and Jung, H.** (1999) Role of conserved Arg40 and Arg117 in the Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 38: 13523-13529.
- Rabenstein, M. D., and Shin, Y. K.** (1995) Determination of the distance between two spin labels attached to a macromolecule. *Proc. Natl. Acad. Sci. U. S. A.* 92: 8239-8243.
- Rapp, M., Seppälä, S., Granseth, E., and von Heijne, G.** (2007) Emulating membrane protein evolution by rational design. *Science* 315: 1282-1284.
- Reed-Tsur, M. D., De la Vieja, A., Ginter, C. S., and Carrasco, N.** (2008) Molecular characterization of V59E NIS, a Na⁺/I⁻ symporter mutant that causes congenital I⁻ transport defect. *Endocrinology* 149: 3077-3084.
- Reizer, J., Reizer, A., and Saier, M. H. Jr.** (1994) A functional superfamily of sodium/solute symporters. *Biochim. Biophys. Acta* 1197: 133-166.
- Reizer, J., and Saier, M. H. Jr.** (1997) Modular multidomain phosphoryl transfer proteins of bacteria. *Curr. Opin. Struct. Biol.* 7: 407-415.
- Ren, Q., and Paulsen, I. T.** (2005) Comparative analysis of fundamental differences in membrane transport capabilities in prokaryotes and eukaryotes. *PLoS Comp. Biol.* 1: 190-201.
- Rimon, A., Gerchman, Y., Kariv, Z., and Padan, E.** (1998) A point mutation (G338S) and its suppressor mutations affect both the pH response of the NhaA-Na⁺/H⁺

antiporter as well as the growth phenotype of *Escherichia coli*. *J. Biol. Chem.* 273: 26470-26476.

- Rimon, A., Hunte, C., Michel, H., and Padan, E.** (2008) Epitope mapping of conformational monoclonal antibodies specific to NhaA Na⁺/H⁺ antiporter: structural and functional implications. *J. Mol. Biol.* 379: 471-481.
- Rink, T., Pfeiffer, M., Oesterhelt, D., Gerwert, K., and Steinhoff, H.-J.** (2000) Unraveling photoexcited conformational changes of bacteriorhodopsin by time resolved electron paramagnetic resonance spectroscopy. *Biophys. J.* 78: 1519-1530.
- Rothman, A., Padan, E., and Schuldiner, S.** (1996) Topological analysis of NhaA, a Na⁺/H⁺ antiporter from *Escherichia coli*. *J. Biol. Chem.* 271: 32288-32292.
- Ruan, Z. S., Anantharam, V., Crawford, I. T., Ambudkar, S. V., Rhee, S. Y., Allison, M. J., and Maloney, P. C.** (1992) Identification, purification, and reconstitution of OxIT, the oxalate: formate antiport protein of *Oxalobacter formigenes*. *J. Biol. Chem.* 267: 10537-10543.
- Rudnick, G.** (1998) Bioenergetics of neurotransmitter transport. *J. Bioenerg. Biomembr.* 30: 173-185.
- Saier, M. H. Jr.** (1998) Molecular phylogeny as a basis for the classification of transport proteins from bacteria, archaea and eukarya. In *Advances in Microbial Physiology* (Poole, R. K., ed.) pp. 81-136, Academic Press, San Diego, CA.
- Saier, M. H. Jr.** (1999) Classification of transmembrane transport proteins in living organisms. In *Biomembrane Transport* (VanWinkle, L., ed.) pp. 265-276, Academic Press, San Diego, CA.
- Saier, M. H. Jr., Beatty, J. T., Goffeau, A., Harley, K. T., Heijne, W. H., Huang, S. C., Jack, D. L., Jähn, P. S., Lew, K., Liu, J., Pao, S. S., Paulsen, I. T., Tseng, T. T., and Virk, P. S.** (1999) The major facilitator superfamily. *J. Mol. Microbiol. Biotechnol.* 1: 257-279.
- Saier, M. H. Jr.** (2003) Tracing pathways of transport protein evolution. *Mol. Microbiol.* 48: 1145-1156.
- Sale, K., Song, L., Liu, Y. S., Perozo, E., and Fajer, P.** (2005) Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *J. Am. Chem. Soc.* 127: 9334-9335.
- Sato, K., and Homma, M.** (2000) Functional reconstitution of the Na⁺-driven polar flagellar motor component of *Vibrio alginolyticus*. *J. Biol. Chem.* 275: 5718-5722.
- Schuldiner, S., and Padan, E.** (1993) Na⁺/H⁺ antiporters in *Escherichia coli*. In *Alkali cation transport systems in prokaryotes*. (Baker, E. P., ed.) pp. 25-51, CRC Press, Boca Raton.
- Schuldiner, W., Lebendiker, M., and Yerushalmi, H.** (1997) EmrE, the smallest ion-coupled transporter, provides a unique paradigm for structure-function studies. *J. Exp. Biol.* 200: 335-341.
- Schuldiner S.** (2007) Controversy over EmrE structure. *Science* 317: 748-751.

- Schwan, W. R., Lehmann, L., and McCormick, J.** (2006) Transcriptional activation of the *Staphylococcus aureus putP* gene by low-proline-high osmotic conditions and during infection of murine and human tissues. *Infect. Immun.* 74: 399-409.
- Screpanti, E., Padan, E., Rimon, A., Michel, H., and Hunte, C.** (2006) Crucial steps in the structure determination of the Na⁺/H⁺ antiporter NhaA in its native conformation. *J. Mol. Biol.* 362: 192-202.
- Screpanti, E., and Hunte, C.** (2007) Discontinuous membrane helices in transport proteins and their correlation with function. *J. Struct. Biol.* 159: 261-267.
- Seeger, M. A., Schiefner, A., Eicher, T., Verrey, F., Diederichs, K., and Pos, K. M.** (2006) Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science* 313: 1295-1298.
- Serrano, A., Perez-Castineira, J. R., Baltscheffsky, H., and Baltscheffsky, M.** (2004) Proton-pumping inorganic pyrophosphatases in some archaea and other extremophilic prokaryotes. *J. Bioenerg. Biomembr.* 36: 127-133.
- Sitte, H. H., Farhan, H., and Javitch, J. A.** (2004) Sodium-dependent neurotransmitter transporters: oligomerization as a determinant of transporter function and trafficking. *Mol. Interv.* 4: 38-47.
- Skulachev, V. P.** (1985) Membrane-linked energy transductions. Bioenergetic functions of sodium: H⁺ is not unique as a coupling ion. *Eur. J. Biochem.* 151: 199-208.
- Skulachev, V. P.** (1991) Chemiosmotic Systems in Bioenergetics: H⁺-cycles and Na⁺-cycles. *Bioscience Reports* 11: 387-444.
- Smirnova, I., Kasho, V., Choe, J. Y., Altenbach, C., Hubbell, W. L., and Kaback, H. R.** (2007) Sugar binding induces an outward facing conformation of LacY. *Proc. Natl. Acad. Sci. U. S. A.* 104: 16504-16509.
- Sonders, M. S., Quick, M., and Javitch, J. A.** (2005) How did the neurotransmitter cross the bilayer? A closer view. *Curr. Opin. Neurobiol.* 15: 296-304.
- Speelmans, G., Poolman, B., Abee, T., and Konings, W. N.** (1993) Energy transduction in the thermophilic anaerobic bacterium *Clostridium fervidus* is exclusively coupled to sodium ions. *Proc. Natl. Acad. Sci. U. S. A.* 90: 7975-7979.
- Spiegelhalter F., and Bremer, E.** (1998) Osmoregulation of the *opuE* proline transport gene from *Bacillus subtilis*: contributions of the sigma A- and sigma B-dependent stress-responsive promoters. *Mol. Microbiol.* 29: 285-296.
- Steinhoff, H.-J., Radzwill, N., Thevis, W., Lenz, V., Brandenburg, D., Antson, A., Dodson, G., and Wollmer, A.** (1997) Determination of interspin distances between spin labels attached to insulin: comparison of electron paramagnetic resonance data with the X-ray structure. *Biophys. J.* 73: 3287-3298.
- Steinhoff, H.-J., Savitsky, A., Wegener, C., Pfeiffer, M., Plato, M., and Möbius, K.** (2000) High-field EPR studies of the structure and conformational changes of site directed spin labeled bacteriorhodopsin. *Biochim. Biophys. Acta* 1457: 253-262.

- Taglicht, D., Padan, E., and Schuldiner, S.** (1991) Overproduction and purification of a functional Na⁺/H⁺ antiporter coded by *nhaA* (*ant*) from *Escherichia coli*. *J. Biol. Chem.* 266: 11289-11294.
- Taglicht, D., Padan, E., and Schuldiner S.** (1993) Proton-sodium stoichiometry of NhaA, an electrogenic antiporter from *Escherichia coli*. *J. Biol. Chem.* 268: 5382-5387.
- Tokuda, H.** (1989) Respiratory Na⁺ pump and Na⁺-dependent energetics in *Vibrio alginolyticus*. *J. Bioenerg. Biomembr.* 21: 693-704.
- Tzubery, T., Rimon, A., and Padan, E.** (2004) Mutation E252C increases drastically the K_m value for Na⁺ and causes an alkaline shift of the pH dependence of NhaA Na⁺/H⁺ antiporter of *Escherichia coli*. *J. Biol. Chem.* 279: 3265-3272.
- Tzubery, T., Rimon, A., and Padan, E.** (2008) Structure-based functional study reveals multiple roles of transmembrane segment IX and loop VIII-IX in NhaA Na⁺/H⁺ antiporter of *Escherichia coli* at physiological pH. *J. Biol. Chem.* 283: 15975-15987.
- Vamvouka, M., Cieslak, J., Van Eps, N., Hubbell, W., and Gross, A.** (2008) The structure of the lipid-embedded potassium channel voltage sensor determined by double-electron-electron resonance spectroscopy. *Protein Sci.* 17: 506-517.
- Veenhoff, L. M., Heuberger, E. H., and Poolman, B.** (2002) Quaternary structure and function of transport proteins. *Trends Biochem. Sci.* 27: 242-249.
- Venturi, M., Rimon, A., Gerchman, Y., Hunte, C., Padan, E., and Michel, H.** (2000) The monoclonal antibody 1F6 identifies a pH-dependent conformational change in the hydrophilic NH₂ terminus of NhaA Na⁺/H⁺ antiporter of *Escherichia coli*. *J. Biol. Chem.* 275: 4734-4742.
- von Blohn, C., Kempf, B., Kappes, R. M., and Bremer, E.** (1997) Osmostress response in *Bacillus subtilis*: characterization of a proline uptake system (OpuE) regulated by high osmolarity and the alternative transcription factor sigma B. *Mol. Microbiol.* 25: 175-187.
- Wegener, C., Tebbe, S., Steinhoff, H.-J., and Jung, H.** (2000) Spin labeling analysis of structure and dynamics of the Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 39: 4831-4837.
- Wegener, A. A., Klare, J. P., Engelhard, M., and Steinhoff, H.-J.** (2001) Structural insights into the early steps of receptor-transducer signal transfer in archaeal phototaxis. *EMBO J.* 20: 5312-5319.
- West, I. C.** (1997). Ligand conduction and the gated-pore mechanism of transmembrane transport. *Biochim. Biophys. Acta* 1331: 213-234.
- Williams, K. A., Geldmacher-Kaufer, U., Padan, E., Schuldiner, S., and Kühlbrandt, W.** (1999) Projection structure of NhaA, a secondary transporter from *Escherichia coli*, at 4.0 Å resolution. *EMBO J.* 18: 3558-3563.
- Williams, K. A.** (2000) Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature* 403: 112-115.

- Wood, J. M.** (2007) Bacterial osmosensing transporters. *Methods Enzymol.* 428: 77-107.
- Wright, E. M., and Turk, E.** (2004) The sodium/glucose cotransport family SLC5. *Pflugers Arch.* 447: 510-518.
- Wright, E. M., Hirayama, B. A., and Loo, D. F.** (2007) Active sugar transport in health and disease. *J. Intern. Med.* 261: 32-43.
- Yamaguchi, T., and Blumwald, E.** (2005) Developing salt-tolerant crop plants: challenges and opportunities. *Trends Plant Sci.* 10: 615-620.
- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E.** (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* 437: 215-223.
- Yamato, I., and Anraku, Y.** (1988) Site-specific alteration of cysteine 281, cysteine 344, and cysteine 349 in the proline carrier of *Escherichia coli*. *J. Biol. Chem.* 263: 16055-16057.
- Yamato, I., and Anraku, Y.** (1990) Mechanism of Na⁺/proline symport in *Escherichia coli*: reappraisal of the effect of cation binding to the Na⁺/proline symport carrier. *J. Membr. Biol.* 114: 143-151.
- Yamato I.** (1992) Ordered binding model as a general mechanistic mechanism for secondary active transport systems. *FEBS Lett.* 298: 1-5.
- Yamato, I., and Anraku, Y.** (1993) Na⁺/substrate symport in prokaryotes. In *Alkali Cation Transport Systems in Prokaryotes* (Bakker, E. P., ed.) pp. 53-76, CRC Press, Inc., Boca Raton, FL.
- Yernool, D., Boudker, O., Jin, Y., and Gouaux, E.** (2004) Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. *Nature* 431: 811-818.
- Yin, Y., He, X., Szewczyk, P., Nguyen, T., and Chang, G.** (2006) Structure of the multidrug transporter EmrD from *Escherichia coli*. *Science* 312: 741-744.
- Zhou, A., Wozniak, A., Meyer-Lipp, K., Nietschke, M., Jung, H., and Fendler, K.** (2004) Charge translocation during cosubstrate binding in the Na⁺/proline transporter of *E. coli*. *J. Mol. Biol.* 343: 931-942.

CHAPTER 2***Role of Ser340 and Thr341 in transmembrane domain IX of the Na⁺/proline transporter PutP of Escherichia coli in ligand binding and transport*****2.1. Abstract**

The Na⁺/solute symporter family comprises more than 400 members of pro- and eukaryotic origin. Using the Na⁺/proline transporter PutP of *Escherichia coli* as a model, the role of two conserved residues, Ser340 and Thr341, is investigated to obtain insights into the mechanism of transport catalyzed by members of this family. Substitution of these amino acids alters the transport kinetics of cells and proteoliposomes containing the PutP variants significantly. In particular, the apparent affinities for Na⁺ and Li⁺ are reduced by two orders of magnitude or more. Also proline binding is affected, albeit to a lesser extent than ion binding. Thereby, the presence of a hydroxyl group at position 341 is essential for high affinity ligand binding. Furthermore, Cys placed at position 340 or 341 reacts with sulfhydryl reagents of different polarity, indicating accessibility from the water phase. In addition, Cys cross-linking suggests proximity of the residues to other amino acids previously shown to be crucial for ligand binding. For these reasons it is suggested that Ser340 and Thr341 are located in a ligand-translocation pathway. Furthermore, it is proposed that the side chain of Thr341 directly participates in Na⁺ binding.

2.2. Introduction

Members of the Na⁺/solute symporter (SSS) family (TC 2.A.21, SLC5) are found in archaea, bacteria, and eukaryotes (Jung, 2002; Jung *et al.*, 2006; Wright and Turk, 2004; Reizer *et al.*, 1994). They have the capability to couple electrochemical Na⁺ gradients with the transport of solutes like glucose, nucleosides, proline, pantothenate, or iodide. Some transporters of the SSS family are implicated in human disease (e.g. the Na⁺/glucose transporter SGLT1) or play an important role in medical therapy (e.g. the human Na⁺/iodide symporter NIS) (Dohán *et al.*, 2006; Wright *et al.*, 2007). Furthermore, bacterial transporters like the Na⁺/proline transporter PutP of *Staphylococcus aureus* contribute to virulence (Schwan *et al.*, 2006).

A crystal structure is not available yet for any member of the SSS family. Beside intensive trials to crystallize SSS proteins, it is obvious that information on structure-function relationships is required in addition to structure to understand the mechanism of respective transport processes at the molecular level. For this purpose we utilize PutP of *Escherichia coli* as a model since it can easily be genetically manipulated and produced in milligram amounts (Jung *et al.*, 1998a; Jung *et al.*, 1998b).

PutP transports Na⁺ and proline with a 1:1 stoichiometry following an ordered binding mechanism at low proline concentrations, whereas at high concentrations binding can be unordered (Chen *et al.*, 1985; Zhou *et al.*, 2004). Current experimental data indicate that the transporter is composed of 13 transmembrane domains (TMs), a motif suggested being common to all SSS proteins (Jung *et al.*, 1998b; Jung, 1998; Turk and Wright, 1997; Wegener *et al.*, 2000; Jeschke *et al.*, 2004). Furthermore, TM II of PutP was recently proposed to form part of the substrate-translocation pathway. The TM contains residues Asp55, Ser57, and Gly58, implicated in Na⁺ and/or proline binding (Pirch *et al.*, 2002; Quick *et al.*, 1996; Quick and Jung, 1997). In addition, it participates in the formation of a water-filled cleft that is open to the cytoplasm in the absence of ligands and closed by proline binding in the presence of Na⁺ (Pirch *et al.*, 2003).

Inspired by the three-dimensional structure of the leucine transporter LeuT_{Aa} (Na⁺/neurotransmitter symporter (NSS) family), which revealed an involvement of two adjacent polar residues (Thr354, Ser355) in Na⁺ binding (Yamashita *et al.*, 2005), PutP was screened for a similar arrangement. Here, first evidence is presented indicating that TM IX of PutP is located close to TM II and contains residues Ser340, Thr341 that are particularly important for high affinity Na⁺ and proline binding. It is suggested that Ser340 and Thr341 are located in a ligand-translocation pathway; thereby, Thr341 probably directly participates in Na⁺ binding.

2.3. Materials and Methods

2.3.1. Bacterial strains and plasmids

E. coli JM109 [*endA1 recA1 gyrA96 thi hsdR17 supE44 relA1 Δ(lac-proAB)* (F⁺*traΔ36pro AB⁺ lacI^f ZΔM15*)] (Yanisch-Perron *et al.*, 1985) and *E. coli* DH5α [F⁺ Φ80d *lacZ ΔM15 Δ(lacZYA-argF) U169 deoR recA1 endA1 hsd R17(r_k⁻,m_k⁺) phoA supE44 λ⁻ thi-1 gyrA96 relA1*] were used as carriers for the plasmids generated in this study. *E. coli* WG170 [F⁺ *trp lacZ rpsL thi Δ(putPA)101 proP219*] (Stalmach *et al.*, 1983), harboring given plasmids, was used for overexpression of the *putP* gene and transport assays. The following plasmids, derivatives of pT7-5 (Tabor and Richardson, 1985) containing the *lac* promoter/operator for expression of the *putP* gene, were used for all gene manipulations; pT7-5/*putP* and pT7-5/*putP*(ΔCys), each of which harboring a cassette version of the *putP* gene encoding PutP-wild-type and an engineered transporter devoid of all five native Cys residues [PutP(ΔCys)], respectively, and a C-terminal-attached amino acid sequence resembling the FLAG epitope and a His₆ tag (Jung *et al.*, 1998a). Vector pTrc99a (Amann *et al.*, 1988) was used for overexpression.

2.3.2. Site-directed mutagenesis

Desired nucleotide substitutions in *putP* were generated by PCR with Taq-DNA polymerase using plasmid pT7-5/*putP* as a template and synthetic mutagenic oligonucleotides. Resulting PCR fragments were digested with *Pst*I and *Afl*III and ligated with similarly treated plasmid pT7-5/*putP* or pT7-5/*putP*(Δ Cys) additionally incubated with alkaline phosphatase. The resulting plasmid DNA was verified by sequencing using an ABI 377 HT device. Double-Cys mutants were generated by restriction fragment replacement with *Pst*I and *Spe*I. For overexpression, the mutated *putP* gene was cut out with restriction endonucleases *Nco*I and *Hind*III and ligated to similarly treated vector pTrc99a additionally incubated with alkaline phosphatase.

2.3.3. Proline transport in intact cells

Active transport was measured in *E. coli* WG170 (PutP^A) harboring derivatives of plasmids pT7-5/*putP* or pT7-5/*putP*(Δ Cys) encoding PutP with given amino acid replacements. The cells were grown aerobically in Luria-Bertani (LB) medium (Miller, 1992) containing 100 μ g/ml ampicillin at 37 °C. Overnight cultures were diluted 25-fold and were allowed to grow up to an optical density at 420 nm (A_{420}) of 1.0 followed by induction with 0.5 mM isopropyl 1-thio- β -D-galactopyranoside for 2 h. Cells were harvested by centrifugation at 13,200 \times g for 10 min and washed up to six times with 100 mM Tris/Mes, pH 6.0, at 4 °C to reduce the Na⁺ contamination below 5 μ M. For transport assays cells were resuspended in the same buffer and adjusted to a total protein concentration of 0.35 mg/ml. Transport of 10 μ M L-[U-¹⁴C]proline (26 Ci/mol) (if not otherwise indicated) was assayed under standard test conditions in the presence of 20 mM D-lactate (Na⁺ salt) and 50 mM NaCl. Transport assays were terminated at various time points (0, 0.17, 0.5, 1, 2, 5, 10, 30 min) using the rapid filtration method as described (Quick and Jung, 1997). Initial rates of transport were calculated from the initial linear portion of the time course and steady-state levels of proline accumulation were taken from time points after leveling off of the uptake curve. Standard deviations (SD) were determined from at least three independent experiments.

2.3.4. Immunological analysis

Relative amounts of PutP with given amino acid replacements in membranes of *E. coli* WG170 were estimated by Western blot analysis with horseradish peroxidase-linked mouse anti-FLAG IgG directed against the FLAG epitope at the C terminus of each PutP variant as described before (Quick *et al.*, 1996).

2.3.5. Cys accessibility analyses

The experiments were performed with right-side-out membrane vesicles prepared from *E. coli* WG170 transformed with pTrc99a/*putP*(Δ Cys) harboring given mutations following the protocol of Kaback (1971).

Before Cys labeling, the right-side-out vesicle suspension (10 mg total membrane protein/ml) was reduced with 1 mM dithiothreitol or 1 mM Tris(2-carboxyethyl)phosphine hydrochloride. The vesicles were washed and resuspended in 100 mM Tris/Mes buffer, pH 6.0. Two hundred- μ l aliquots of the suspension (10 mg of total membrane protein/ml) were incubated with 500 μ M [14 C]NEM (8 μ Ci/ μ mol) at 25 °C for 10 min [PutP(Δ Cys)-T341C] or 15 min [PutP(Δ Cys)-S340C]. The reaction was stopped by the addition of 2 mM dithiothreitol. When indicated, 500 μ M methanethiosulfonate ethylammonium (MTSEA), methanethiosulfonate ethylsulfonate (MTSES), or methyl methanethiosulfonate (MMTS) were added before [14 C]NEM. Similarly, the effect of PutP ligands on Cys accessibility was analyzed by incubation with 50 mM Na⁺, Li⁺, and/or 10 mM proline Na⁺/proline before [14 C]NEM labeling. After labeling, PutP was solubilized by the addition of 1.5% β -dodecyl-maltoside under stirring at 4 °C for 30 min. Then PutP was purified via Ni²⁺-NTA affinity chromatography as described before (Pirch *et al.*, 2002). After protein determination, equal amounts of protein were subjected to a 10% SDS-PAGE. Gels were stained with Coomassie Blue and scanned for estimation of relative amounts of protein. Radioactivity was detected with a PhosphorImager SI of GE Healthcare. The software ImageQuant was used for quantitative analysis of the images.

2.3.6. In situ site-directed cross-linking

For cross-linking experiments, 0.4 liters of *E. coli* WG170 transformed with pT7-5/*putP*(Δ Cys) encoding PutP with single- or double-Cys replacements and an introduced double thrombin cleaving site were grown and harvested as described above. Inverted membrane vesicles were prepared as described (Jung *et al.*, 1998a) with 1 mM Tris(2-carboxyethyl)phosphine hydrochloride as the reducing compound.

Before cross-linking, the membrane vesicles were washed twice and resuspended in 50 mM KP_i buffer, pH 7.0. One hundred μ g total membrane protein (1 mg total membrane protein/ml) were incubated with 0.5 mM Cu(1,10-phenanthroline)₃ (CuPh) or with 2 mM concentrations of one of the homobifunctional cross-linker 1,6-bis-(maleimido) hexane (BMH) or *N,N*-*p*-phenylenedimaleimide (*p*-PDM). Stock solutions of 30 mM CuPh (in ethanol) and 200 mM concentrations of the maleimide reagents (in Me₂SO) were prepared. Reactions were performed at 4 °C for 60 (CuPh) or 120 min (maleimide reagents) and at 25 °C for 30 or 60 min to estimate the dependence of cross-linking efficiency on thermal backbone motions.

After termination with 10 mM EDTA (CuPh) or 10 mM dithiothreitol (BMH and *p*-PDM), the inverted membrane vesicles were subjected to ultracentrifugation at 250,000 x *g* for 30 min at 4 °C. The membranes were resuspended in thrombin cleavage buffer (Invitrogen) containing 1% β -dodecyl-maltoside and, after a 2-fold dilution in thrombin cleavage buffer, digested overnight at 4 °C with 1.3 units of thrombin protease. The digested proteins were subjected to non-reducing SDS-PAGE and immunoblot analysis as described above.

2.3.7. Purification and reconstitution of PutP

PutP was purified by Ni²⁺-NTA affinity chromatography and reconstituted into proteoliposomes prepared from *E. coli* polar lipid extracts (Avanti Polar Lipids, Inc., Alabaster, Alabama) at a lipid to protein ratio of 100 to 1 (w/w) as described before (Jung *et al.*, 1998a).

2.3.8. Transport assay with proteoliposomes

L-[¹⁴C]proline uptake was measured with proteoliposomes containing PutP(FH) (~2 mg of protein/ml) preloaded with 100 mM KP_i, pH 7.5, 2 mM β -mercaptoethanol, 5 mM MgSO₄. The proteoliposomes were extruded through a 400 nm filter before use. Transport was started by a 400-fold dilution of aliquots of the proteoliposome suspension into the desired buffer. The following buffers were used: Na⁺-free 100 mM Tris/Mes, pH 7.5, 2 mM β -mercaptoethanol, 5 mM MgSO₄ containing 50 mM NaCl (creation of $\Delta\mu_{\text{Na}^+}$) or 50 mM LiCl (creation of $\Delta\mu_{\text{Li}^+}$) or no further additions (creation of $\Delta\Psi$); Na⁺-free 100 mM Tris, Mes, pH 6.0, 2 mM β -mercaptoethanol, 5 mM MgSO₄ (creation of $\Delta\mu_{\text{H}^+}$) or 100 mM KP_i, pH 7.5, 2 mM β -mercaptoethanol, 5 mM MgSO₄ (facilitated diffusion, control). All buffers contained 10 μ M L-[¹⁴C]proline (26 Ci/mol) and 0.2 μ M valinomycin. For the measurement of proline counterflow (cfl), proteoliposomes preloaded with 10 mM L-proline were diluted 400-fold into 100 mM KP_i, pH 7.5, 2 mM β -mercaptoethanol containing L-[¹⁴C]proline at a final concentration of 25 μ M (15 Ci/mol). Transport assays were terminated at various time points (0, 0.17, 0.5, 1, 2, 5, 10 min) by quenching of the reaction with 3 ml of ice-cold 100 mM KP_i, pH 6.6, 100 mM LiCl and immediate filtration using Millipore filters type GSTF 02500, 0.2- μ m pore size as described (Jung *et al.*, 1998a). Initial rates of transport were determined from the initial linear part of the transport curve. SD were determined from at least three independent experiments.

2.3.9. Determination of Na⁺

Na⁺ concentrations in buffers used for transport assays and Cys accessibility analyses were determined with a Varian AA240 atomic absorption spectrometer.

2.3.10. Protein determination

Determination of protein was performed according to a modified Lowry method (Peterson, 1977) for total membrane protein, according to Bradford (1976) for detergent-solubilized protein, and by the Amido Black method (Schaffner and Weissmann, 1973) for protein in proteoliposomes.

2.4. Results

2.4.1. Conserved amino acids of TM IX

Searching for amino acid residues of potential structural and/or functional significance (e.g. residues involved in ligand binding), a multiple sequence alignment of members of the SSS family was performed. The analysis revealed that Ser340 and Thr341 of *E. coli* PutP are conserved within prokaryotic and eukaryotic family members irrespective of their substrate specificity (e.g. transporters for proline, pantothenate, sugars, or iodide). In fact, the transporters analyzed contained Ser at the position corresponding to Ser340 and Ser or Thr at the position corresponding to Thr341 of *E. coli* PutP (suppl. Fig. 2.1.). The high degree of conservation suggests that both residues fulfill similar functions in all related transporters. In addition, a similar amino acid arrangement (Thr354, Ser355) was found to be involved in Na⁺ binding in the non-homologous leucine transporter LeuT_{Aa} (Yamashita *et al.*, 2005). Therefore, we set out to elucidate the role of Ser340 and Thr341 in PutP of *E. coli*.

2.4.2. Effect of the substitution of Ser340 and Thr341 on proline uptake and putP expression in *E. coli* WG170

Ser340 and Thr341 were individually replaced by Ala, Cys, Thr (Ser340) or Cys, Ser, Val (Thr341). In addition, double mutants S340A/T341C and S340A/T341V were generated.

In a first step, time courses of L-proline uptake into *E. coli* WG170 were recorded under standard test conditions (70 mM Na⁺, 10 μM L-[¹⁴C]proline) (Fig. 2.1.a). Based on the observed activities, the mutants were divided into three groups: I) mutants with moderately reduced initial rates (S340A 20%, S340T 40%, T341S 60% of the wild-type rate) but steady-state values of L-proline accumulation close to wild-type (80-100% of the wild-type value); II) mutants with highly reduced initial rates (S340C 9%, T341C 3%, T341V 3.5% of the wild-type rate) and reduced steady-state values (S340C 60%, T341C 35%, T341V 20% of the wild-type value); III) mutants with only marginal levels of activity (<1% of the wild type) (S340A/T341C, S340A/T341V). Western blot analysis of membranes prepared from cell suspensions used for the transport assays indicated that the observed reduced proline uptake rates were not because of reduced amounts of transporter molecules in the membrane (Fig. 2.1.b).

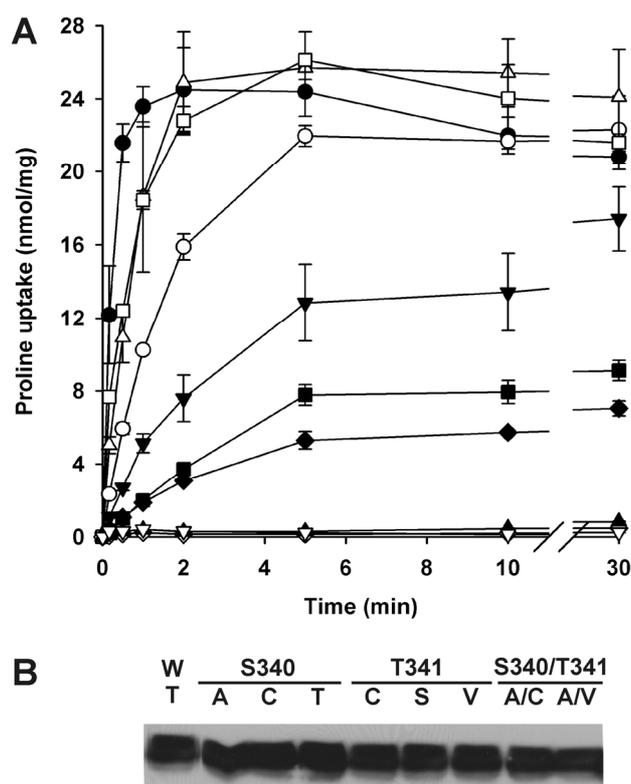


Figure 2.1.: Influence of amino acid replacements in TM IX on proline transport and *putP* expression. **A**, time courses of proline uptake into intact cells of *E. coli* WG170 were assayed in the presence of 10 μM L-[U- ^{14}C]proline, 50 mM NaCl, and 20 mM D-lactate (Na-salt) as the electron donor at 25 °C under aerobic conditions. At given time points transport was terminated by rapid filtration and filter-bound radioactivity was determined (Quick *et al.*, 1996). Shown are the mean values and SD of at least three measurements. **B**, *putP* expression was estimated by Western blot analysis. 25 μg of total cell protein from each culture were separated by 10% SDS-PAGE. Proteins were then transferred onto a nitrocellulose membrane (0.45- μm pore size) and probed with horse-radish peroxidase-linked mouse anti-FLAG M2 monoclonal antibody. Detection was performed according to the enhanced chemiluminescence method. ●, wild type (W/T); ○, S340A (A); ▼, S340C (C); △, S340T (T); ■, T341C (C); □, T341S (S); ◆, T341V (V); ◇, S340A/T341C (AC); ▲, S340A/T341V (AV); ▽, pT7-5.

To test whether the inhibitory effects of the Cys substitution at position 340 or 341 were due to intra- or intermolecular disulfide bridge formation (e.g. with one of the five native Cys residues), the Cys substitutions were introduced into PutP(ΔCys). Recording of time courses of transport yielded similar inhibitory effects as observed in the wild-type background (data not shown). The results suggested that the altered amino acid side chains *per se* and not subsequent reactions caused the detected transport defects.

Taken together the results indicated that the nature of the amino acid side chains at positions 340 and 341 of PutP of *E. coli* was crucial for the proline uptake rate. Highest rates were observed when hydroxyl groups were retained at both positions (e.g. S340T, T341S). Overall, the steady-state value of proline accumulation was much less affected than the uptake rate.

2.4.3. Influence of the substitutions on the ion selectivity of PutP

To test for possible changes of the ion selectivity, the energetic requirements of the transport process were investigated under more defined conditions using purified and

reconstituted transport protein. It should be noted that the protocol used for reconstitution yields proteoliposomes containing PutP in an inside-out orientation (Jung *et al.*, 1998a).

First, determination of the initial rates of Na⁺-coupled (sodium motive force (*smf*)-driven) transport in proteoliposomes confirmed in principle the defects described above for intact cells, although individual values varied to some extent (Fig. 2.2.). Mutants of group I (S340A, S340T, T341S) showed again the highest activities. However, differing from intact cells, neither the S340A nor the T341S substitution caused significant inhibition of transport. This difference may be attributed to the previously described functional asymmetry of the transporter (Jung *et al.*, 1998a; Yamato and Anraku, 1993). The initial rates of the mutants T341C and T341V of group II were highly reduced (about 2.5% of the wild type) as in intact cells. PutP-S340C (9% residual activity in intact cells) showed only marginal activity (<1% of the wild type) in proteoliposomes (Fig. 2.2.). Reducing conditions were maintained during all steps of the preparation to prevent a potential participation of Cys340 in formation of a disulfide bridge with a native Cys. Nevertheless, oxidation of Cys340 during PutP preparation could have added to the deleterious effect of the substitution in proteoliposomes. Finally, activities of mutants of group III (S340A/T341C, S340/T341V) were reduced to marginal values (<2% of the wild type) similar as observed with intact cells (not shown).

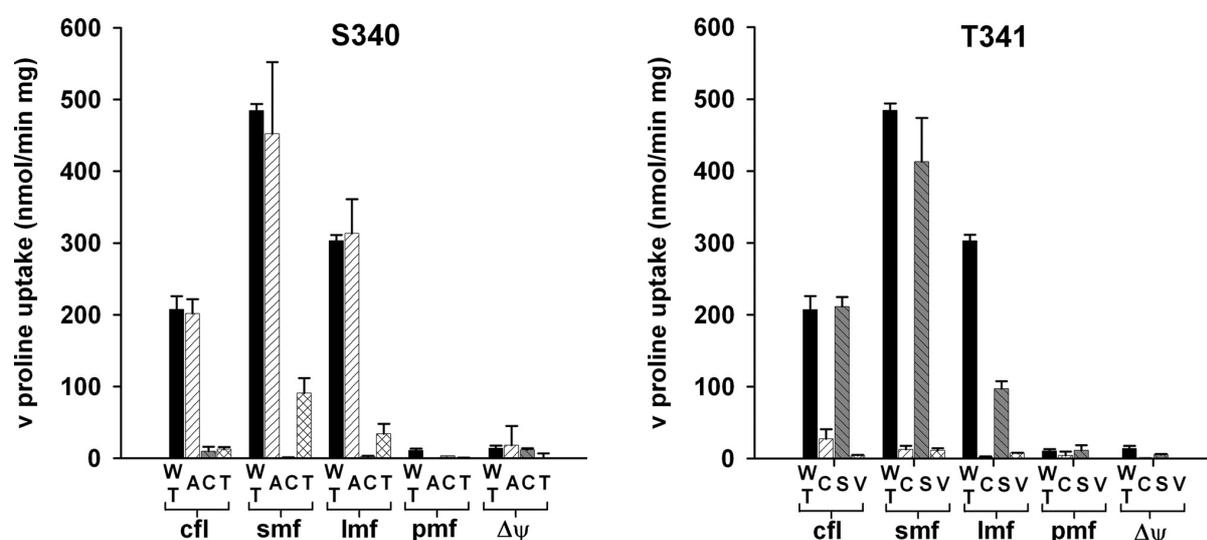


Figure 2.2.: Ion dependence of proline uptake into proteoliposomes containing PutP with given replacements. PutP was purified and reconstituted as described previously (Jung *et al.*, 1998a). Proteoliposomes in 100 mM KP_i, pH 7.5, were diluted 400-fold into 100 mM Tris/Mes, pH 7.5, 2 mM β-mercaptoethanol, 5 mM MgSO₄, 0.2 μM valinomycin, 10 μM [¹⁴C]proline. The following additions were made to generate inwardly directed electrochemical ion gradients: 50 mM NaCl for *smf* or 50 mM LiCl for *lmf* or no further additions for an electrical potential only (Δψ). In addition, proteoliposomes were diluted into 100 mM Tris/Mes, pH 6.0, containing 0.2 μM valinomycin to generate a proton motive force (*pmf*). Dilution into 100 mM KP_i, pH 7.5, served as a negative control. For the measurement of proline counterflow (*cfl*), proteoliposomes preloaded with 10 mM L-proline were diluted 400-fold into 100 mM KP_i, pH 7.5, 2 mM β-mercaptoethanol containing L-[¹⁴C]proline at a final concentration of 25 μM. Transport was stopped by rapid filtration and analyzed as described under "Experimental Procedures." A, S340A; C, S340C; T, S340T; C, T341C; S, T341S; V, T341V.

Secondly, analysis of various energetic conditions revealed that in all cases transport was strictly dependent on an inwardly directed electrochemical Na⁺ (*smf*) or Li⁺ gradient (lithium motive force (*lmf*)) (Fig. 2.2.). ΔΨ-Driven transport of the mutants was only marginal and most likely due to Na⁺ impurities of buffer components (leading to Na⁺ concentrations in

the assay of about 5-10 μM). Establishment of a H^+ gradient in addition to $\Delta\Psi$ (proton motive force-driven transport) did not further stimulate transport. These measurements indicated that the replacements at position 340 or 341 did not cause a dramatic change in ion selectivity (e.g. from Na^+ - to H^+ -coupled transport). However, comparison of *smf*- and *lmf*-coupled transport of PutP-T341S (*smf* >90% and *lmf* 30% of wild-type activity) hints at subtle alterations of the ion preferences of the transporter (Fig. 2.2.).

Besides ion-coupled transport, the impact of the substitutions on proline counterflow was analyzed. This transport mode neither required a membrane potential nor an ion gradient. Instead, accumulation of externally applied L-[^{14}C]proline in proteoliposomes was driven by downhill movement of unlabeled L-proline from preloaded proteoliposomes. Substitution of Ser340 (by Cys or Thr) or Thr341 (by Cys or Val) led to significantly reduced counterflow activities (remaining activities below 13% of the wild-type value) (Fig. 2.2.). Exceptions were PutP-S340A and PutP-T341S, which catalyzed counterflow almost like the wild type. The latter results confirmed the functional significance of Ser340 and Thr341 and suggested that these residues are important for proline binding or transport.

2.4.4. Kinetic analyses

More detailed kinetic analyses revealed substantial defects in proline AND Na^+ binding. Thereby, binding parameters of group II mutants were more severely affected than the ones of group I mutants. The maximum rate of proline uptake was in general less dramatically affected (reduced to maximum 50% of wild-type V_{max}) than binding parameters (Table 2.1.; Fig. 2.3.). Transport by group III mutants (S340A/T341V; S340A/T341C) was not significantly stimulated by increasing ligand concentrations. In more detail, kinetic analysis with intact cells demonstrated that removal of the hydroxyl group at position 340 or 341 decreased the apparent affinity for proline 15-60-fold in comparison to wild type, suggesting a defect in proline binding (Table 2.1., Fig. 2.3.). Only PutP-S340T and PutP-T341S yielded $K_{m(\text{pro})}$ values similar to that of the unaltered transporter. The data indicated that hydroxyl groups at positions 340 and 341 were crucial for proline binding.

The apparent affinities of PutP for the coupling ions Na^+ and Li^+ ($K_{0.5(\text{Na}^+)}$ and $K_{0.5(\text{Li}^+)}$) were even more severely affected by the amino acid replacements than the proline affinity. So, for all three PutP-S340 variants, 100-1000-fold increases of $K_{0.5(\text{Na}^+)}$ and $K_{0.5(\text{Li}^+)}$ were determined (Table 2.1., Fig. 2.3.). The data did not hint at a specific role of the hydroxyl group at position 340 in ion binding since Ala or Thr at position 340 led to about the same decrease of the apparent ion affinity (Fig. 2.3.). Furthermore, although PutP-T341S showed apparent ion affinities close to the wild type, proline uptake into cells catalyzed by group II mutants PutP-T341C and -T341V was no longer significantly stimulated by Na^+ (T341V also

not by Li^+). Therefore, contrary to position 340, high affinity ion binding required a hydroxyl group at position 341.

Table 2.1.: Proline uptake kinetics of PutP bearing replacements of given amino acid residues. To determine $K_{m(\text{pro})}$ and V_{max} initial rates of L-[U- ^{14}C]proline uptake by *E. coli*, WG170 producing either PutP wild type or PutP with given replacements were measured in the presence of 50 mM NaCl and 20 mM D-lactate (Na^+ salt) at proline concentrations from 0.2 to 500 μM . For determination of the ion-specific parameters ($K_{0.5(\text{Na}^+)}$, $K_{0.5(\text{Li}^+)}$), transport of 10 μM L-[U- ^{14}C]proline was measured in the presence of 0.005-250 mM NaCl or LiCl at 25 °C. The resulting data were plotted according to Eadie-Hofstee.

PutP	$K_{m(\text{pro})}$ μM	V_{max} $\text{nmol/min} \times \text{mg of protein}$	$K_{0.5(\text{Na}^+)}$ mM	$K_{0.5(\text{Li}^+)}$ mM
wild-type	2.1 \pm 0.2	44.5 \pm 3.9	0.038 \pm 0.007 (0.7 \pm 0.09) ^a	0.295 \pm 0.057
S340A	31.3 \pm 3.7	39.5 \pm 7.1	5.3 \pm 1.0 (5.2 \pm 0.49)	53.0 \pm 7.5
S340C	46.3 \pm 7.8	53.7 \pm 2.6	40.1 \pm 5.7	32.3 \pm 7.9
S340T	0.7 \pm 0.2	19.7 \pm 5.8	3.3 \pm 0.3 (11.9 \pm 1.7)	32.1 \pm 2.8
T341C	60.8 \pm 7.8	21.6 \pm 0.8	NS ^b (>250)	>100
T341S	3.0 \pm 1.0	40.2 \pm 11.2	0.051 \pm 0.001	2.0 \pm 0.29
T341V	117.1 \pm 4.6	27.5 \pm 2.0	NS ^b (>250)	NS

^a Values in parentheses were obtained with proteoliposomes containing the purified PutP variants.

^b NS, no stimulation of proline uptake into intact cells by increasing concentrations of Na^+ or Li^+ .

Because the lack of Na^+ (Li^+) stimulation of proline uptake by the latter group II mutants could be due to insufficient proline binding, the proline concentration in the ion affinity assay was raised from 10 to 250 μM . However, substitution of Thr341 by Cys or Val again prevented a significant stimulation of transport by Na^+ (and in case of T341V, also by Li^+). The same elevation of the proline concentration increased the apparent Na^+ and Li^+ affinities of PutP-wild type and -S340C about 10-fold (other mutants not tested). These data suggested cooperative interactions between ion- and proline-binding sites. Importantly, ion binding remained impaired also in the presence of saturating proline concentrations, suggesting that the observed defects in ion binding were not secondary to proline binding defects.

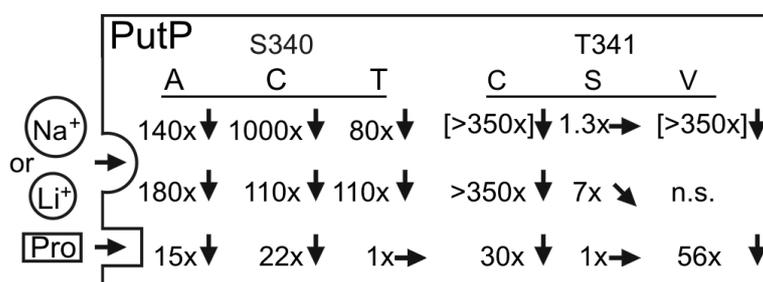


Figure 2.3.: Influence of amino acid substitutions on apparent ion and proline affinities of PutP. Given are changes of apparent affinity parameters for proline, Na^+ and Li^+ ($K_{m(\text{pro})}$, $K_{0.5(\text{Na}^+)}$, $K_{0.5(\text{Li}^+)}$) as multiples of the corresponding wild-type value. Calculation is based on the kinetic data of Table 2.1.. Values in brackets were obtained with proteoliposomes containing the purified PutP variants since proline uptake into intact cells was not significantly stimulated by Na^+ . n.s., no stimulation of proline uptake into intact cells of *E. coli* WG170 by increasing Li^+ concentrations. A, S340A; C, S340C; T, S340T; C, T341C; S, T341S; V, T341V.

Because the ion dependence of PutP-T341C and -T341V in intact cells remained unclear, the apparent Na^+ affinities were analyzed in the more defined proteoliposome system. Here, proline uptake into proteoliposomes containing either purified PutP-T341C or -

T341V was finally stimulated by Na⁺ concentrations of 50 mM or higher, and the $K_{0.5(\text{Na}^+)}$ was estimated to exceed 250 mM. A precise determination of $K_{0.5(\text{Na}^+)}$ was not possible since NaCl concentrations in the molar range inhibited PutP-dependent transport in proteoliposomes (Racher *et al.*, 1999). In contrast to the Cys and Val substitution, PutP with Ser in place of Thr341 showed transport kinetics close to wild type. These results confirmed the particular significance of a hydroxyl group at position 341 for high affinity Na⁺ binding.

Contrary to position 341, evidence for a specific function the hydroxyl group at position 340 was not obtained with the proteoliposome system. Rather, size and polarity of the residue mattered for transport as already observed for the uptake kinetics of intact cells. Therefore, the kinetic data obtained with proteoliposomes confirmed the results with intact cells (compare above).

Taken together, the kinetic analyses indicated that Ser340 and Thr341 were essential for the high affinities of PutP for coupling ion and substrate. Furthermore, the data suggested a particular crucial role of the hydroxyl group at position 341 in binding of Na⁺ or Li⁺.

2.4.5. Accessibility to sulfhydryl reagents

To determine whether positions 340 and 341 are in contact with the water phase, the accessibility of Cys placed at either one of these positions in PutP(Δ Cys) to sulfhydryl reagents of various sizes and polarities was analyzed. Labeling experiments were performed with right-side-out membrane vesicles containing the corresponding PutP variant.

In initial studies it was shown that PutP(Δ Cys)-S340C and PutP(Δ Cys)-T341C readily react with [¹⁴C]NEM. Reaction of Cys340 resulted in about 60% labeling of the protein within 10 min. Reaction of the maleimide with Cys341 was faster, yielding about 100% protein labeling within the same time. These data suggested a somewhat more buried position of Cys340. In contrast, the negative control PutP(Δ Cys) did not react with [¹⁴C]NEM under the conditions of the experiment even after extended periods of time (Fig. 2.4.a, b). Because reaction of the maleimide required deprotonation of the sulfhydryl group, the results indicated a polar environment at positions 340 and 341. The latter conclusion was strongly supported by the observation that charged reagents methanethiosulfonate ethylammonium (MTSEA) and methanethiosulfonate ethylsulfonate (MTSES) as well as neutral methyl methanethiosulfonate (MMTS) efficiently inhibited [¹⁴C]NEM labeling of Cys at both positions (Fig. 2.4.b). Taken together, the data suggested that Ser340 and Thr341 are located in a water-filled cavity in the membrane that is at least transiently open to the periplasm. Finally, the influence of PutP ligands on Cys modification by [¹⁴C]NEM was investigated. Repeated experiments did not reveal a significant effect of Na⁺, Li⁺, proline, or a combination thereof on the reaction (data not shown).

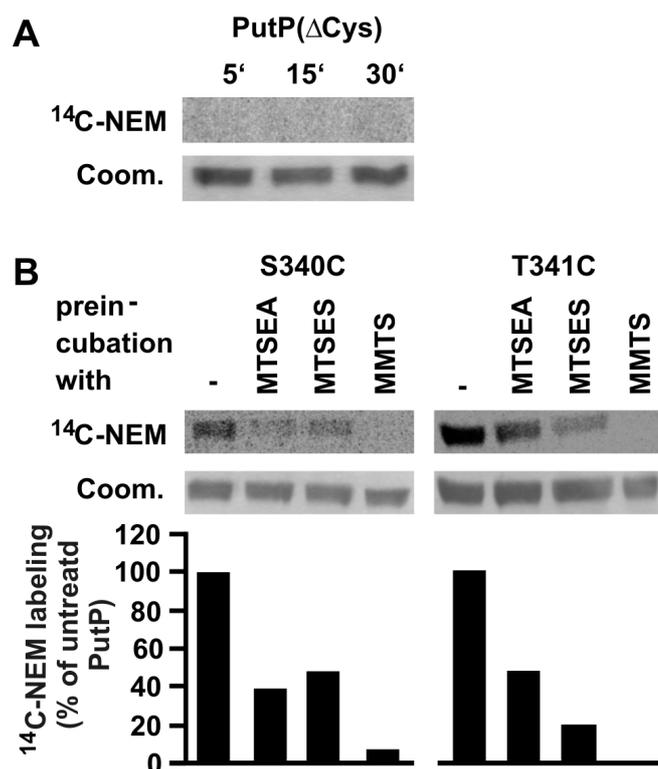


Figure 2.4.: Accessibility of Cys at positions 340 and 341 of PutP to sulfhydryl reagent. 200 μ l of a suspension of right-side-out membrane vesicles of *E. coli* WG170 containing given PutP derivatives (10 mg of total membrane protein/ml) were incubated with 500 μ M [14 C]NEM for given periods of time. Labeling reactions were stopped by the addition of 2 mM β -mercaptoethanol. After solubilization with 1.5% *n*-dodecyl- β -maltoside and purification via Ni $^{2+}$ -NTA affinity chromatography, 5 μ g of protein were loaded onto a 10% SDS-PAGE and separated. Radioactivity was detected with a PhosphorImager SI of GE Healthcare. Gels were stained with Coomassie Blue (*Coom.*) and scanned for an estimation of relative amounts of protein. The software ImageQuant was used for quantitative analysis of the images. **A**, time course of [14 C]NEM labeling of PutP(Δ Cys), which was used as a negative control. **B**, labeling of PutP(Δ Cys)-S340C or -T341C. The [14 C]NEM labeling reaction was performed for 10 min. The right-side-out vesicles were preincubated with 500 μ M methanethiosulfonate ethylammonium (*MTSEA*; positive charge), methanethiosulfonate ethylsulfonate (*MTSES*; negative charge), or methyl methanethiosulfonate (*MMTS*; neutral) for 10 min where indicated. The diagram gives the percentage of [14 C]NEM labeling after preincubation with the methanethiosulfonate (*MTS*) reagents, thereby labeling in the absence of *MTS*-compounds was set to 100%.

2.4.6. Proximity between TM II and TM IX

Because the dramatic changes of the apparent Na $^{+}$ and proline affinities suggested an involvement of TM IX in ligand binding, a possible proximity between residues of TM IX and the functionally important TM II was tested by *in situ* Cys cross-linking. For this purpose the following pairs of Cys residues were individually introduced in TMs II and IX of PutP(Δ Cys): S41C/C344, D55C/S340C and D55C/T341C, S57C/S340C and S57C/T341C, M62C/A327C. In addition, thrombin cleavage sites were introduced into periplasmic loop 7 (pL7) to facilitate determination of cross-linking products (compare the supplemental material for details on the establishment of protease cleavage sites and activities of all PutP variants used for cross-linking).

Cross-linking of Cys in TMs II and IX was performed with BMH (3.47-15.64 Å), *p*-PDM (9.2-12.29 Å), and CuPh (zero-length) (Green *et al.*, 2001) (Fig. 2.5.). Thrombin digestion of PutP variants in the absence of cross-linker resulted in C-terminal fragments of the expected

apparent size of 30 kDa, whereas cross-linked and cleaved PutP migrated at the size of intact PutP (44 kDa) (Fig. 2.5.).

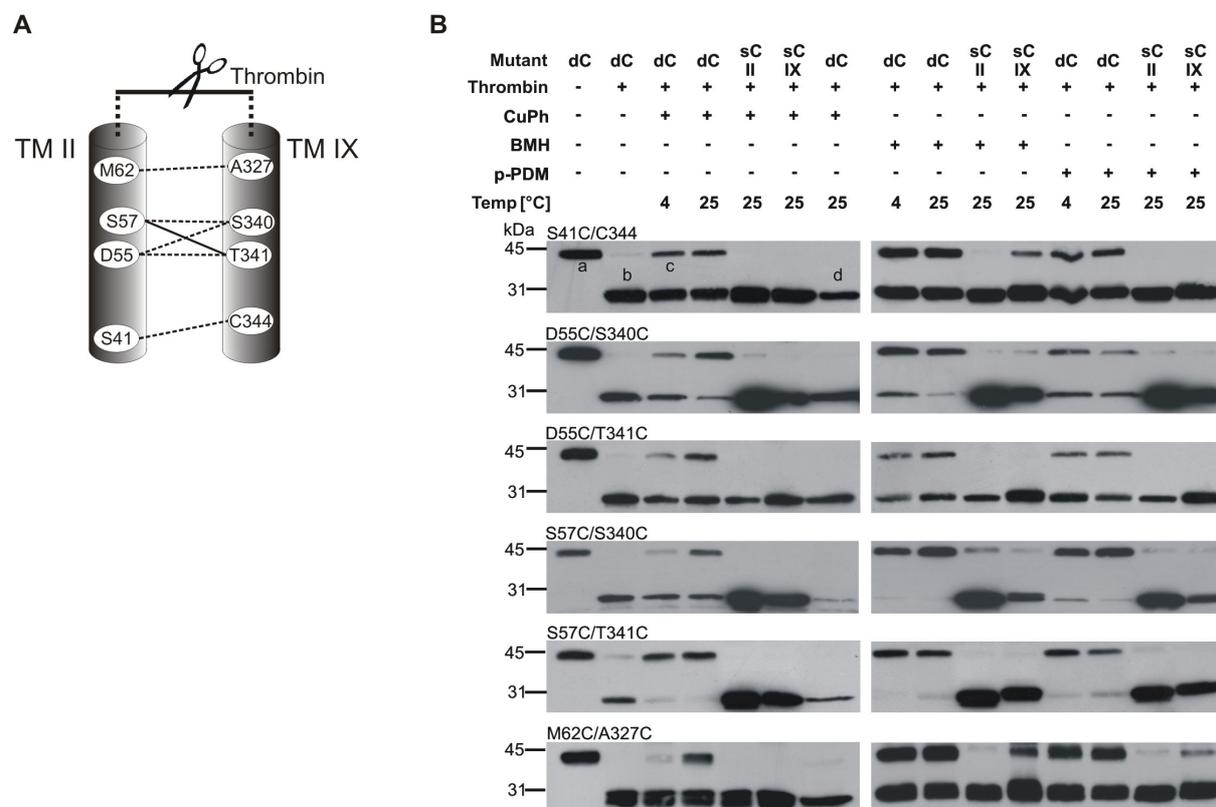


Figure 2.5.: *In situ* cross-linking of TM II and TM IX of PutP. A, model of TMs II and IX showing residues successfully cross-linked. The *solid line* connects positions most efficiently cross-linked with CuPh; the *dashed lines* mark positions less efficiently cross-linked with CuPh. Two thrombin cleavage sites were inserted into pL7 to facilitate detection of intramolecular cross-linking. B, primary results of the cross-linking analysis. 100 μ l of membranes (1 mg of total membrane protein/ml) containing the given PutP derivatives were subjected to 0.5 mM zero-length cross-linker CuPh (*third through seventh vertical lanes*) or to 2 mM maleimide cross-linker BMH and *p*-PDM at 4 °C or 25 °C (*eighth through fifteenth vertical lanes*). Cross-linking reactions were stopped by adding 10 mM dithiothreitol (maleimide cross-linker) or 10 mM EDTA (CuPh). After solubilization with 1% β -dodecyl-maltoside protein was digested with thrombin, and the C-terminal FLAG epitope was detected after SDS-PAGE separation and Western blotting with mouse anti-FLAG IgG as described before (Quick *et al.*, 1996). Intact PutP migrated at a size of 44 kDa (a). Thrombin digestion of PutP in the absence of cross-linker resulted in C-terminal fragments of 30 kDa (b). Thrombin digestion of PutP successfully cross-linked between TMs II and IX resulted in bands of the same size as intact PutP (c). *dC* indicates that the double-Cys variant of PutP given on top of each gel was used. The respective single-Cys PutP variants (*sC II*, Cys in only in TM II; *sC IX*, Cys only in TM IX) served as controls. CuPh-cross-linked PutP was treated with β -mercaptoethanol to check reversibility of disulfide formation (d).

All three cross-linkers (CuPh, BMH, *p*-PDM) caused intramolecular cross-linking between Cys in TMs II and IX, although the efficiency of cross-linking varied with the cross-linker and position analyzed (Fig. 2.5.). Cys residues at positions 57 and 341 were most efficiently cross-linked (>90% intramolecular cross-linking) not only with the zero-length cross-linker CuPh but also with BMH and *p*-PDM, indicating a temporal close proximity of both positions. The fact that the two bifunctional maleimides were also well accommodated hints at a structural flexibility of this part of the protein. Positions 57 and 340 were efficiently cross-linked (>90%) by BMH and *p*-PDM and to a significant lesser extent by CuPh (15-50% depending on temperature), suggesting a predominant distance between these positions in the range of 9.2 Å and 12.3 Å. Also, the Cys pairs S41C/C344, D55C/S340C, and

D55C/T341C were cross-linked by all three cross-linkers albeit the cross-linking efficiencies were lower (20-40% intramolecular cross-linking) than that of S57C/T341C. Finally, M62C/A327C showed significant cross-linking by CuPh (30% intramolecular cross-linking) only at 25 °C, whereas reaction at 4 °C was ineffective, proposing that a conformational alteration preceded disulfide bridge formation. In contrast to CuPh, M62C/A327C could be cross-linked with BMH or *p*-PDM (40 to 45% intramolecular cross-linking) already at 4 °C, suggesting that under this condition the distance between both positions lied in the range of 9.2 Å and 12.3 Å (distance spanned by rigid *p*-PDM).

To test for possible ligand-induced changes of distances between TMs II and IX, the effect of Na⁺ and proline on Cys cross-linking was analyzed (suppl. Fig. 2.4.). The results did not reveal any significant effect of the ligands on cross-linking.

Taken together, these results demonstrated an at least temporal proximity of TMs II and IX. The distances between selected residues of these domains were not well defined probably due to the structural flexibility of this protein region.

2.5. Discussion

In a search for residues involved in ion and/or substrate binding in the SSS family member PutP, the role of two polar amino acids, Ser340 and Thr341, in Na⁺/proline transport is analyzed. Both amino acids are located in the middle of TM IX and are conserved in pro- and eukaryotic members of the family independent of their substrate specificity (e.g. proline, pantothenate, glucose, iodide). This investigation demonstrates that the two residues are of particular functional significance. This is documented by the fact that already relative conservative replacements (S340A, S340C or S340T; T341C, T341S, or T341V) cause a significant inhibition of proline uptake into intact cells as well as proteoliposomes containing the reconstituted PutP variants. In addition, simultaneous substitution of both residues (group III mutants) inhibits transport almost completely. Kinetic analyses reveal that inhibition of transport of group I (S340A, S340T, T341S) and group II mutants (S340C, T341C, T341V) are primarily due to drastic changes of the apparent affinities of PutP for its ligands. By contrast, maximum rates of proline uptake are only marginally affected. The results suggest that events associated with ligand binding rather than with subsequent translocation steps are hampered by the alterations.

Most severe defects are observed for binding of the coupling ion. Substitution of Ser340 decreases the apparent Na⁺ and Li⁺ affinities of PutP by 2-3 orders of magnitude (Fig. 2.3.). These dramatic changes ascribe Ser340 a specific role in Na⁺ (and Li⁺) binding. However, the size of the side chain at this position had a stronger impact on activity than the presence of a hydroxyl group. Therefore, it is proposed that the residue is located in the direct

surroundings of the ion-binding site of PutP. Alternatively, it may participate in ion binding via its main-chain carbonyl group.

Substitution of the neighboring Thr341 results in a somewhat different phenotype, indicating a specific role for the hydroxyl group of the side chain in ion binding. Although the most conservative replacement with Ser has only little effect on ion binding, Na⁺ does not significantly stimulate proline uptake into intact cells containing PutP-T341C or -T341V. Using the more defined proteoliposomes system, Na⁺ stimulation of transport containing either one of these mutants is observed only at elevated ion concentrations, suggesting a highly reduced (minimum two orders of magnitude) ion affinity. Why is this stimulation not seen with intact cells? Clearly, it is easier to establish defined energetic conditions with proteoliposomes than with intact cells in which not only PutP but a number of other membrane proteins (e.g. Na⁺/H⁺ antiporter) interfere with electrochemical ion gradients. Furthermore, high concentrations of Na⁺ have an inhibitory effect on various proteins and cellular processes (Padan and Krulwich, 2000). In fact, proline uptake into PutP wild-type cells is inhibited already at Na⁺ concentrations of 250 mM and higher. This is exactly the concentration range by which proline uptake into PutP-T341C or -T341V proteoliposomes becomes stimulated. Therefore, because of the complex scenario in intact cells, stimulatory and inhibitory effects of Na⁺ on proline transport may partially compensate each other. In any event, the observed distinctive effects of the Thr341 substitutions on the apparent ion affinity implicate the residue in binding of the coupling ion.

Besides the effects of the amino acid replacements on ion binding, proline binding and transport is also affected. The substitutions inhibit also counterflow activity, which does not require an electrochemical ion gradient. In addition, the proline affinity is affected, albeit to a lesser extent than the ion affinities (Fig. 2.3.). These results provide additional evidence for the particular functional significance of Ser340 and Thr341 and indicate that both residues are important for ion and proline binding. Similar observations have already been made upon replacement of other residues (Asp55, Ser57, Gly58 of TM II) supposed to be involved in ligand binding (Pirch *et al.*, 2002; Quick *et al.*, 1996; Quick and Jung, 1997). Obviously, there is an interdependence of ion and proline binding as shown by this and other investigations (Jung *et al.*, 1998a; Zhou *et al.*, 2004; Yamato and Anraku, 1993; Yamato and Anraku, 1990). A simple explanation for the observed phenomena is a close proximity or even an overlapping of ion- and proline-binding sites. Such an arrangement would ensure tight coupling of ion and substrate transport. But what would be the primary function of Ser340 and Thr341? The following results argue for a primary role of the residues in ion binding: (i) the amino acid replacements have a stronger impact on ion than on proline binding; ii) apparent ion affinities remain reduced even in presence of a saturating proline concentration; (iii) Cys at position 340 or 341 is not protected by proline as, e.g. observed for positions 57

and 58, which are suggested to participate in proline binding (Pirch *et al.*, 2002; Quick *et al.*, 1996).

A location of Ser340 and Thr341 in or close to a ligand-translocation pathway is further supported by the result that Cys placed at either position in the middle of TM IX reacts with methanethiosulfonate compounds of different polarities. This result indicates an accessibility of both positions from the water phase. The lack of efficient Cys protection by Na⁺ may be attributed to the highly reduced Na⁺ affinities of the mutants. In addition, coupling ions are small and, therefore, *a priori* less efficient in protecting Cys at or close to a ligand-binding site than a larger co-transported substrate (e.g. amino acid, sugar).

If Ser340 and Thr341 participate in ligand binding, are the residues located close to other amino acids implicated in Na⁺ (Asp55) and proline binding (Ser57, Gly58)? Here, it is shown that pairs of Cys residues placed at these positions 55/340, 55/341, 57/340, and 57/341 form disulfide bridges. Albeit the cross-linking efficiency varies between the pairs, the results support the idea of a functional interaction of TMs II and IX. However, the necessary amino acid replacements in (or close to) the proposed ligand-binding sites of PutP impair transport activity almost completely. Therefore, pairs of Cys residues are also introduced at either end of TMs II and IX. The resulting functional PutP variants are also cross-linked, indicating an at least temporal proximity of TMs II and IX. Furthermore, the observation that cross-linkers of varying lengths cross-link positions in both TMs suggests a conformational flexibility of these parts of the transporter. The latter idea is further supported by the temperature dependence of cross-linking of M62C/A327C. The proposed proximity between TMs II and IX is in general agreement with earlier analyses reporting a clustering of mutations causing an altered substrate specificity within the N- and C-terminal domains of PutP (Dila and Maloy, 1986).

Despite comprehensive trials, a significant effect of Na⁺ and proline binding on cross-linking efficiencies (e.g. via altering intramolecular distances) was not observed, although a participation of TM II in ligand-induced conformational alterations was previously reported (Pirch *et al.*, 2003). Also, in view of the in-part dramatic effects of the Cys substitutions on apparent ligand affinities, it is assumed that the PutP variants are even in the presence of ligands still dynamic enough to allow efficient cross-linking of Cys pairs within the given labeling period. In addition, Cys cross-linking is known to underestimate distances (Abramson *et al.*, 2003), thereby potentially obscuring ligand-induced changes.

Taken together, the data lead to a working hypothesis according to which amino acids of TM II (Asp55) and TM IX (Ser340, Thr341) form part of the ion- and/or substrate-translocation pathway of PutP, thereby Asp55 and Thr341 may directly participate in Na⁺ binding (Fig. 2.6.). Furthermore, since binding of Na⁺ and proline is interdependent and side-chain alterations affect not only binding of Na⁺ but also of proline, it is proposed that binding

of both ligands occurs in close proximity. Thereby, the carboxylate of proline may even directly interact with the coupling ion.

Such an arrangement resembles in part the situation found in the three-dimensional structure of the leucine transporter LeuT_{Aa} (NSS family) (Yamashita *et al.*, 2005). Here, binding of two Na⁺ ions occurs in cavities located also in about the middle of the membrane with Na₁⁺ being in direct contact with the substrate leucine. Furthermore, ligand binding also involves residues in N-terminal and C-terminal domains (e.g. Na₁⁺, TMs I and 6; Na₂⁺, TMs I and VIII). Na₂⁺, for example, makes direct contact to five oxygen atoms out of which three originate from main-chain carbonyl groups and two come from side-chain hydroxyl groups of two adjacent polar amino acids, Thr354 and Ser355 (TM VIII). Although the neighborhood of Thr and Ser resembles the situation in PutP, the results presented here suggest a direct involvement in Na⁺ binding only for the side chain of Thr341 of PutP.

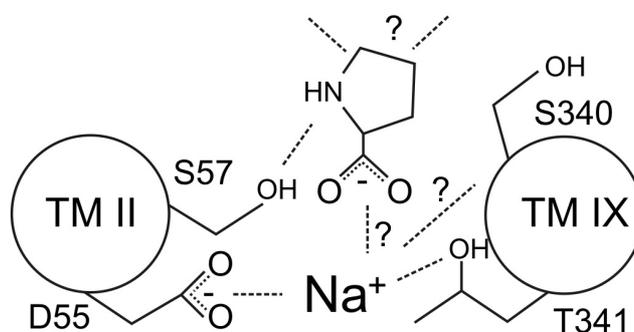


Figure 2.6.: Model showing the participation amino acids of TM II and IX in ligand binding and transport. The model is based on the current analysis as well as on previous investigations (Pirch *et al.*, 2002; Quick *et al.*, 1996; Quick and Jung, 1997). It represents a working hypothesis according to which TMs II and IX participate in the formation of a ligand-translocation pathway; Asp55, Ser57, and Thr341 directly participate in Na⁺ or proline binding. Further structural information is necessary to prove or disprove the model.

Can the model presented in Fig. 2.6. be generalized for all members of the SSS family? The conclusions drawn here are in good agreement with previous analyses of NIS, which like PutP is a member of the SSS family. In fact, the corresponding amino acids in NIS (Ser353 and Thr354) are demonstrated to be particularly crucial for Na⁺ binding and/or translocation (De la Vieja *et al.*, 2007). The general importance of Asp55 and Ser57 is less clear as these residues are not as conserved within the SSS family as Ser340 and Thr341. So, it is possible that Na⁺ binding to the N-terminal domains differs significantly between members of the SSS family. These differences may be attributed to the different Na⁺:substrate stoichiometries (e.g. 1:1 in PutP; 2:1 in NIS, SGLT1). In addition, because ion- and substrate-binding sites are supposed to closely interact with each other, the properties of the substrates (e.g. charged proline *versus* polar sugar) may also have influenced the precise mechanism of Na⁺ binding. In human SGLT1, the N-terminal half of the protein is thought to contain two Na⁺-binding sites, whereas the C-terminal part is made responsible for sugar binding and translocation (Meinild *et al.*, 2001; Panayotova-Heiermann *et al.*, 1996; Panayotova-Heiermann *et al.*, 1997; Kumar *et al.*, 2007). Thereby, sugar binding is supposed

to occur in a hydrophilic cavity formed by TMs X-XIII (Diez-Sampedro *et al.*, 2001; Wright *et al.*, 2007). Residues directly participating in Na⁺ binding are not yet identified.

Finally, it must be considered that the complete sets of residues involved in ion and/or substrate binding are neither known for PutP, SGLT1, or other members of the SSS family. Furthermore, although the results presented here support an involvement of Ser340 and Thr341 of PutP in ligand binding, it must be stated that the observed binding defects can also be explained by indirect effects, e.g. distortion of the structural arrangement of neighboring residues participating in the formation of a binding cavity. Along this line, Ser340 and Thr341 could simply be required for stabilizing intramolecular contacts between N- and C-terminal domains, which in turn may be necessary for high affinity ligand binding. Clearly, more information on transporter structure is required to fully understand the phenomena described here.

In summary, the results demonstrate that Ser340 and Thr341 of PutP are particularly important for high affinity Na⁺ and proline binding. Being placed in the middle of TM IX, the residues are accessible from the water phase and located close to residues of TM II, which are also implicated in ligand binding. For these reasons it is suggested that Ser340 and Thr341 are located in a ligand-translocation pathway. Furthermore, it is proposed that Thr341 directly participates in Na⁺ binding.

2.6. Acknowledgments

We thank Monika Nietschke (University of Osnabrück) for excellent technical assistance. Sophie Hansen (University of Osnabrück) participated in the analysis in a practical course.

2.7. References

- Abramson, J., Smirnova, I., Kasho, V., Verner, G., Iwata, S., and Kaback, H. R.** (2003) The lactose permease of *Escherichia coli*: overall structure, the sugar-binding site and the alternating access model for transport. *FEBS Lett.* 555: 96-101.
- Amann, E., Ochs, B., and Abel, K. J.** (1988) Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene (Amst.)* 69: 301-315.
- Bradford, M. M.** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248-254.
- Chen, C. C., Tsuchiya, T., Yamane, Y., Wood, J. M., and Wilson, T. H.** (1985) Na⁺ (Li⁺)-proline cotransport in *Escherichia coli*. *J. Membr. Biol.* 84: 157-164.

- De la Vieja, A., Reed, M. D., Ginter, C. S., and Carrasco, N.** (2007) Amino acid residues in transmembrane segment IX of the Na⁺/I⁻ symporter play a role in its Na⁺ dependence and are critical for transport activity. *J. Biol. Chem.* 282: 25290-25298.
- Diez-Sampedro, A., Wright, E. M., and Hirayama, B. A.** (2001) Residue 457 controls sugar binding and transport in the Na⁺/Glucose cotransporter. *J. Biol. Chem.* 276: 49188-49194.
- Dila, D. K., and Maloy, S. R.** (1986) Proline transport in *Salmonella typhimurium*: *putP* permease mutants with altered substrate specificity. *J. Bacteriol.* 168: 590-594.
- Dohán, O., De la Vieja, A., and Carrasco, N.** (2006) Hydrocortisone and purinergic signaling stimulate sodium/iodide symporter (NIS)-mediated iodide transport in breast cancer cells. *Mol. Endocrinol.* 20: 1121-1137.
- Green, N. S., Reisler, E., and Houk, K. N.** (2001) Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. *Protein Sci.* 10: 1293-1304.
- Jeschke, G., Wegener, C., Nietschke, M., Jung, H., and Steinhoff, H.-J.** (2004) Interresidual distance determination by four-pulse double electron-electron resonance in an integral membrane protein: The Na⁺/proline transporter PutP of *Escherichia coli*. *Biophys. J.* 86: 2551-2557.
- Jung, H.** (1998) Topology and function of the Na⁺/proline transporter of *Escherichia coli*, a member of the Na⁺/solute cotransporter family. *Biochim. Biophys. Acta* 1365: 60-64.
- Jung, H., Tebbe, S., Schmid, R., and Jung, K.** (1998a) Unidirectional reconstitution and characterization of purified Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 37: 11083-11088.
- Jung, H., Rübenhagen, R., Tebbe, S., Leifker, K., Tholema, N., Quick, M., and Schmid, R.** (1998b) Topology of the Na⁺/proline transporter of *Escherichia coli*. *J. Biol. Chem.* 273: 26400-26407.
- Jung, H.** (2002) The sodium/substrate symporter family: structural and functional features. *FEBS Lett.* 529: 73-77.
- Jung, H., Pirch, T., and Hilger, D.** (2006) Secondary transport of amino acids in prokaryotes. *J. Membr. Biol.* 213: 119-133.
- Kaback, H. R.** (1971) Bacterial membranes. *Methods Enzymol.* 22: 99-120.
- Kumar, A., Tyagi, N. K., Goyal, P., Pandey, D., Siess, W., and Kinne, R. K.** (2007) Sodium-independent low-affinity D-glucose transport by human sodium/D-glucose cotransporter 1: critical role of tryptophan 561. *Biochemistry* 46: 2758-2766
- Meinild, A. K., Loo, D. D., Hirayama, B. A., Gallardo, E., and Wright, E. M.** (2001) Evidence for the involvement of Ala 166 in coupling Na⁺ to sugar transport through the human Na⁺/glucose cotransporter. *Biochemistry* 40: 11897-11904.

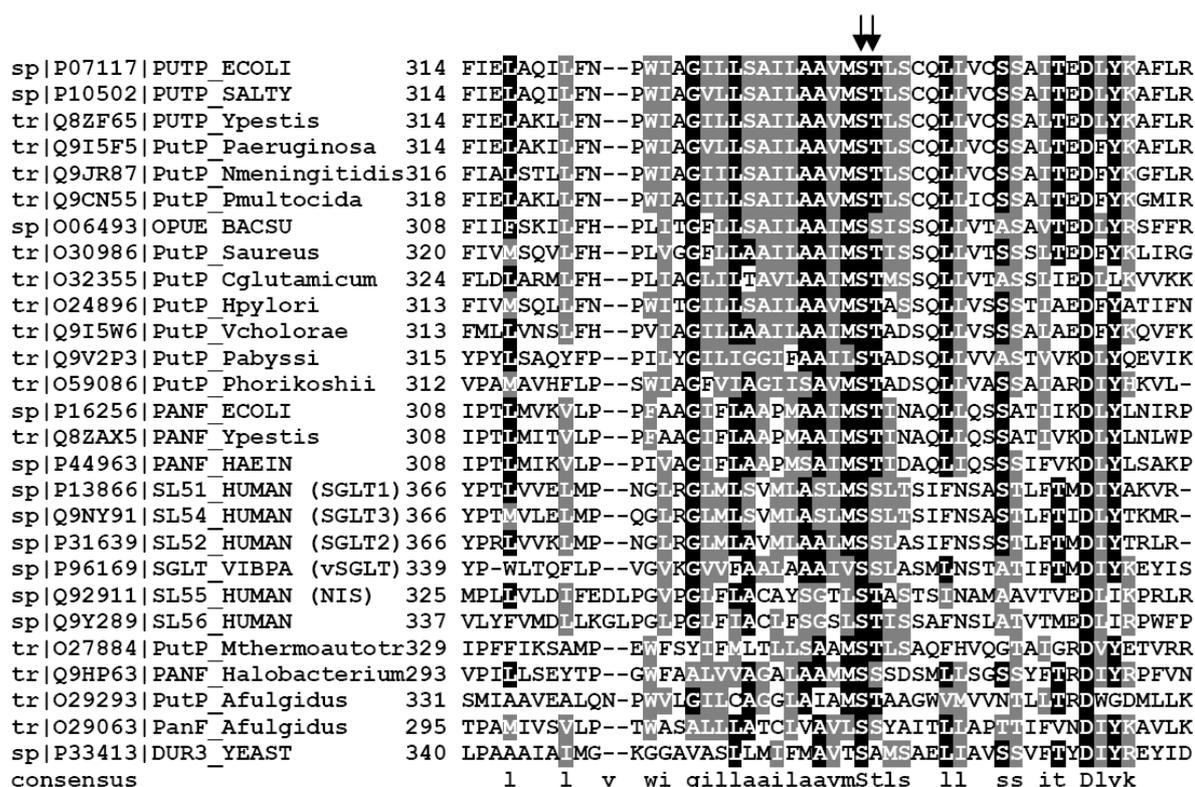
- Miller, J. H.** (1992) *A short course in bacterial genetics. A laboratory manual and handbook for Escherichia coli and related bacteria*, p. 439, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Padan, E., and Krulwich, T. A.** (2000) Sodium Stress. In *Bacterial stress responses* (Storz, G., and Hengge-Aronis, R., eds.) pp. 117-130, American Society for Microbiology, Washington, DC.
- Panayotova-Heiermann, M., Loo, D. D., Kong, C. T., Lever, J. E., and Wright, E. M.** (1996) Sugar binding to Na⁺/Glucose cotransporters is determined by the carboxyl-terminal half of the protein. *J. Biol. Chem.* 271: 10029-10034.
- Panayotova-Heiermann, M., Eskandari, S., Turk, E., Zampighi, G. A., and Wright, E. M.** (1997) Five transmembrane helices form the sugar pathway through the Na⁺/Glucose cotransporter. *J. Biol. Chem.* 272: 20324-20327.
- Peterson, G. L.** (1977) A simplification of the protein assay method of Lowry *et al.* which is more generally applicable. *Anal. Biochem.* 83: 346-356.
- Pirch, T., Quick, M., Nietschke, M., Langkamp, M., and Jung, H.** (2002) Sites important for Na⁺ and substrate binding in the Na⁺/proline transporter of *Escherichia coli*, a member of the Na⁺/solute symporter family. *J. Biol. Chem.* 277: 8790-8796.
- Pirch, T., Landmeier, S., and Jung, H.** (2003) Transmembrane domain II of the Na⁺/proline transporter PutP of *Escherichia coli* forms part of a conformationally flexible, cytoplasmic exposed aqueous cavity within the membrane. *J. Biol. Chem.* 278: 42942-42949.
- Quick, M., Tebbe, S., and Jung, H.** (1996) Ser57 in the Na⁺/proline permease of *Escherichia coli* is critical for high-affinity proline uptake. *Eur. J. Biochem.* 239: 732-736.
- Quick, M., and Jung, H.** (1997) Aspartate 55 in the Na⁺/proline permease of *Escherichia coli* is essential for Na⁺-coupled proline uptake. *Biochemistry* 36: 4631-4636.
- Racher, K. I., Voegelé, R. T., Marshall, E. V., Culham, D. E., Wood, J. M., Jung, H., Bacon, M., Cairns, M. T., Ferguson, S. M., Liang, W. J., Henderson, P. J., White, G., and Hallett, F. R.** (1999) Purification and reconstitution of an osmosensor: transporter ProP of *Escherichia coli* senses and responds to osmotic shifts. *Biochemistry* 38: 1676-1684.
- Reizer, J., Reizer, A., and Saier, M. H., Jr.** (1994) A functional superfamily of sodium/solute symporters. *Biochim. Biophys. Acta* 1197: 133-166.
- Schaffner, W., and Weissmann, C.** (1973) A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Anal. Biochem.* 56: 502-514.
- Schwan, W. R., Lehmann, L., and McCormick, J.** (2006) Transcriptional activation of the *Staphylococcus aureus putP* gene by low-proline-high osmotic conditions and during infection of murine and human tissues. *Infect. Immun.* 74: 399-409.
- Stalmach, M. E., Grothe, S., and Wood, J. M.** (1983) Two proline porters in *Escherichia coli* K-12. *J. Bacteriol.* 156: 481-486.

- Tabor, S., and Richardson, C. C.** (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U. S. A.* 82: 1074-1078.
- Turk, E., and Wright, E. M.** (1997) Membrane topology motifs in the SGLT cotransporter family. *J. Membr. Biol.* 159: 1-20.
- Wegener, C., Tebbe, S., Steinhoff, H.-J., and Jung, H.** (2000) Spin labeling analysis of structure and dynamics of the Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 39: 4831-4837.
- Wright, E. M., and Turk, E.** (2004) The sodium/glucose cotransport family SLC5. *Pfluegers Arch. Eur. J. Physiol.* 447: 510-518.
- Wright, E. M., Hirayama, B. A., and Loo, D. F.** (2007) Active sugar transport in health and disease. *J. Intern. Med.* 261: 32-43.
- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E.** (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* 437: 215-223.
- Yamato, I., and Anraku, Y.** (1990) Mechanism of Na⁺/proline symport in *Escherichia coli*: reappraisal of the effect of cation binding to the Na⁺/proline symport carrier. *J. Membr. Biol.* 114: 143-151.
- Yamato, I., and Anraku, Y.** (1993) Na⁺/substrate symport in prokaryotes. In *Alkali cation transport systems in prokaryotes* (Bakker, E. P., ed.) pp. 53-76, CRC Press, Inc., Boca Raton, FL.
- Yanisch-Perron, C., Vieira, J., and Messing, J.** (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene (Amst.)* 33: 103-119.
- Zhou, A., Wozniak, A., Meyer-Lipp, K., Nietschke, M., Jung, H., and Fendler, K.** (2004) Charge translocation during cosubstrate binding in the Na⁺/proline transporter of *E. coli*. *J. Mol. Biol.* 343: 931-942.

2.8. Supplementary Material

2.8.1. Sequence alignment of members of the SSS family

Searching for amino acid residues of potential structural and/or functional significance (e.g., residues involved in ligand binding) a multiple sequence alignment of members of the SSSF was performed. The analysis revealed that Ser340 and Thr341 of *E. coli* PutP are conserved within prokaryotic and eukaryotic family members irrespective of their substrate specificity (e.g., transporters for proline, pantothenate, sugars, or iodide). In fact, the transporters analyzed contained Ser at the position corresponding to Ser340, and Ser or Thr at the position corresponding to Thr341 of *E. coli* PutP (suppl. Fig. 2.1.).



Supplementary Figure 2.1.: Multiple sequence alignment of the region around TM IX of pro- and eukaryotic members of the SSSF. A multiple alignment was generated using ClustalW (1.82) with sequences (FASTA) obtained from a blastp search (www.ncbi.nlm.nih.gov/BLAST) against PutP of *E. coli* K12 (swiss-prot entry P07117) at the EBI [www-server](http://www.ebi.ac.uk/cluster/) (www.ebi.ac.uk/cluster/) (Pearson and Lipman, 1988; Pearson, 1990). The chosen matrix was blosum62, penalties were 10 (open gap), 0.05 (extending gap) and 10 (gap distance). The resulting alignment was optically improved by the Boxshade 3.21 program (www.ch.embnet.org), highlighting conserved residues in black (>60% identical to consensus) or grey (>60% similar to consensus). The term PutP refers to the homology with PutP of *E. coli*, experimental evidence for Na⁺/proline transporter activity exists only for *Salmonella typhimurium* (SALTY), *Staphylococcus aureus* (Saureus), and OpuE of *Bacillus subtilis* (BACSU). Arabic numbers indicate the position in the primary sequence and refer to the first amino acid of the alignment.

2.8.2. Establishment of protease cleaving sites in PutP

Factor Xa or thrombin cleavage sites were introduced into periplasmic (L7) or cytoplasmic loops (L8) of PutP(Δ Cys) to facilitate determination of cross-linking between the TMs II and IX (suppl. Table 2.1.). Introduction of the cleavage sites was performed by oligonucleotide-directed site-specific mutagenesis. After PCR amplification, mutagenic

fragments were treated with *NheI* and *PstI* and ligated with similarly restricted plasmid pT7-5/*putP*(Δ Cys) encoding PutP with the given single or double-Cys substitutions.

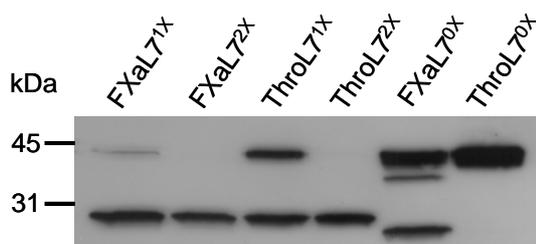
Supplementary Table 2.1.: Localization and organization of protease cleaving-sites introduced in PutP(Δ Cys).

Localization and number of protease cleaving-site	Primary sequence ^a
Single thrombin cleaving-site in loop 7 (ThroL7 ^{1X})	...DSLEVIKQK <u>L</u> VPRG <u>S</u> IENVDMLK...
Double thrombin cleaving-site in loop 7 (ThroL7 ^{2X})	...DSLEVIKQK <u>L</u> VPRG <u>S</u> L <u>V</u> PRG <u>S</u> IENVDMLK...
Single factor Xa cleaving-site in loop 7 (FXaL7 ^{1X})	...DSLEVIKQKS <u>I</u> EGRNVDMLK...
Double factor Xa cleaving-site in loop 7 (FXaL7 ^{2X})	..DSLEVIKQKS <u>I</u> EGR <u>I</u> EGRNVDMLK...
Single factor Xa cleaving-site in loop 8 (FXaL7 ^{1X})	...RFMAADSHHS <u>I</u> EGRVHARRIS...

^a The single or double protease recognition-sites are underlined. Inserted amino acids (1-letter-code) are displayed with bold letters.

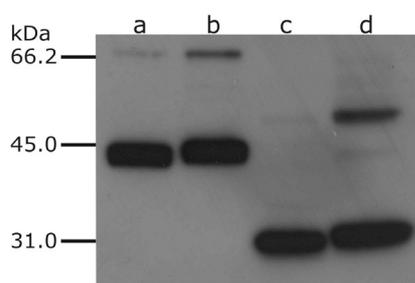
The capability of the resulting PutP variants to catalyze L-[¹⁴C]-proline uptake into *E. coli* WG170 was analyzed as described in the main manuscript. While insertion of protease cleavage sites into L8 inhibited transport completely, transporters with insertions in L7 were all functional [50 to 100% of PutP(Δ Cys) activity depending on the nature and number of the introduced sites]. Accessibility of the cleaving sites was tested after protein solubilization from the membranes as described in the main manuscript. C-terminal fragments resulting from cleavage in loop 7 were detected with anti-FLAG antibody and migrated at an apparent M_r of about 30 kDa, while uncleaved PutP had an apparent M_r of about 44 kDa (Suppl. Fig. 2.2.). Complete cleavage at the desired site without byproducts occurred only when two adjacent thrombin sites were inserted into periplasmic loop 7 (ThroL7^{2X}). Therefore, the latter construct was used in the following experiments.

For *in situ* Cys cross-linking experiments, the following pairs of Cys residues were individually introduced in TMs II and IX of PutP(ThroL7^{2X}): S41C/C344, D55C/T341C, S57C/T341C, M62C/A327C). Na⁺-coupled proline uptake into strain WG170 was demonstrated for PutP(ThroL7^{2X})-S41C/344C and -M62C/A327C [e.g., 80% and 10%, respectively, of the PutP(Δ Cys) steady-state level of accumulation], whereas transport activity of PutP(ThroL7^{2X})-S57C/S340C and -S57C/T341C was only detectable at elevated proline concentrations (2% and 3%, respectively, of the PutP(Δ Cys) steady-state level of accumulation in presence of 250 μ M ¹⁴C-proline). PutP(ThroL7^{2X})-D55C/S340C and -D55C/T341C were inactive. Western blot analysis showed, that the differences described between PutP(Δ Cys) and the double-Cys mutants with inserted thrombin cleavage-sites could not be attributed to defective expression, insertion into the membrane or instability of the PutP variants. The results were in full agreement with the described functional significance of the residues (Pirch *et al.*, 2002; Quick *et al.*, 1996; Quick and Jung, 1997; this article).



Supplementary Figure 2.2.: Efficiency of site-specific protease cleavage of PutP. Membranes prepared from WG170 containing PutP(Δ Cys) with and without given cleavage sites were solubilized with 1% DDM and subjected to digestion by Factor Xa (FXa) or Thrombin (Thro) at 4°C for 16 hours. The protein fragments were resolved in a SDS Gel and the C-terminal FLAG epitope of PutP was detected by Western blotting with HRP-linked monoclonal FLAG antibody. The exponent displays the number of inserted protease cleaving sites in loop 7 (L7). PutP(Δ Cys) without FXa (FXaL7^{0x}) or Thro cleaving site (ThroL7^{0x}) was used as a control.

Single-Cys PutP variants were used as negative controls in the cross-linking experiments. Besides the expected bands for full size and cleaved PutP, an additional band of about 55 kDa was observed after thrombin cleavage especially when the long and flexible cross-linker BMH was applied to membranes containing PutP variants with a single Cys in TM IX (Suppl. Fig. 2.3.). The band was most prominent for single-Cys PutP(ThroL7^{2x})-A327C in presence of BMH. A mass spectrum analysis identified the protein as the C-terminal fragment (30.34 kDa) obtained after thrombin cleavage which presumably dimerized in presence of BMH. Since the electrophoretic mobility of the dimer band differed significantly from the other bands, intra- and intermolecular cross-linking could clearly be distinguished.

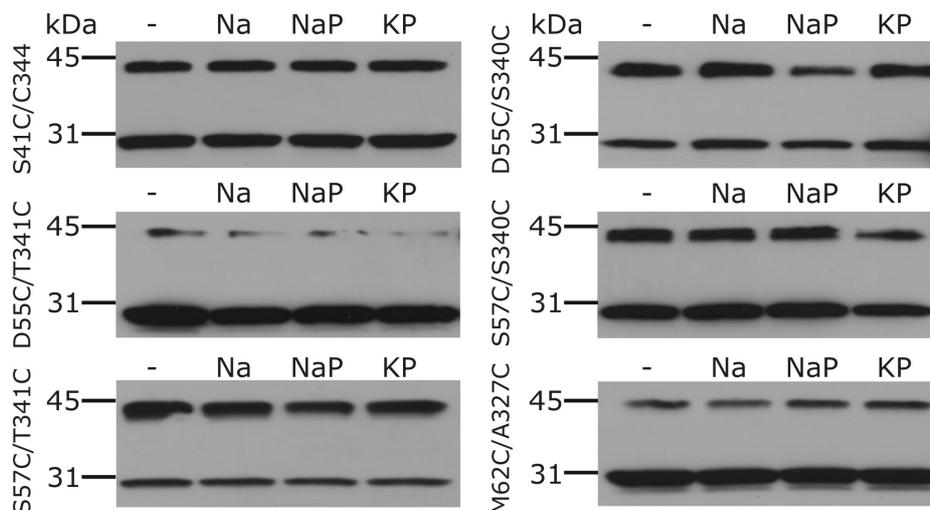


Supplementary Figure 2.3.: *In situ* intermolecular cross-linking PutP(Δ Cys)-A327C. Membranes prepared from WG170 containing PutP(Δ Cys)-A327C with inserted thrombin cleaving sites in loop 7 were subjected to 2 mM BMH at 25°C. Cross-linking reactions were stopped by adding 10 mM DTT. After solubilization with 1% DDM, protein was digested by thrombin or was incubated with protease storage buffer at 4°C for 16 hours. The protein fragments were resolved in a SDS Gel and the C-terminal FLAG epitope of PutP was detected by Western blotting with HRP-linked monoclonal FLAG antibody. In the absence of BMH, undigested protein (lane a, 44 kDa band) and digested protein (lane c, 30 kDa band) were used as controls. Intermolecular crosslinking in the presence of BMH yielded a band of 68 kDa before (lane b) and of 55 kDa after thrombin cleavage (lane d).

2.8.3. Effect of ligand on Cys cross-linking

To analyze the effect of Na⁺ and proline on distances between TMs II and IX, Cys crosslinking experiments were performed as described in the main manuscript with the following changes: Where indicated membrane vesicles were pre-incubated with 50 mM Na⁺, 10 mM L-proline or both followed by addition of cross-linker. Cross-linking was performed at 25 °C for 10 min. Three independent repeats of the experiment did not reveal a significant

impact of the ligands on the cross-linking efficiency. The results of a typical experiment are presented in Suppl. Fig. 2.4..



Supplementary Figure 2.4.: Influence of PutP ligands on cross-linking of TM II and TM IX. One-hundred μ l membranes (1 mg of total membrane protein/ml), containing given PutP derivatives, were subjected to 0.5 mM of CuPh in the presence of 50 mM sodium (Na, lane 2), 50 mM sodium and 10 mM proline (NaP, lane 3), or 50 mM potassium and 10 mM proline (KP, lane 4). Cross-linking in the absence of ligands was performed as a control (lane 1). After an incubation time of 10 minutes, the cross-linking reactions were stopped by adding 10 mM of EDTA. After solubilisation with 1% DDM, protein was digested with thrombin and the C-terminal FLAG epitope was detected after SDS-Page separation and Western blotting with mouse anti-FLAG IgG. Thrombin digestion of PutP successfully cross-linked between TMs II and IX resulted in bands of about 44 kDa. The double-Cys variants of PutP are given on the left hand side of each gel.

2.9. References

- Pearson, W. R. and Lipman, D. J. (1988)** Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* 85: 2444-2448.
- Pearson, W. R. (1990)** Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods. Enzymol.* 183: 63-98.
- Pirch, T., Quick, M., Nietschke, M., Langkamp, M., and Jung, H. (2002)** Sites important for Na^+ and substrate binding in the Na^+ /proline transporter of *Escherichia coli*, a member of the Na^+ /solute symporter family. *J. Biol. Chem.* 277: 8790-8796.
- Quick, M., Tebbe, S., and Jung, H. (1996)** Ser57 in the Na^+ /proline permease of *Escherichia coli* is critical for high-affinity proline uptake. *Eur. J. Biochem.* 239: 732-736.
- Quick, M., and Jung, H. (1997)** Aspartate 55 in the Na^+ /proline permease of *Escherichia coli* is essential for Na^+ -coupled proline uptake. *Biochemistry* 36: 4631-4636.

CHAPTER 3**Function of transmembrane domain IX in the Na⁺/proline transporter PutP****3.1. Abstract**

Selected residues of transmembrane domain (TM) IX were previously shown to play key roles in ligand binding and transport in members of the Na⁺/solute symporter family. Using the Na⁺/proline transporter PutP as a model, a complete Cys scanning mutagenesis of TM IX (positions 324 to 351) was performed here to further investigate the functional significance of the domain. Gly328, Ser332, Gln345, and Leu346 were newly identified as important for Na⁺-coupled proline uptake. Placement of Cys at one of these positions altered $K_{m(\text{pro})}$ (S332C and L346C, 3- and 21-fold decreased, respectively; Q345C, 38-fold increased), $K_{0.5(\text{Na}^+)}$ (S332C, 13-fold decreased; Q345C, 19-fold increased), and/or V_{max} [G328C, S332C, Q345C, and L346C, 3-, 22-, 2-, and 8-fold decreased compared to PutP(wild type), respectively]. Membrane-permeant *N*-ethylmaleimide inhibited proline uptake into cells containing PutP with Cys at distinct positions in the middle (T341C) and cytoplasmic half of TM IX (C344, L347C, V348C, and S351C) and had little or no effect on all other single-Cys PutP variants. The inhibition pattern was in agreement with the pattern of labeling with fluorescein-5-maleimide. In addition, Cys placed into the cytoplasmic half of TM IX (C344, L347C, V348C, and S351C) was protected from fluorescein-5-maleimide labeling by proline while Na⁺ alone had no effect. Membrane-impermeant methanethiosulfonate ethyltrimethylammonium modified Cys in the middle (A337C and T341C) and periplasmic half (L331C) but not in the cytoplasmic half of TM IX in intact cells. Furthermore, Cys at the latter positions was partially protected by Na⁺ but not by proline. Based on these results, a model is discussed according to which residues of TM IX participate in the formation of ligand-sensitive, hydrophilic cavities in the protein that may reconstitute part of the Na⁺-and/or proline-translocation pathway of PutP.

3.2. Introduction

The Na⁺/solute symporter family (SSSF) (TC 2.A.21, SLC5) is an evolutionary related collection of secondary transporters found in all three kingdoms of life (Jung, 2002; Wright and Turk, 2004; Reizer *et al.*, 1994). Prokaryotic members of the family utilize a sodium motive force (*smf*) to specifically accumulate, for example, proline (PutP, OpuE), galactose and glucose (SGLT), pantothenate (PanF), or phenylacetate (Ppa) in cells (Jung, 2001). Most of these transporters (PutP, SGLT, and Ppa) feed catabolic pathways or play a role in cell adaptation to osmotic stress conditions (OpuE). PanF scavenges extracellular

pantothenate for coenzyme A biosynthesis in the cells (Jackowski and Alix, 1990). In eukaryotes, SSSF members are responsible for the transport of, for example, monosaccharides (SGLT), myoinositol (SMIT), vitamins (SMVT), iodide (NIS), and urea (DUR) (Dohán *et al.*, 2006; Wright *et al.*, 2007; Kojima *et al.*, 2007). So far, missing information on the three-dimensional structure of SSSF members limits access to the molecular mechanism of Na⁺-coupled substrate transport.

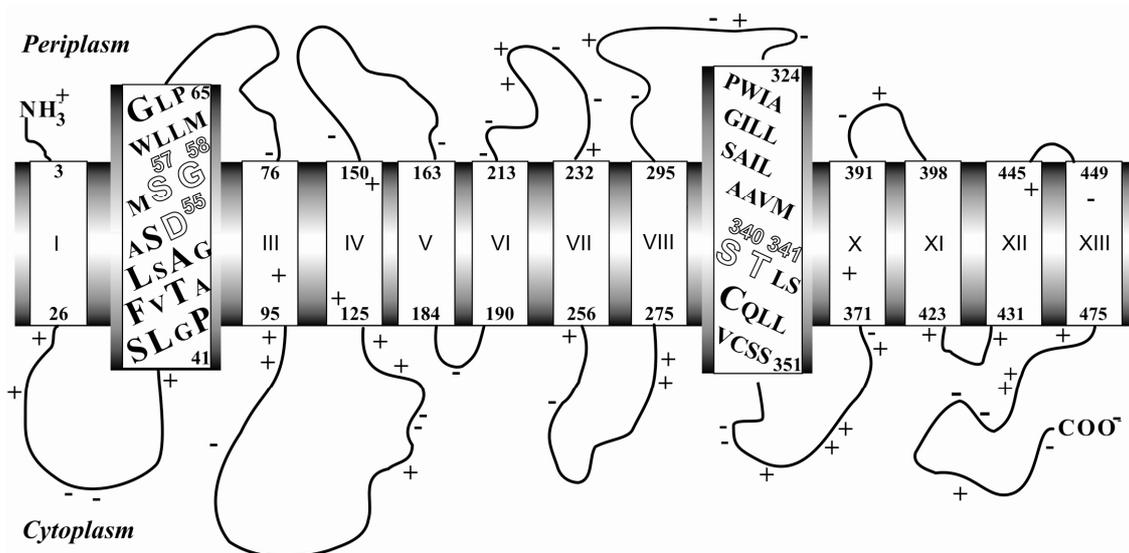


Figure 3.1.: Secondary structure model of PutP of *E. coli* highlighting functional important residues. The model is based on a gene fusion approach, Cys accessibility analyses, site-directed spin labeling, and site-specific proteolysis (Jung *et al.*, 1998a; Wegener *et al.*, 2000). Putative TMs are represented as rectangles and numbered with Roman numerals. Amino acid residues of TMs II and IX previously shown to be crucial for Na⁺ and/or proline binding are indicated by open letters (Quick *et al.*, 1996; Quick and Jung, 1997; Pirch *et al.*, 2002; Hilger *et al.*, 2008). Residues of TM II proposed to line a ligand-sensitive hydrophilic cavity in the transporter are highlighted (enlarged letters) (Pirch *et al.*, 2003). In TM IX, only Cys344 (enlarged letter) is known so far to react with sulfhydryl reagents and is protected by substrate (Yamato and Anraku, 1993). TMs II and IX are adjacent in the tertiary structure of the protein as shown by Cys cross-linking (Hilger *et al.*, 2008).

We use PutP of *Escherichia coli* as a model to obtain insights into structure and molecular mechanisms of function of SSSF members. Genetic, protein chemical, and spectroscopic analyses suggest a secondary-structure model according to which PutP contains 13 transmembrane domains (TMs) with the N terminus located on the periplasmic side of the membrane and the C terminus facing the cytoplasm (Fig. 3.1.) (Jung *et al.*, 1998a; Wegener *et al.*, 2000). Furthermore, Cys accessibility experiments propose a participation of TM II in the formation of a hydrophilic cavity in the transporter and implicate the TM in ligand-induced conformational alterations (Pirch *et al.*, 2003). In addition, Asp55 in TM II is essential for function and proposed to participate in binding of the coupling ion (Quick and Jung, 1997). Ser57 and Gly58 of the same TM are important for Na⁺ and proline binding, suggesting partly overlapping binding sites for both ligands (Quick *et al.*, 1996; Pirch *et al.*, 2002). More recent Cys cross-linking data propose that TM II is close to TM IX in the tertiary structure of the protein. In fact, also, the latter TM contains residues (Ser340 and Thr341) that are required for high-affinity Na⁺ and proline binding (Hilger *et al.*, 2008). Beside PutP, polar residues of TM IX were shown to be particularly crucial also for function of human NIS

(De la Vieja *et al.*, 2007), suggesting that this domain is of general significance for members of the SSSF.

To gain further insights into the role of TM IX, all amino acids of the domain of our model system PutP (Pro324 to Ser351) were individually replaced with Cys and the functional consequences were analyzed. Furthermore, accessibilities of Cys placed throughout TM IX to sulfhydryl reagents of various sizes and polarities were determined. Finally, the influence of ligands (Na⁺, proline) on the labeling reactions was investigated. The results led to a model according to which TM IX contributes to the formation of hydrophilic cavities in PutP, which may reconstitute part of a ligand-translocation pathway of PutP.

3.3. Materials and Methods

3.3.1. Bacterial strains and plasmids

E. coli JM109 [*endA1 recA1 gyrA96 thi hsdR17 supE44 relA1 Δ(lac-proAB)* (*F'**traΔ36pro AB⁺ lacI^q ZΔM15*)] (Yanisch-Perron *et al.*, 1985) and *E. coli* DH5α [*F⁻ φ80d lacZ ΔM15 Δ(lacZYA-argF) U169 deoR recA1 endA1 hsd R17(r_k⁻,m_k⁺) phoA supE44 λ⁻ thi-1 gyrA96 relA1*] were used as carriers for the plasmids generated in this study. *E. coli* WG170 [*F⁻ trp lacZ rpsL thi Δ(putPA)101 proP219*] (Stalmach *et al.*, 1983), harboring given plasmids, was used for overexpression of the *putP* gene and transport assays. The following plasmids, derivatives of pT7-5 (Tabor and Richardson, 1985), containing the *lac* promoter/operator for expression of the *putP* gene were used for all gene manipulations: pT7-5/*putP* and pT7-5/*putP*(ΔCys), each of which harbors a cassette version of the *putP* gene encoding PutP-wild-type and an engineered transporter devoid of all five native Cys residues [PutP(ΔCys)], respectively, and a C-terminal-attached amino acid sequence resembling the FLAG epitope and a 6His tag (Jung *et al.*, 1998a; Wegener *et al.*, 2000).

3.3.2. Site-directed mutagenesis

Every amino acid of TM IX was individually replaced by Cys by site-directed mutagenesis using plasmid pT7-5/*putP*(ΔCys) as a template and synthetic mutagenic oligonucleotides in a one- or two-step PCR reaction with *Taq*-DNA polymerase. PCR fragments were digested with *Pst*I and *Spe*I and ligated with similarly treated plasmid pT7-5/*putP*(ΔCys) additionally incubated with alkaline phosphatase. Constructs were verified by sequencing the resulting plasmid DNA. For overexpression, the mutated *putP* gene was cut out with restriction endonucleases *Nco*I and *Hind*III and ligated to the similarly treated vector pTrc99a additionally incubated with alkaline phosphatase.

3.3.3. Purification and reconstitution of PutP

PutP was purified by Ni²⁺-NTA affinity chromatography and reconstituted into proteoliposomes prepared from *E. coli* polar lipid extracts (Avanti Polar Lipids, Inc., Alabaster, AL) at a lipid-to-protein ratio of 100:1 (w/w) as described before (Jung *et al.*, 1998b).

3.3.4. Proline transport assays

Transport was measured in *E. coli* WG170 transformed with plasmid pT7-5/*putP*(Δ Cys) or pT7-5/*putP* harboring the given mutations. Cells were grown and prepared as previously described (Quick *et al.*, 1996). *Smf*-driven proline uptake was assayed under standard conditions with 10 μ M L-[U-¹⁴C]proline (26 Ci/mol) in the presence of 20 mM D-lactate (Na⁺ salt) and 50 mM NaCl using the rapid filtration method as described previously (Chen *et al.*, 1985). Initial rates of transport were calculated from the initial linear portion of the time course. Steady-state levels of proline accumulation were determined from time points after leveling out of the uptake curve. Kinetic parameters were determined by plotting initial rates and substrate concentrations according to Eadie-Hofstee. The resulting $K_{m(\text{pro})}$, $K_{0.5(\text{Na}^+)}$, and V_{max} values of pT7-5/*putP*-encoded PutP(wild type) closely match the parameters of the native, fully induced, chromosome-encoded PutP in the absence of PutA (Wood and Zadworny, 1979; Chen *et al.*, 1985) suggesting that deviations of kinetic parameters are solely due to the site-specifically introduced amino acid replacements and not to overexpression artifacts.

For transport inhibition analyses, cells containing the respective single-Cys variant were incubated with 2 mM sulfhydryl reagent (NEM or MTSET) at room temperature for 15 min. The reaction was stopped by twofold dilution of the reaction mixture with 100 mM Tris/4-morpholineethanesulfonic acid (Mes) buffer, pH 6.0, containing 0.1% bovine serum albumin (BSA). Cells were washed and resuspended in the same buffer without BSA, and transport was analyzed as described above.

For analysis of the effect of ligand on MTSET labeling, cells containing the respective single-Cys variant were washed four times in sodium-free 100 mM Tris/Mes buffer, pH 6.0, and preincubated in the absence or presence of 30 mM NaCl, 10 mM proline, or 30 mM NaCl/10 mM proline at 25 °C for 10 min. Subsequently, 2 mM MTSET was added and incubation was continued for 10 min. The reaction was stopped by twofold dilution of the reaction mixture with 100 mM Tris/Mes buffer, pH 6.0, containing 0.1% BSA, and washed two times with the same buffer without BSA. The transport was analyzed as described above.

Smf-driven proline uptake into proteoliposomes was analyzed as described before (Jung *et al.*, 1998b).

3.3.5. Cys accessibility to fluorescein-5-maleimide

E. coli WG170 transformed with pTrc99a/*putP*(Δ Cys) were grown, disrupted by sonification, and randomly oriented membrane vesicles were prepared as previously described (Pirch *et al.*, 2003). For accessibility analyses, 150- μ L aliquots of the membrane suspension containing 10 mg/mL total protein were incubated without or with Na⁺, K⁺, and/or L-proline at given concentrations at 25 °C for 10 min. Subsequently, 200 μ M FM was added and incubation was continued at 25 °C for an additional 10 min. Reactions were stopped by addition of 5 mM β -mercaptoethanol. After labeling, PutP was solubilized with 1.5% (w/v) *n*-dodecyl- β -D-maltoside under stirring at 4 °C for 30 min. Then, samples were purified *via* Ni²⁺-NTA spin columns and eluted with 200 mM imidazole as described before (Pirch *et al.*, 2002). After protein determination, equal amounts of protein were subjected to 10% SDS-PAGE. Fluorescent bands of PutP were visualized using the Multimager™ device (BioRad, Munich) and quantified using the software ImageQuant 5.0. After analysis of fluorescent PutP bands, the gel was stained with Coomassie Blue to detect total amounts of protein.

3.3.6. Western Blot analysis

Relative amounts of PutP with given amino acid replacements in membranes of *E. coli* WG170 were estimated by Western Blot analysis with horseradish peroxidase-linked mouse anti-FLAG immunoglobulin G directed against the FLAG epitope at the C terminus of each PutP variant as described before (Quick *et al.*, 1996).

3.3.7. Protein determination

Determination of protein was performed according to Bradford (1976) for purified protein and a modified Lowry method (Peterson, 1977) for total membrane protein with bovine serum albumin as standard.

3.4. Results

3.4.1. Cys scanning mutagenesis of TM IX

TM IX of PutP was subjected to Cys scanning mutagenesis in order to identify residues of particular structural and/or functional significance. For this purpose, each amino acid of TM IX (Pro324 to Ser351) was individually replaced with Cys in a functional PutP molecule

devoid of all five native Cys residues [PutP(Δ Cys)]. A Western blot analysis revealed that these replacements resulted in amounts of transporter molecules in the membrane similar as observed for PutP(Δ Cys), indicating that possible differences in proline uptake kinetics cannot be attributed to different amounts of transporter molecules in the membrane (Fig. 3.2.).



Figure 3.2.: Immunological detection of PutP containing replacements of amino acids in TM IX. Twenty-five micrograms of total membrane protein of each mutant was separated by 10% SDS-PAGE. Proteins were transferred onto a nitrocellulose membrane (0.45 μ m pore size) and probed with mouse anti-FLAG M2 monoclonal antibody linked to horseradish peroxidase. Detection was performed according to the enhanced chemiluminescence method. Cells transformed with plasmid pT7-5 without *putP* served as negative control (NC).

Active transport was then measured under standard test conditions (70 mM Na⁺, 10 μ M proline) by using *E. coli* WG170 complemented with plasmid-encoded PutP variants. Under these conditions, PutP(Δ Cys) exhibited an initial rate of 23 ± 5 nmol/min-mg [50% of PutP(wild type)] and a steady-state level of proline accumulation of 24 ± 3 nmol/mg [100% of PutP(wild type)].

Initial analyses of rates and steady-state levels of proline accumulation newly identified Gly328, Ser332, Gln345, and Leu346 as important for PutP function (Fig. 3.3.). The most severe defect was detected for PutP(Δ Cys)-S332C [2.6% and 15.8% of the initial rate and steady-state level of proline accumulation of PutP(Δ Cys)]. Intermediate inhibitory effects [30-40% of the initial rate and 70-100% of steady-state level of proline accumulation of PutP(Δ Cys)] were observed for cells containing PutP(Δ Cys)-G328C, -Q345C, or -L346C. In contrast, cells with PutP(Δ Cys)-A337C or -C344 exhibited about twofold stimulated initial rates and unaltered steady-state levels of proline accumulation compared to PutP(Δ Cys). It should be noted that placement of Cys at position 344 [Ser in PutP(Δ Cys)] restored a native Cys residue. All remaining substitutions had only little or no impact on transport kinetics (Fig. 3.3.).

3.4.2. Effect of Cys replacements in TM IX of PutP(wild type) on transport kinetics

Substitutions in PutP(Δ Cys) causing alterations of transport were also analyzed in the wild-type background. Cys in place of Gly328, Ser332, Gln345, and Leu346 reduced the initial rate of transport to 30%, 6%, 17%, and 28%, respectively, of the wild-type value and, in the case of the three latter substitutions, also led to reduced maximum levels of proline accumulation (19-56% of the wild-type value). In contrast, replacement of Ala337 increased the initial rate of transport by a factor of 1.8 with no significant effect on the steady-state level

of accumulation (data not further illustrated). These results confirmed the observations made in the PutP(Δ Cys) background.

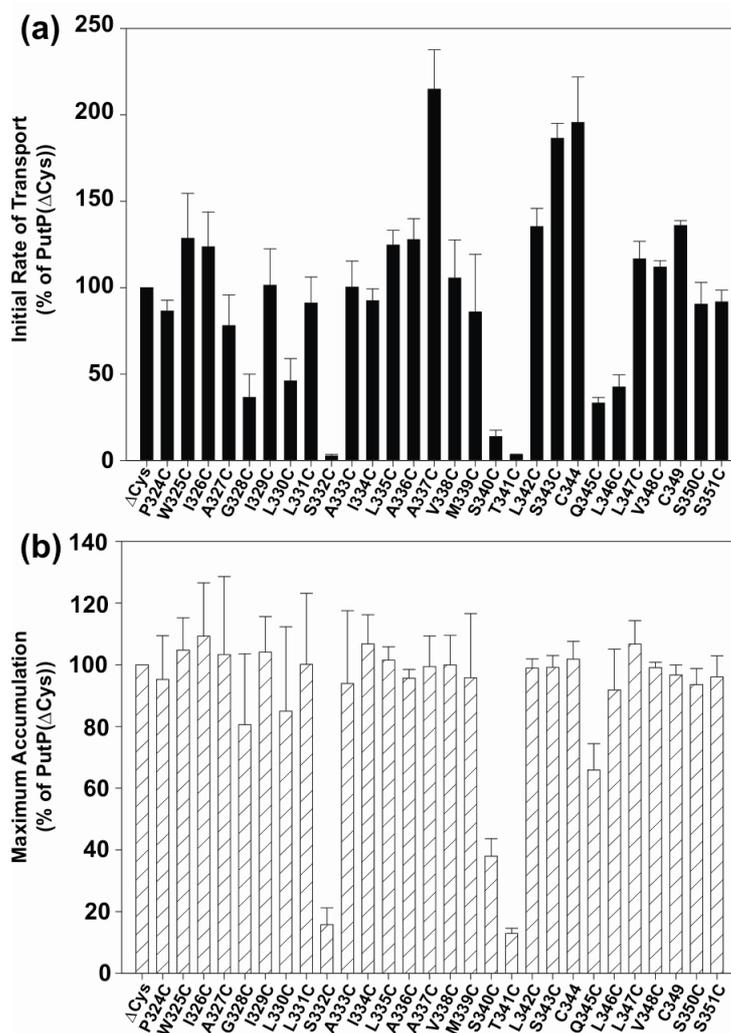


Figure 3.3.: Influence of amino acid replacements in TM IX on Na⁺-coupled proline uptake. Transport of L-[U-¹⁴C]proline (10 μ M final concentration) into cells of *E. coli* WG170 was assayed in the presence of 50 mM NaCl and 20 mM D-lactate (Na⁺ salt) as electron donor at 25 °C under aerobic conditions using a rapid filtration method. (a) Initial rates of proline uptake. (b) Steady-state levels of proline accumulation. All values are expressed as percentage of the corresponding value of PutP(Δ Cys) (initial rate, 23 \pm 5 nmol/min-mg; steady-state level of proline accumulation, 24 \pm 3 nmol/mg). Standard deviations were calculated from a minimum of three independent experiments.

A more detailed kinetic analysis of the functional consequences of the Cys substitutions revealed diverse effects on Michaelis-Menten parameters (Table 3.1.).

Replacement of Gly328 by Cys affected V_{max} (3-fold decreased) and left $K_{m(\text{pro})}$ and $K_{0.5(\text{Na}^+)}$ almost unaltered compared to PutP(wild type). PutP-S332C and -L346C exhibited decreased $K_{m(\text{pro})}$ (3- and 21-fold, respectively) and V_{max} values (22- and 8-fold, respectively). While $K_{0.5(\text{Na}^+)}$ of PutP-L346C was only slightly altered, the value could not be reliably determined for PutP-S332C due to the low transport activity (signal-to-noise ratio) in intact cells. Therefore, the analysis was performed with proteoliposomes containing the purified transporter in an inside-out orientation. Under these conditions, S332C exhibited a $K_{0.5(\text{Na}^+)}$, which was 13-fold reduced compared to the wild type in proteoliposomes (Table 3.1.). In

contrast to PutP-S332C and -L346C, placement of Cys at the position of Gln345 increased $K_{m(\text{pro})}$ and $K_{0.5(\text{Na}^+)}$ 38- and 19-fold, respectively, compared to PutP(wild type). V_{max} was about 2-fold reduced (Table 3.1.). The latter results fit to previous analyses of substitutions for Ser340 and Thr341, albeit the impact of alterations at the latter positions on $K_{m(\text{pro})}$ and $K_{0.5(\text{Na}^+)}$ was even stronger than that of the Gln345 replacement (Hilger *et al.*, 2008). Finally, placement of Cys at the position of Ala337, which was shown to stimulate proline uptake in the initial analyses, slightly increased $K_{0.5(\text{Na}^+)}$ and V_{max} (3- and 1.6-fold) and left $K_{m(\text{pro})}$ unaltered (Table 3.1.).

Table 3.1.: Proline uptake kinetics of PutP bearing replacements of given amino acid residues

PutP	$K_{m(\text{pro})}$ μM	V_{max} $\text{nmol/min} \times \text{mg of protein}$	$K_{0.5(\text{Na}^+)}$ mM
Wild-type	1.3±0.2	44.6±1.6	0.038±0.007 (0.7±0.09) ^a
G328C	1.5±0.5	16.7±1.0	0.018±0.004
S332C	0.4±0.3	2.1±0.2	n.s. ^b (0.055±0.014) ^a
A337C	1.7±0.5	72.5±3.8	0.117±0.025
Q345C	49.3±1.0	23.0±1.8	0.705±0.088
L346C	0.063±0.016	5.8±0.3	0.071±0.011

To determine $K_{m(\text{pro})}$ and V_{max} , we measured initial rates of L-[U-¹⁴C]proline uptake by *E. coli* WG170 producing either PutP(wild type) or PutP with given replacements in the wild-type background in the presence of 50 mM NaCl and 20 mM D-lactate (Na⁺ salt) at proline concentrations from 0.2 to 500 μM . For determination of the ion specific parameters [$K_{0.5(\text{Na}^+)}$], transport of 10 μM L-[U-¹⁴C]proline was measured in the presence of 0.005 to 250 mM NaCl or LiCl at 25 °C. The resulting data were plotted according to Eadie-Hofstee.

^a Values in parentheses were obtained with proteoliposomes containing the purified PutP variants.

^b n.s., no stimulation of proline uptake into intact cells by increasing concentrations of Na⁺.

3.4.3. Influence of Cys modification on transport activity

Site-directed sulfhydryl modification of single-Cys proteins *in situ* has been shown to be particularly useful for studying both static and dynamic features of transporters (Guan and Kaback, 2007; Kaback *et al.*, 2007; Karlin and Akabas, 1998). Here, we used the method to probe a potential participation of TM IX of PutP in the formation of a hydrophilic cavity in the transporter. In a first approach, accessibility of Cys placed throughout TM IX to membrane-permeant *N*-ethylmaleimide (NEM) and membrane-impermeant methanethiosulfonate ethyltrimethylammonium (MTSET) was analyzed via the impact of the reagents on proline uptake into intact cells. Only freshly prepared cells were used for transport and additional energization by D-lactate was omitted to minimize indirect inhibitory effects (e.g., caused by reaction of NEM or MTSET with components of the respiratory chain). In this way, NEM- or MTSET-treated cells containing PutP(Δ Cys) maintained a minimum of 80% and 96%, respectively, of the activity of untreated cells. PutP(Δ Cys) was not chemically modified as demonstrated below and elsewhere (Hilger *et al.*, 2008).

NEM avidly reacted with Cys at distinct positions in the cytoplasmic half of TM IX (C344, L347C, and S351C) as indicated by the highly reduced activities of the respective

NEM-treated cells (8% or less remaining activity compared to untreated cells) (Figs. 3.4.a and 3.5.). Intermediate inhibitory effects of NEM were observed with cells containing PutP(Δ Cys)-T341C or -V348C (40% and 48% remaining activity, respectively). The activity of all other single-Cys PutP variants was not or only slightly altered by NEM treatment. Here, either Cys was not accessible to NEM or labeling had only a minor impact on transport activity. Since alkylation reagents are known to react with sulfhydryl groups in polar but not in apolar environments (Zhou *et al.*, 1995; Poelarends and Konings, 2002), the observed NEM inhibition pattern suggested that Thr341, Cys344, Leu347, Val348, and Ser351 are in contact with a polar (hydrophilic) environment.

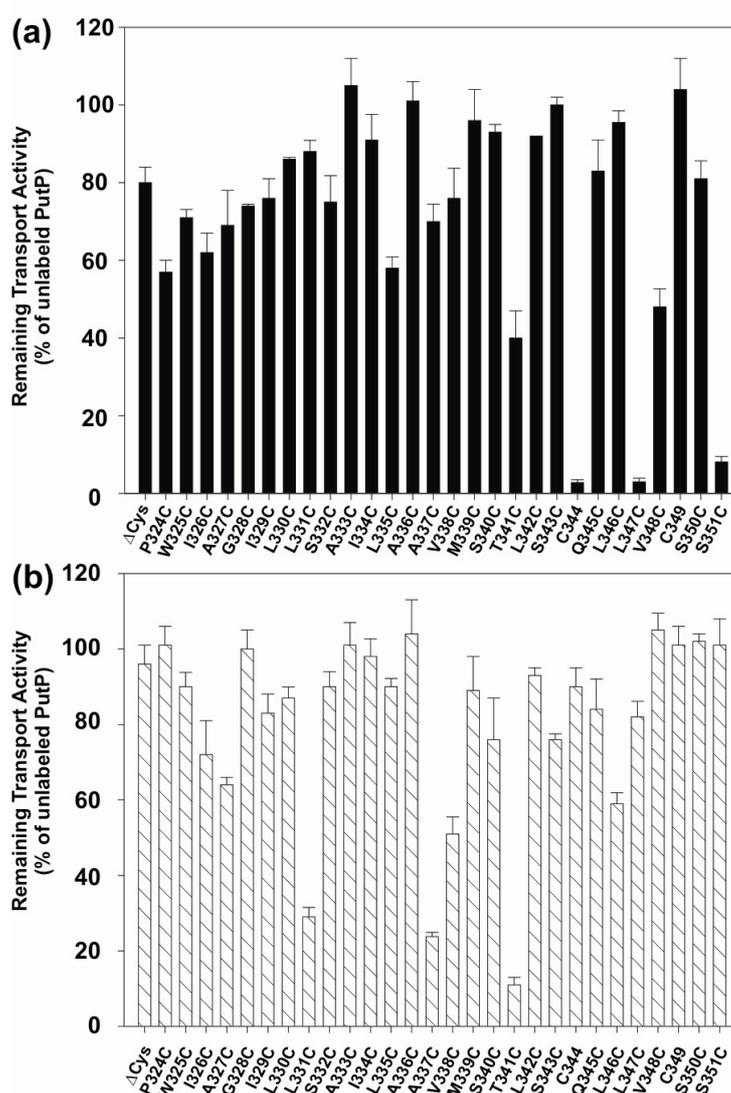


Figure 3.4.: Influence of sulfhydryl reagents on proline uptake into cells containing single-Cys PutP variants of TM IX. Cells were incubated with 2 mM sulfhydryl reagent (NEM and MTSET) at room temperature for 15 min. The reaction was stopped by dilution with 100 mM Tris/Mes buffer, pH 6.0, containing 0.1% BSA. Cells were washed and resuspended in the same buffer without BSA, and transport was analyzed as described in the legend to Fig. 3.3.. Activities of NEM-treated (a) or MTSET-treated (b) cells are presented as percentage of the activity of the respective unlabeled PutP variant. Standard deviations were calculated from a minimum of three measurements.

The inhibition pattern of membrane-impermeant, positively charged MTSET differed significantly from the NEM inhibition pattern (Figs. 3.4.b and 3.5.). MTSET inhibited transport of single-Cys PutP variants when Cys was placed in the periplasmic half (L331C, 29% remaining activity) or the middle of TM IX (A337C, 23%; T341C, 11% remaining activity). Also, PutP(Δ Cys)-A327C and -V338C were inhibited by MTSET albeit to a significantly lesser extent (64% and 51% remaining activity, respectively). Contrary to NEM, MTSET did not inhibit single-Cys PutP activity when Cys was placed into the cytoplasmic half of the domain (Figs. 3.4.b and 3.5.). These results suggested that the transporter contained a narrow cavity (funnel), allowing the positively charged MTSET with a width similar to the diameter of hydrated Na^+ to reach amino acids in the middle of TM IX from the periplasmic side of the membrane. The more rigid and, as compared to MTS reagents, chemically less reactive NEM presumably could not efficiently enter this cavity or the transporter variants were insensitive to Cys modification by the neutral reagent.

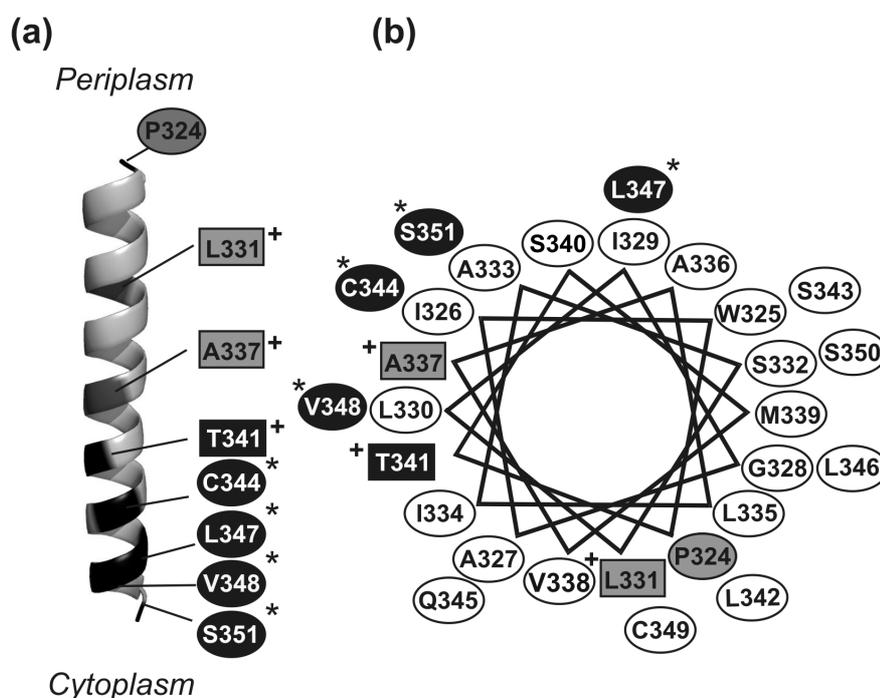


Figure 3.5.: Schematic presentation of Cys accessibility in TM IX. The figure summarizes the Cys accessibility data shown in Figs. 3.4., 3.6.-3.8.. (a) Side view of TM IX. (b) Helical wheel projection of TM IX. The helix was visualized by PyMOL. Black ellipses indicate single-Cys PutP variants inhibited by NEM and modified by FM; gray ellipses denote single-Cys PutP variants not inhibited by NEM but modified by FM; white ellipses represent single-Cys PutP variants not or only marginally inhibited or modified by NEM and FM; gray rectangles indicate single-Cys PutP variants inhibited by MTSET; black rectangles denote single-Cys PutP variant inhibited by NEM and MTSET and modified by FM. *Cys protected by proline; †Cys partially protected by Na^+ . The accessibility of the remaining residues was not affected by Na^+ or proline.

3.4.4. Accessibility of Cys to fluorescein-5-maleimide

Since the analysis of transport inhibition did not allow discrimination between unlabeled and labeled but functionally unaffected proteins, Cys modification was directly determined by analyzing fluorescence of single-Cys PutP variants incubated with membrane-permeant fluorescein-5-maleimide (FM). Labeling was performed with randomly oriented

membrane vesicles; labeled PutP derivatives were solubilized from the membranes, purified by Ni²⁺-NTA affinity chromatography and subjected to SDS-PAGE. Fluorescence was detected by UV excitation and normalized to the protein amount. The normalized fluorescence intensity of PutP(Δ Cys)-T341C was arbitrarily set to 100. Unspecific labeling was estimated with PutP(Δ Cys) as negative control and corresponded to three arbitrary units (AU) or less (Fig. 3.6.).

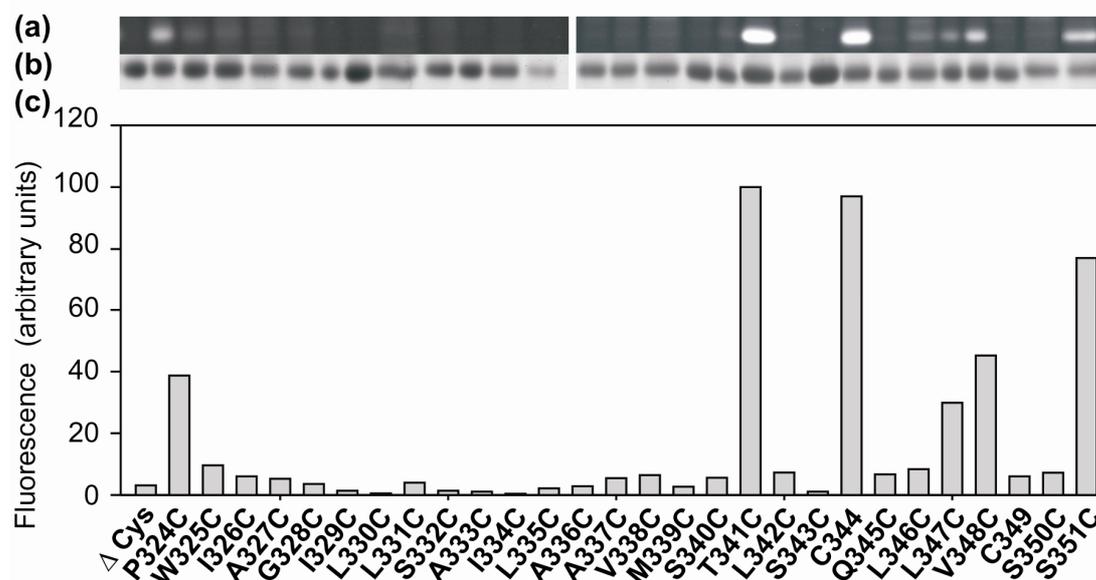


Fig. 3.6.: Accessibility of Cys individually introduced throughout TM IX to FM. Randomly oriented membrane vesicles containing the indicated PutP derivative were prepared as described in Materials and Methods. One hundred fifty microliters of a suspension of membranes of *E. coli* WG170 containing given PutP derivatives (10 mg of total membrane protein/mL) were incubated with 200 μ M FM for 10 min. Labeling reactions were stopped by addition of 5 mM β -mercaptoethanol. After solubilization with 1.5% (w/v) *n*-dodecyl- β -D-maltoside and purification via Ni²⁺-NTA affinity chromatography, 5 μ g protein was loaded onto 10% SDS-PAGE and separated. Fluorescent bands were detected using a MultimagerTM (BioRad). Afterwards, the same gel was stained with Coomassie Blue. Fluorescence intensities detected by Multimager were corrected for background fluorescence and quantitatively analyzed using software ImageQuant 5.0 and plotted. (a) Fluorescent protein bands. (b) Coomassie Blue-stained protein. (c) Relative fluorescence intensities normalized to the amount of protein. PutP(Δ Cys)-T341C showed the highest fluorescence intensity, and the respective value was arbitrarily set to 100. Relative fluorescence intensity (rel. fl.) is expressed in arbitrary units (AU).

FM reacted essentially with all single-Cys PutP variants inhibited by NEM [PutP(Δ Cys)-T341C, -C344, -L347C, -V348C, and -S351C; fluorescence intensities between 30 and 100 AU] (Figs. 3.4.-3.6.). In addition, PutP(Δ Cys)-P324C was modified by FM (fluorescence intensity, 40 AU). Pro324 is located at the periplasmic border of TM IX. Following TM IX from the periplasmic side, accessibility decreased abruptly, and Cys at positions 325 to 340 was not significantly accessible to the sulfhydryl reagent (relative labeling yields between 0.4% and 7%) including positions at which Cys was modified by MTSET (Figs. 3.5. and 3.6.). The only single-Cys PutP variant modified by all three sulfhydryl reagents was PutP(Δ Cys)-T341C. As alkylation by maleimide-based reagents requires the sulfhydryl group to be in the deprotonated form (Zhou *et al.*, 1995; Poelarends and Konings, 2002), the high accessibility and reactivity of Cys placed into the middle or cytoplasmic half of TM IX suggest that the respective residues are exposed to a hydrophilic environment as already deduced above from the NEM inhibition pattern.

3.4.5. Influence of ligand on Cys accessibility

The accessibility analyses described above suggest that various residues of TM IX are exposed to a hydrophilic environment as they originate, for example, from hydrophilic cavities or funnels formed by the transport protein. To explore the physiological significance of the putative cavities, the influence of ligands (Na^+ , proline) on Cys accessibility was tested throughout TM IX. For this purpose, FM labeling of randomly oriented membrane vesicles containing individual single-Cys PutP variants was performed in the absence and presence of ligand(s). Subsequently, PutP was isolated and FM fluorescence was detected as described above.

FM labeling of Cys at positions 324 and 341 was not significantly influenced by ligands (Fig. 3.7.). However, labeling of the accessible positions 344, 347, 348, and 351 was inhibited by proline in the presence of Na^+ , while Na^+ alone had no effect (Fig. 3.7.). Also, proline alone inhibited Cys modification at all four positions albeit somewhat less efficient than upon addition of Na^+ (data not shown). Finally, neither Na^+ nor proline stimulated reaction of FM with Cys at all remaining positions reported inaccessible in the absence of ligands (data not shown).

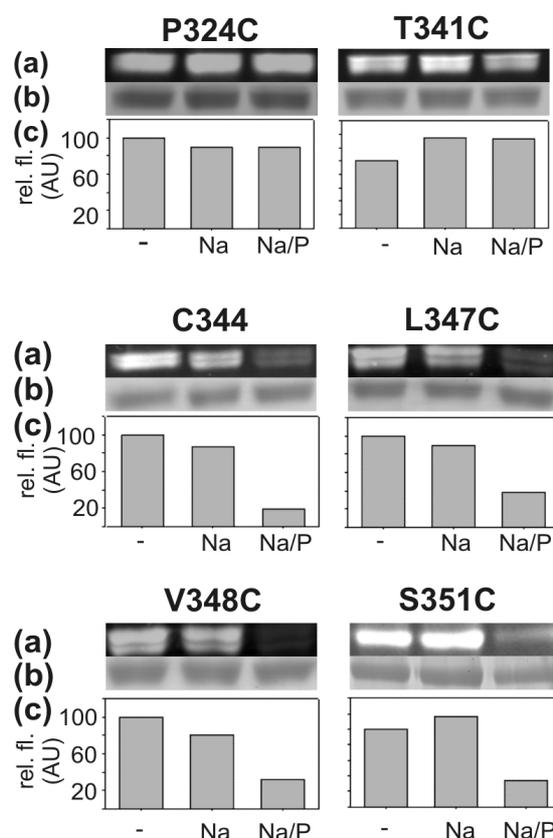


Figure 3.7.: Influence of ligands on Cys accessibility to FM. Membrane vesicles containing the indicated PutP derivatives were prepared as described in the legend to Fig. 3.5.. Aliquots (150 μL) of a suspension of membranes of *E. coli* WG170 containing given PutP(ΔCys) derivatives (10 mg of total membrane protein/mL) were preincubated in the absence (–) or presence of 30 mM NaCl (Na) and 30 mM NaCl/10 mM proline (Na/P) at 25 $^{\circ}\text{C}$ for 10 min. Subsequently, labeling of the samples with 200 μM FM was performed for 10 min. The reaction was stopped by addition of 5 mM β -mercaptoethanol. Labeled PutP was isolated and analyzed as described in the legend to Fig. 3.4.. (a) Fluorescent protein bands. (b) Coomassie Blue-stained protein. (c) Relative fluorescence intensities (rel. fl.) in arbitrary units (AU) normalized to the amount of protein.

The effect of ligands on MTSET labeling of PutP(Δ Cys)-L331C, -A337C, and -T341C was investigated. For this purpose, freshly prepared cells containing the respective single-Cys PutP derivatives were labeled with MTSET in the presence and absence of ligands, and proline uptake was analyzed after cell washing. The effect of ligand on MTSET labeling is expressed as the ratio of the activities measured after MTSET labeling in the presence and absence of ligand to better compare protective effects between the respective proteins (Fig. 3.8.). In contrast to the protective effect of proline on Cys in the cytoplasmic half of TM IX described above, the substrate alone did not significantly inhibit labeling of Cys placed into the periplasmic half or middle of TM IX (position 331, 337, or 341) with MTSET (Fig. 3.8.). Instead, the results of the activity assay suggested a partial protective effect of Na^+ . In fact, MTSET labeling of cells containing PutP(Δ Cys)-L331C, -A337C, or -T341C in the presence of Na^+ led to proline uptake rates that were 1.4-, 1.7-, or 1.4-fold higher, respectively, than that of cells labeled in the absence of Na^+ (Fig. 3.8.). The protective effect was somewhat strengthened than when both Na^+ and proline were present during the labeling reaction. However, it should be noted that protection of Cys at position 341 was not very efficient as even under conditions of maximum protection, uptake activities of MTSET-treated cells were reduced to 30% of the value of unlabeled cells. Cells containing PutP(Δ Cys)-L331C or -337C showed a maximum of 60-70% of the activity of unlabeled cells if labeling was performed in the presence of Na^+ and proline.

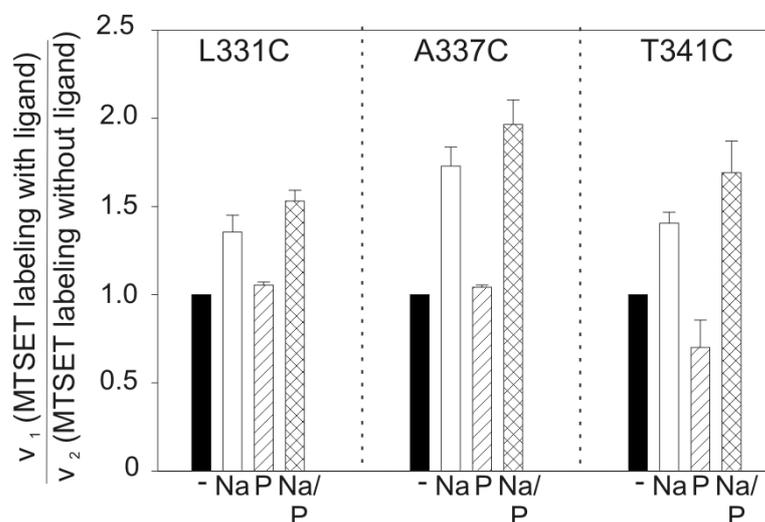


Figure 3.8.: Influence of ligands on Cys accessibility to MTSET. Cells containing the respective single-Cys variant were preincubated in the absence (-) or presence of 30 mM NaCl (Na), 10 mM proline, or 30 mM NaCl/10 mM proline (Na/P) at 25 °C for 10 min. Subsequently, 2 mM MTSET was added and incubation was continued for 10 min. The reaction was stopped, cells were washed and transport was analyzed as described in the legend to Fig. 3.4.. The effect of ligand is shown as the ratio of the activities measured after MTSET labeling in the presence (v_1) and absence (v_2) of ligand. Ratios > 1 indicate inhibition, whereas ratios < 1 denote stimulation of MTSET labeling.

3.5. Discussion

Ser340 and Thr341 in TM IX of PutP have previously been suggested to play key roles in ligand binding and transport; thereby, the side chain of Thr341 may directly

participate in Na⁺ binding (Hilger *et al.*, 2008). The current analysis of structure-function relationships in PutP newly identified Gly328, Ser332, Gln345, and Leu346 of the same domain as important for Na⁺-coupled proline uptake. In addition, the Cys accessibility pattern of TM IX suggests the existence of hydrophilic cavities in the transporter that are affected by ligand binding.

What are the roles of the newly identified structural and/or functional important residues? While none of the four residues appears to be crucial for membrane insertion and stability of PutP, functional analyses of the effect of Cys replacements reveal significant alterations of kinetic parameters of transport. $K_{m(\text{pro})}$ is most severely affected (more than one order of magnitude) upon replacement of Gln345 or Leu346. Although the effects at both positions are diametrical, the data suggest that ligand binding is highly sensitive to alterations in this part of the protein. In particular, the properties of PutP-Q345C [increased $K_{m(\text{pro})}$ and $K_{0.5(\text{Na}^+)}$, comparatively small changes of V_{max}] are, to some extent, reminiscent of the results obtained upon replacement of Ser340 or Thr341 (Hilger *et al.*, 2008). These observations further support the idea that these polar residues of TM IX are located at or close to a ligand-binding site of PutP. Furthermore, as both $K_{m(\text{pro})}$ and $K_{0.5(\text{Na}^+)}$ are highly affected upon replacement of Ser340, Thr341, or Gln345, the idea that binding of coupling ion and substrate occurs in close proximity is supported. Furthermore, the reduced V_{max} values observed with PutP-G328C, -S332C, and -L346C hint at defects in the transport cycle subsequent to ligand binding, for example, inhibition of conformational alterations or a hampered release of ligands. The latter idea is, in principle, supported by the decreased Michaelis-Menten parameters found for PutP-S332C [$K_{m(\text{pro})}$ and $K_{0.5(\text{Na}^+)}$] and PutP-L346C [$K_{m(\text{pro})}$] (Table 3.1.).

If, in fact, residues of TM IX are located at or close to a ligand-binding site, they might be accessible from one side and/or from the other side of membrane. First evidence for the existence of an outwardly oriented cavity (funnel) in PutP comes from labeling of PutP-S340C and -T341C in right-side-out membrane vesicles with various methane thiosulfonate compounds (Hilger *et al.*, 2008). Here, we present new evidence for the existence of a funnel that may connect Thr341 with the periplasmic surface of PutP. This evidence comes from the observation that Na⁺-coupled proline uptake into intact cells containing PutP(Δ Cys)-L331C, -A337C, or -T341C is efficiently inhibited by MTSET. Na⁺ but not proline inhibits the labeling reaction, although it cannot completely protect the respective Cys residues from modification possibly due to the dynamics of the system. Considering the Na⁺-sensitive accessibility of these residues to positively charged MTSET with a width similar to the diameter of hydrated Na⁺, as well as the previously reported particular significance of Thr341 for Na⁺ binding (Hilger *et al.*, 2008), it is tempting to speculate that the proposed funnel is used by Na⁺ as an entrance to a Na⁺-binding site in about the middle of the membrane (Fig. 3.9.). Clearly,

inhibition of MTSET labeling of the reported residues could also be achieved by conformational alterations induced by Na⁺ binding elsewhere.

Unlike the periplasmic half of TM IX, Cys placed at positions in the cytoplasmic half of the domain (C344, L347C, V348C, and S351C) is highly accessible to NEM and FM but not to MTSET (if MTSET is added from the periplasmic side of the membrane). The labeling pattern suggests the existence of an inwardly oriented hydrophilic cavity (Fig. 3.9.). Cavities open to the cytoplasm have also been seen, for example, by Cys accessibility analyses and X-ray crystallography of lactose permease and glycerol-3-phosphate transporter (both members of the major facilitator family), and are suggested for members of the neurotransmitter/sodium symporter family (TC 2.A.22, SLC6) (Kaback *et al.*, 2007; Abramson *et al.*, 2003; Huang *et al.*, 2003; Guan *et al.*, 2007; Quick *et al.*, 2006; Zhang and Rudnick, 2006). In the case of PutP, proline but not Na⁺ effectively blocks modification of Cys at the respective positions. Cys protection may be achieved by direct steric hindering or closure of the cavity due to either movement of neighboring TMs relative to each other or a gating mechanism. The latter mechanism has recently been suggested for NIS and members of the neurotransmitter/sodium symporter family and is supposed to involve amino acids of adjoining loops that participate in salt bridge formation (De la Vieja *et al.*, 2007; Yamashita *et al.*, 2005; Loland *et al.*, 2004). For lactose permease, site-directed spin labeling and EPR spectroscopy revealed a substrate-induced closure of an inwardly oriented cytoplasmic cavity (Smirnova *et al.*, 2007). Cavities of this kind are frequently thought to participate in the release of ligands from the transporter into the cytoplasm.

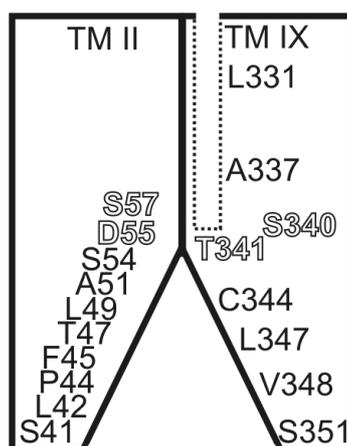


Figure 3.9.: Model showing the participation of TMs II and IX in the formation of hydrophilic cavities in PutP. The model is based on the current analysis as well as on previous investigations (Quick and Jung, 1997; Quick *et al.*, 1996; Pirch *et al.*, 2002; Hilger *et al.*, 2008; Pirch *et al.*, 2003). Amino acids of TMs II and IX supposed to line a hydrophilic cavity (funnel) are indicated. Residues shown to be important for ligand binding and/or translocation are represented as open characters. There is no evidence so far that the periplasmic half of TM II participates in the formation of the outwardly oriented funnel (drawn as dotted box) suggested to be lined by residues of TM IX.

Thr341 is the only position at which Cys is highly accessible to all three sulfhydryl reagents (MTSET, NEM, and FM). In this context, it appears interesting to note that a recent backbone model of TM IX generated based on intrahelical distance measurements by EPR

spectroscopy shows a pronounced kink near Thr341, which is associated with a short loop-like structure (Hilger *et al.*, submitted). While these results agree, in principle, with the idea of Thr341 being located at the interface between periplasmic and cytoplasmic cavities, a direct involvement in the transition of ligands between these cavities remains elusive.

Alternatively to the cavity model, the accessibility of Cys to sulfhydryl reagents could be taken as a hint at a location of respective positions in a cytoplasmic or periplasmic loop rather than in a TM. The following arguments are contrary to this interpretation: (i) the *in silico* predicted location and orientation of TM IX of PutP is confirmed by *phoA* and *lacZ* gene fusion analyses and site-specific proteolysis (Jung *et al.*, 1998a); (ii) the FM labeling pattern of the periplasmic half of TM IX agrees perfectly with the current topology model of PutP (Figs. 3.1., 3.5. and 3.6.); (iii) the FM labeling pattern of the cytoplasmic half of TM IX hints at an α -helical arrangement of this part of the domain (Figs. 3.5. and 3.6.); (iv) intrahelical distances determined by site-directed spin labeling and EPR spectroscopy are in full agreement with a kinked α -helical backbone structure of TM IX within the borders indicated in the current topology model (Hilger *et al.*, submitted); (v) the membrane topography of TM II is confirmed by site-directed spin labeling and EPR spectroscopy (Wegener *et al.*, 2000) and Cys cross-linking (Hilger *et al.*, 2008) aligns TMs II and IX within the borders of the current topology model; and (vi) our current model is in agreement with topological analyses of other SSSF members (Turk *et al.*, 1996; Levy *et al.*, 1998).

Finally, it should be noted that a Cys accessibility pattern very similar to the cytoplasmic half of TM IX was previously observed for TM II (Fig. 3.1.) (Pirch *et al.*, 2003). Proline blocked modification of Cys at respective positions in TM II as detected here for TM IX. Since pairs of Cys residues placed in both domains can efficiently be cross-linked even by the “zero length” cross-linker $\text{Cu}(1,10\text{-phenanthroline})_3$ (Hilger *et al.*, 2008), we suggest that TM IX is joined by TM II in forming an inwardly oriented cavity (Fig. 3.9.).

Taken together, the results indicate that Gly328, Ser332, Gln345, and Leu346 in TM IX of PutP are important for ligand binding and/or translocation in addition to the previously identified Ser340 and Thr341. Furthermore, it is suggested that distinct residues of the domain participate in the formation of an outwardly oriented funnel (Leu331, Ala337, and Thr341) and a larger inwardly directed cavity (Thr341, Cys344, Leu347, Val348, and Ser351).

3.6. Acknowledgments

We thank Maret Böhm for generating PutP-S332C. This work was financially supported by the Deutsche Forschungsgemeinschaft (Ju333/3-2, Ju333/4-2 and Exc114-1).

3.7. References

- Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R., and Iwata, S.** (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* 301: 610-615.
- Bradford, M. M.** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248-254.
- Chen, C. C., Tsuchiya, T., Yamane, Y., Wood, J. M., and Wilson, T. H.** (1985) Na⁺ (Li⁺)-proline cotransport in *Escherichia coli*. *J. Membr. Biol.* 84: 157-164.
- De la Vieja, A., Reed, M. D., Ginter, C. S., and Carrasco, N.** (2007) Amino acid residues in transmembrane segment IX of the Na⁺/I⁻ symporter play a role in its Na⁺ dependence and are critical for transport activity. *J. Biol. Chem.* 282: 25290-25298.
- Dohán, O., De la Vieja, A., and Carrasco, N.** (2006) Hydrocortisone and purinergic signaling stimulate sodium/iodide symporter (NIS)-mediated iodide transport in breast cancer cells. *Mol. Endocrinol.* 20: 1121-1137.
- Guan, L., and Kaback, H. R.** (2007) Site-directed alkylation of cysteine to test solvent accessibility of membrane proteins. *Nat. Protoc.* 2: 2012-2017.
- Guan, L., Mirza, O., Verner, G., Iwata, S., and Kaback, H. R.** (2007) Structural determination of wild-type lactose permease. *Proc. Natl. Acad. Sci. U. S. A.* 104: 15294-15298.
- Hilger, D., Böhm, M., Hackmann, A., and Jung, H.** (2008) Role of Ser-340 and Thr-341 in transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli* in ligand binding and transport. *J. Biol. Chem.* 283: 4921-4929.
- Hilger, D., Polyhach, Y., Jung, H., and Jeschke, G.** (2009) Backbone structure of transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli*. *Biophys. J.* submitted.
- Huang, Y., Lemieux, M. J., Song, J., Auer, M., and Wang, D. N.** (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* 301: 616-620.
- Jackowski, S., and Alix, J. H.** (1990) Cloning, sequence, and expression of the pantothenate permease (*panF*) gene of *Escherichia coli*. *J. Bacteriol.* 172: 3842-3848.
- Jung, H., Rübenhagen, R., Tebbe, S., Leifker, K., Tholema, N., Quick, M., and Schmid, R.** (1998a) Topology of the Na⁺/proline transporter of *Escherichia coli*. *J. Biol. Chem.* 273: 26400-26407.
- Jung, H., Tebbe, S., Schmid, R., and Jung, K.** (1998b) Unidirectional reconstitution and characterization of purified Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 37: 11083-11088.

- Jung, H.** (2001) Towards the molecular mechanism of Na⁺/solute symport in prokaryotes. *Biochim. Biophys. Acta* 1505: 131-143.
- Jung, H.** (2002) The sodium/substrate symporter family: structural and functional features. *FEBS Lett.* 529: 73-77.
- Kaback, H. R., Dunten, R., Frillingos, S., Venkatesan, P., Kwaw, I., Zhang, W., and Ermolova, N.** (2007) Site-directed alkylation and the alternating access model for LacY. *Proc. Natl. Acad. Sci. U. S. A.* 104: 491-494.
- Karlin, A., and Akabas, M. H.** (1998) Substituted-cysteine accessibility method. *Methods. Enzymol.* 293: 123-145.
- Kojima, S., Bohner, A., Gassert, B., Yuan, L., and Wiren, N. v.** (2007) AtDUR3 represents the major transporter for high-affinity urea transport across the plasma membrane of nitrogen-deficient *Arabidopsis* roots. *Plant J.* 52: 30-40.
- Levy, O., De la Vieja, A., Ginter, C. S., Riedel, C., Dai, G., and Carrasco, N.** (1998) N-linked glycosylation of the thyroid Na⁺/I⁻ symporter (NIS). Implications for its secondary structure model. *J. Biol. Chem.* 273: 22657-22663.
- Loland, C. J., Granas, C., Javitch, J. A., and Gether, U.** (2004) Identification of intracellular residues in the dopamine transporter critical for regulation of transporter conformation and cocaine binding. *J. Biol. Chem.* 279: 3228-3238.
- Peterson, G. L.** (1977) A simplification of the protein assay method of Lowry *et al.* which is more generally applicable. *Anal. Biochem.* 83: 346-356.
- Pirch, T., Quick, M., Nietschke, M., Langkamp, M., and Jung, H.** (2002) Sites important for Na⁺ and substrate binding in the Na⁺/proline transporter of *Escherichia coli*, a member of the Na⁺/solute symporter family. *J. Biol. Chem.* 277: 8790-8796.
- Pirch, T., Landmeier, S., and Jung, H.** (2003) Transmembrane domain II of the Na⁺/proline transporter PutP of *Escherichia coli* forms part of a conformationally flexible, cytoplasmic exposed aqueous cavity within the membrane. *J. Biol. Chem.* 278: 42942-42949.
- Poelarends, G., and Konings, W. N.** (2002) The transmembrane domains of the ABC multidrug transporter LmrA form a cytoplasmic exposed, aqueous chamber within the membrane. *J. Biol. Chem.* 277: 42891-42898.
- Quick, M., Tebbe, S., and Jung, H.** (1996) Ser57 in the Na⁺/proline permease of *Escherichia coli* is critical for high-affinity proline uptake. *Eur. J. Biochem.* 239: 732-736.
- Quick, M., and Jung, H.** (1997) Aspartate 55 in the Na⁺/proline permease of *Escherichia coli* is essential for Na⁺-coupled proline uptake. *Biochemistry* 36: 4631-4636.
- Quick, M., Yano, H., Goldberg, N. R., Duan, L., Beuming, T., Shi, L., Weinstein, H., and Javitch, J. A.** (2006) State-dependent conformations of the translocation pathway in the tyrosine transporter Tyt1, a novel neurotransmitter:sodium symporter from *Fusobacterium nucleatum*. *J. Biol. Chem.* 281: 26444-26454.

- Reizer, J., Reizer, A., and Saier, M. H., Jr.** (1994) A functional superfamily of sodium/solute symporters. *Biochim. Biophys. Acta* 1197: 133-166.
- Smirnova, I., Kasho, V., Choe, J. Y., Altenbach, C., Hubbell, W. L., and Kaback, H. R.** (2007) Sugar binding induces an outward facing conformation of LacY. *Proc. Natl. Acad. Sci. U. S. A.* 104: 16504-16509.
- Stalmach, M. E., Grothe, S., and Wood, J. M.** (1983) Two proline porters in *Escherichia coli* K-12. *J. Bacteriol.* 156: 481-486.
- Tabor, S., and Richardson, C. C.** (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U. S. A.* 82: 1074-1078.
- Turk, E., Kerner, C. J., Lostao, M. P., and Wright, E. M.** (1996) Membrane topology of the human Na⁺/glucose cotransporter SGLT1. *J. Biol. Chem.* 271: 1925-1934.
- Wegener, C., Tebbe, S., Steinhoff, H.-J., and Jung, H.** (2000) Spin labeling analysis of structure and dynamics of the Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 39: 4831-4837.
- Wood, J. M., and Zadworny, D.** (1979) Characterization of an inducible porter required for L-proline catabolism by *Escherichia coli* K12. *Can. J. Biochem.* 57: 1191-1199.
- Wright, E. M., and Turk, E.** (2004) The sodium/glucose cotransport family SLC5. *Pflugers Arch.* 447: 510-518.
- Wright, E. M., Hirayama, B. A., and Loo, D. F.** (2007) Active sugar transport in health and disease. *J. Intern. Med.* 261: 32-43.
- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E.** (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* 437: 215-223.
- Yamato, I., and Anraku, Y.** (1993) Na⁺/substrate symport on prokaryotes. In *Alkali cation transport systems in prokaryotes* (Bakker, E. P., ed.) pp. 53-76, CRC-Press, Boca Raton.
- Yanisch-Perron, C., Vieira, J., and Messing, J.** (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33: 103-119.
- Zhang, Y. W., and Rudnick, G.** (2006) The cytoplasmic substrate permeation pathway of serotonin transporter. *J. Biol. Chem.* 281: 36213-36220.
- Zhou, J., Fazio, R. T., and Blair, D. F.** (1995) Membrane topology of the MotA protein of *Escherichia coli*. *J. Mol. Biol.* 251: 237-242.

CHAPTER 4**Backbone structure of transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli*****4.1. Abstract**

The backbone structure is determined by site-directed spin labeling, double electron-electron resonance measurements of distances, and modeling in terms of a helix-loop-helix construct for a transmembrane domain that is supposed to line the translocation pathway in the 54.3 kDa Na⁺/proline symporter PutP of *Escherichia coli*. The conformational distribution of the spin labels is accounted for by a rotamer library. An ensemble of backbone models with a root mean-square deviation of less than 2 Å is obtained. These models exhibit a pronounced kink near residue T341, which is involved in substrate binding. The kink may be associated with a hinge that allows the protein to open and close an inwardly oriented cavity.

4.2. Introduction

Despite recent advances (Bowie, 2005; Opella and Marassi, 2004; Tamm and Liang, 2006), the process of determining the three-dimensional structures of membrane proteins is still highly challenging and progress is rather slow (Lacapere *et al.*, 2007). This applies in particular to integral α -helical membrane proteins with more than a few transmembrane (TM) domains that are too large for current NMR approaches and are often very difficult to crystallize. Many transporters, which are potential pharmaceutical targets, belong to this category of proteins. New techniques for their structural characterization are thus required. Site-directed spin labeling (SDSL) can complement crystallographic techniques because membrane proteins can be studied in liposomes or detergent micelles, and NMR because there is no significant limitation to the size of the protein (Hubbell *et al.*, 2000). For instance, a structural model of the KcsA channel in its open state has been derived from the crystal structure of the closed state and continuous-wave electron paramagnetic resonance (EPR) measurements on spin-labeled mutants (Liu *et al.*, 2001). By combining SDSL with pulsed EPR techniques, longer-range distances can be measured and distance distributions derived (Jeschke and Polyhach, 2007; Schiemann and Prisner, 2007). Thus, a detailed picture of large-amplitude conformational motion in the ATP-binding cassette transporter MsbA can be obtained (Borbat *et al.*, 2007). In the lactose permease LacY, sugar-triggered conformational rearrangements on both sides of the molecule have been observed, strongly supporting the alternative access mechanism of transport (Smirnova *et al.*, 2007). Furthermore, the structures of the soluble protein complex of histidine kinase CheA and its coupling protein

CheW (Bhatnagar *et al.*, 2007), and of the dimer of the Na⁺/H⁺ antiporter NhaA of *Escherichia coli* (Hilger *et al.*, 2007) have been determined by rigid-body refinement based on structures of the component proteins. The latter study demonstrated that the use of labels with the size of a few Ångstroms does not pose a resolution limit if the conformational distribution of the label is modeled (Jeschke and Polyhach, 2007; Sale *et al.*, 2005) so that distance distributions can be fitted to the experimental data.

In this work we extend this approach to *ab initio* structure determination. Specifically, we work out the backbone structure of TM domain IX of the Na⁺/proline transporter PutP of *E. coli* by obtaining distance distribution measurements for 16 spin-labeled double mutants using the four-pulse double electron-electron resonance (DEER) technique (Pannier *et al.*, 2000). PutP belongs to the sodium/solute symporter family (SSSF, TC 2.A.21, SLC5), which contains several hundred members of prokaryotic and eukaryotic origin (Jung, 2002). Proteins of this family utilize a sodium motive force to drive uphill transport of substrates such as sugars, amino acids, vitamins, ions, *myo*-inositol, phenyl acetate, and urea. Some transporters are implicated in human disease (e.g., Na⁺/glucose transporter (SGLT1), Na⁺/iodide symporter (NIS)) or play an important role in medical therapy (e.g., NIS) (Dohán *et al.*, 2006, Wright *et al.*, 2007). Furthermore, bacterial transporters such as the Na⁺/proline transporter PutP of *Staphylococcus aureus* contribute to virulence (Schwan *et al.*, 2006). Until very recently, no high-resolution structure was available for any member of the protein family. The first X-ray structure in this protein family, for the sodium/galactose transporter vSGLT of *Vibrio parahaemolyticus* (Faham *et al.*, 2008), is surprisingly analogous to the structure of the leucine transporter LeuT_{Aa} of *Aquifex aeolicus* from the neurotransmitter/sodium symporter (NSS) family (Yamashita *et al.*, 2005), although proteins from both families are dissimilar on the sequence level.

Recent kinetic and protein chemical analyses suggest that amino acids of TM domain IX (e.g., S340 and T341) form part of the ion- and/or substrate-translocation pathway of PutP (Hilger *et al.*, 2008). The number of residues conserved within the SSS family, together with the previously shown functional significance of TM IX of NIS (De la Vieja *et al.*, 2007), suggests that the domain may fulfill similar functions in the transport cycle of the other members of the family. Therefore, we set out to determine the structure of TM IX of PutP.

Such TM domains are often discontinuous helices in related transporters whose X-ray structures have been determined (Screpanti and Hunte, 2007). Therefore, we describe the backbone structure of TM domain IX in terms of a helix-loop-helix model. This model consists of two sections with an ideal α -helical secondary structure that are joined by an intervening loop. The length l of this loop and its position p in the sequence are fit parameters. Furthermore, the dihedral angles ϕ_i and ψ_i ($i = p \dots p+l-1$) for the residues in the loop section are fitted, whereas the dihedral angles in the helical sections are fixed. Thus, altogether $2l+1$

parameters have to be determined. This is achieved by acquiring experimental data on distance distributions for 16 pairs of spin labels that scan the shape of the TM domain, and by fitting the 12 best primary experimental data sets by distance distributions simulated for different helix-loop-helix models. The remaining four data sets are used to derive lower bounds for distances and to reject structures that violate these lower bounds. This procedure provides an ensemble of structural models with low root-mean-square deviation (RMSD) between simulated and experimental data. The backbone RMSD of $<2 \text{ \AA}$ of this ensemble indicates that structural models of membrane proteins can be derived from SDSL EPR data with a resolution that allows discussion of the structure-function relations.

4.3. Materials and Methods

4.3.1. Sample preparation

The *putP* alleles encoding double-Cys PutP molecules used in this study were generated by site-directed mutagenesis using plasmid pT7-5/*putP*(Δ Cys) as a template and synthetic mutagenic oligonucleotides in one- or two-step PCR reactions using *Taq*-DNA polymerase. Plasmid pT7-5/*putP*(Δ Cys) is a derivative of pT7-5 (Tabor and Richardson, 1985) and contains the *lac* promoter/operator for expression of the *putP* gene and an engineered cassette version of this gene that is devoid of all five native Cys residues. PCR fragments were digested with *Pst*I and *Spe*I or *Afl*III and ligated with similarly treated plasmid pT7-5/*putP*(Δ Cys). For overexpression, the *putP* alleles were cloned into plasmid pTrc99a (Amann *et al.*, 1988) using restriction endonucleases *Nco*I and *Hind*III. The resulting plasmids were transformed into *E. coli* WG170 [F^- *trp lacZ rpsL thi* Δ (*putPA*)₁₀₁ *proP*₂₁₉] (Stalmach *et al.*, 1983). Cells were grown, membranes were prepared, and PutP was solubilized and purified by Ni-nitrilotriacetic acid affinity chromatography as previously described (Jung *et al.*, 1998). The double-Cys PutP was labeled with (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate (MTSSL; Toronto Research Chemicals, Toronto, Canada) on the column. For this purpose, 1 mM MTSSL in buffer W (50 mM KP_i , pH 8.0, 300 mM KCl, 10 mM imidazole, 10% glycerol (v/v), 0.04% β -D-dodecylmaltoside (w/v)) was applied to the column and incubated at 4°C for 3 h. The unbound label was removed by washing the column with buffer W, and the labeled protein was eluted with 200 mM imidazole in buffer W. After elution, the protein was reconstituted under nonreducing conditions into liposomes composed of *E. coli* lipids (67% phosphatidylethanolamine, 23.2% phosphatidylglycerol, and 9.8% cardiolipin; Avanti Polar Lipids, Alabaster, AL) at a lipid/protein ratio of 20:1 (w/w) as previously described (Jung *et al.*, 1998). Finally, the proteoliposomes were washed twice with 50 mM KP_i , pH 7.5, and resuspended in the same

buffer to yield a PutP concentration of 100-250 μM . The proteoliposomes were frozen and stored in liquid nitrogen until use.

Unless the native Cys residues Cys344 and Cys349 were used for labeling, they were mutated to Ser residues. The following spin-labeled double mutants were prepared: W325R1 combined with C344R1 and C349R1, and G328R1 combined with L330R1, S332R1, A336R1, M339R1, S340R1, T341R1, L342R1, S343R1, C344R1, Q345R1, L346R1, L347R1, V348R1, and C349R1.

The effect of double-Cys replacements on transport activity was determined under standard test conditions (70 mM Na^+ , 10 μM proline) by using *E. coli* WG170 complemented with plasmid-encoded PutP variants as previously described (Hilger *et al.*, 2008; Raba *et al.*, 2008). Under these conditions, the majority of double-Cys mutants exhibited an initial rate of transport of minimum 23% of PutP(ΔCys), which in turn has 50% of the initial rate of transport and 100% of the steady-state level of proline accumulation of the wild-type. Exceptions were mutants 328/332, 328/340, 328/341, 328/345, and 328/346, with initial rates of ~14%, 9%, 1.8%, 6%, and 9.7% of PutP(ΔCys). The latter inhibitory effects on transport activity were in agreement with the previously reported importance of corresponding single-Cys mutants for Na^+ -coupled proline uptake (Hilger *et al.*, 2008; Raba *et al.*, 2008). The effect of *in situ* alkylation of single-Cys mutants was tested by Raba *et al.* (2008) using membrane-permeant *N*-ethylmaleimide (NEM). Among the Cys-mutants used in this work, C344 and L347C exhibited highly reduced activity of less than 8% after NEM labeling, whereas T341C and V348C exhibited moderately reduced activity (40% and 48%, respectively). The transport activity of all other single-Cys mutants was not significantly altered by NEM treatment.

4.3.2. DEER measurements

Four-pulse DEER measurements were performed with a Bruker Elexsys 580 spectrometer (Bruker Biospin GmbH, Karlsruhe, Germany) equipped with a 3 mm split-ring resonator under conditions of strong overcoupling ($Q \approx 100$) at a temperature of 50 K. Before insertion into the probe head, the samples were shock-frozen in liquid nitrogen to avoid crystallization of water. The four-pulse DEER sequence $(\pi/2)_{\nu_1} - \tau_1 - (\pi)_{\nu_1} - t' - (\pi)_{\nu_2} - \tau_1 + \tau_2 - t' - (\pi)_{\nu_1} - \tau_2 - \text{echo}$ was used (Pannier *et al.*, 2000). The $\pi/2$ and π pulses at the observer frequency ν_1 had equal pulse lengths of 32 ns to ensure equal excitation bandwidths, while the π pulse at the pump frequency ν_2 had a length of 12 ns to maximize modulation depth while still keeping the two excitation bands separate. The long interpulse delay was $\tau_2 = 2000$ ns, except for double mutants W325R1/C344R1 and W325R1/C349R1, where it was 2300 ns. An initial value of $t' = 80$ ns and an increment $\Delta t' =$

8 ns were used to acquire the time trace. To suppress proton modulation, data were added for eight equidistant values of τ_1 between 200 and 256 ns. A phase cycle $[(+x), -(-x)]$ was applied to the first pulse. The pump frequency ν_2 (typically 9.33 GHz) was set to the center of the resonator mode and to coincide with the global maximum of the nitroxide spectrum. The observer frequency ν_1 was set to the local maximum at the low-field edge of the spectrum ($\nu_1 - \nu_2 = 65$ MHz). Accumulation times for the data sets varied between 8 and 14 h. Data were analyzed for dipolar evolution times $t = t' - \tau_1 \geq 0$. Analysis of the data in terms of distance distributions and mean distances was performed with the program DeerAnalysis2006 (Jeschke *et al.*, 2006). The number of spins per molecules was determined based on the calibration performed previously with a series of biradicals and a triradical (Hilger *et al.*, 2005).

4.3.3. Rotamer library

The lowest-energy MTSSL rotamer R- of the previous library (Hilger *et al.*, 2007) was supplemented with peptide backbone atoms to an $\text{NH}_2\text{-C}^\alpha\text{RH-COH}$ structure, geometry-optimized with ADF 2004.01 using the BLYP density functional and a DZP basis set with effective core potentials for inner electrons (te Velde *et al.*, 2001). From this reference structure a set of 108 initial rotamer structures was generated by systematically varying side-chain dihedrals χ_1 , χ_2 and χ_4 in steps of 120° and χ_3 and χ_5 in steps of 180° , according to the expected shape of the angular potential (Tombolato *et al.*, 2006). All initial starting structures were geometry-optimized twice, first in vacuum and then in a dielectric continuum with permittivity 80.1, corresponding to water, using the COSMO solvation model and the same functional and basis set as above. As a result of strong intramolecular clashes, 10 rotamers had to be discarded. The water structures of the remaining 98 rotamers comprise the library.

4.3.4. Structure determination

An ideal α -helix backbone with dihedral angles $\phi_i = -57^\circ$ and $\psi_i = -48^\circ$ for all residues was generated as a starting structure and the appropriate side chains were added with SCWRL3.0 (Canutescu *et al.*, 2003). For a given position p and length l of the loop region, the $2l$ dihedral angles ϕ_i and ψ_i were determined in the following way: All 98 rotamers of the library were attached and their total Lennard-Jones interaction potentials with the protein were computed based on the optimized potential for liquid simulations force field (Jorgensen and Tirado-Rives, 1988). Rotamer populations were calculated assuming a Boltzmann distribution at a temperature of 175 K corresponding to solidification of a lipid membrane. The distance distribution of a spin label pair was computed by generating a histogram of N-O

midpoint distances for all possible combinations of rotamers weighted with the product of their populations. Rotamers with populations below a cutoff value of 0.001 were considered as clashing rotamers and discarded. The distance distribution was convoluted with a Gaussian function with standard deviation of 0.5 Å to account for librational fluctuation of dihedrals within their respective potential minima. DEER form factors were computed from the distance distributions and fitted together with a two-dimensionally homogeneous background function to the primary experimental data using DeerAnalysis2006 (Jeschke *et al.*, 2006) subroutines. The global RMSD between simulated and experimental data was computed by summing the individual RMSDs weighted by the inverse noise level corresponding to a maximum likelihood estimate. For the best-fit structure, side chains were repacked with SCWRL3.0, rotamer populations were recomputed, and the whole fit procedure was repeated once, giving the final structure for these values of p and l .

Such best-fit structures were computed for loop lengths $l = 4 \dots 6$ and all feasible loop starting positions p within TM domain IX. At least 10 structures with the smallest RMSD within a shallow RMSD minimum basin were provisionally accepted. Up to five additional structures were accepted if their RMSD values were only significantly larger than that of the worst accepted structure. These structures were checked against those lower limit distance constraints from DEER data sets that were not used in model fitting. Structures that violated at least one of the lower limits were rejected. The remaining structures form the final structural ensemble (see also supplementary Information). A flow chart of the algorithm is given in Fig. 4.2..

4.4. Results and Discussion

4.4.1. Test of the helix-loop-helix model for TM domains in transporters

Most membrane proteins belong to the structure type of α -helical bundles, and many TM helices are not ideal straight helices - they are often bent or kinked (Yohannan *et al.*, 2004). Any technique for elucidating the structural models of such proteins thus must be able to characterize not only the relative arrangement of helices (Bhatnagar *et al.*, 2007) but also their backbone shape. The main limitation of SDSL/EPR techniques in this context is the relatively small number of constraints that can be obtained. Structural models thus need to be coarse-grained, i.e., defined by a minimum of adjustable parameters. To derive such a model, we selected three α -helical transporters with known structures: the Na⁺/H⁺ antiporter NhaA of *E. coli* (PDB code: 1ZCD) (Hunte *et al.*, 2005), the leucine transporter LeuT_{Aa} of *Aquifex aeolicus* (PDB code: 2A65) (Yamashita *et al.*, 2005; Singh *et al.*, 2007; Zhou *et al.*, 2007), and the H⁺/Cl⁻ antiporter CIC of *E. coli* (PDB code: 1KPK) (Dutzler *et al.*, 2002). For all

TM domains in these transporters, we analyzed dihedral angles ϕ_i and ψ_i . A plot typical of a strongly kinked helix is shown in Fig. 4.1.. Here, apparently most of the dihedral angles are very close to values for an ideal α -helix, with significant deviations being confined to a continuous loop region that comprises residues 14-19.

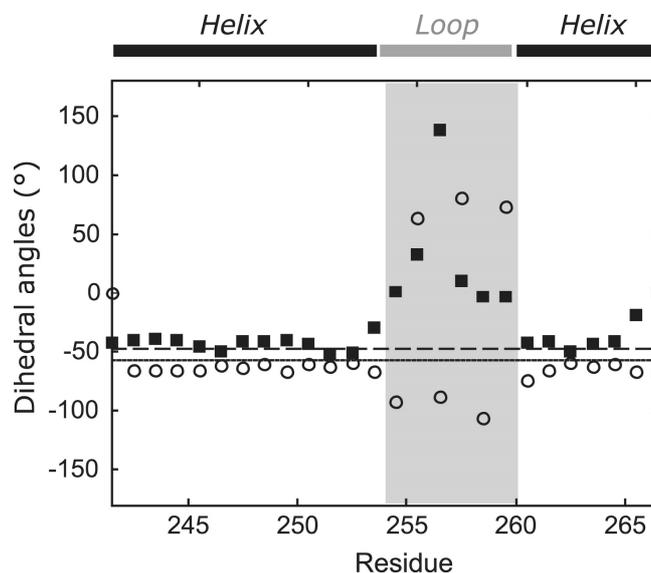


Figure 4.1.: Dihedral angles ϕ (open circles) and ψ (filled squares) for residues for the typical strongly kinked TM domain VI of LeuT_{Aa}. The solid (-57°) and dashed (-48°) lines correspond to values for an ideal α -helix for ϕ and ψ , respectively. The range assigned to a loop in a helix-loop-helix model is marked as a gray box.

This suggests a model for kinked helical TM domains that consists of two ideal α -helices joined by a loop whose position p in the domain and length l are variable. Ideal α -helix geometry is assumed for the helical parts, whereas the backbone of the loop is completely defined by $2l$ dihedral angles. We fitted this idealized model with loop lengths varying from $l = 1$ to 7 for all possible loop positions to all TM domains found in the crystal structures of NhaA, LeuT_{Aa}, and CIC. The results for the closest relative of PutP, LeuT_{Aa}, are given in Table 4.1.. Except for TM domain XI, all domains are modeled with backbone RMSDs of less than 0.5 \AA , which compares favorably to the experimental uncertainty of the X-ray data. TM domain XI is doubly kinked near its ends and can be fitted with a backbone RMSD of 0.11 \AA , assuming two loop regions (lengths of two residues each, positioned at residues 8 and 24 out of 30). Similar results are obtained for the other two transporters NhaA and CIC (see suppl. Tables 4.1. and 4.2.).

Such simplified backbone models have to be supplemented with side-group coordinates to fit them to primary DEER data because the fitting procedure requires predicted conformational distributions of the spin labels (Hilger *et al.*, 2007), which in turn depend on the conformations of neighboring side groups. Because we have no independent information on side-group conformations, these have to be predicted by an *ab initio* method, such as SCWRL3 (Canutescu *et al.*, 2003). Therefore, we tested how well experimental

coordinates of side groups are predicted by SCWRL3 for transporters. We found side-group RMSDs of 1.76 Å (NhaA), 1.52 Å (LeuT_{Aa}), and 1.27 Å (CIC). Although significant, such deviations below 2 Å are not expected to bias the predicted conformational distribution of the spin label to an extent that exceeds other uncertainties of our methodology.

Table 4.1.: Performance of a helix-loop-helix model in describing the X-ray structure (Yamashita *et al.*, 2005) of TM domains in the leucine transporter LeuT_{Aa} of *Aquifex aeolicus* (PDB identifier 2A65).

TM domain	Loop position, p	Loop length, l	Backbone RMSD (Å)
I	14	6	0.15
II	14	2	0.17
III	22	1	0.45
IV	11	2	0.05
V	7	3	0.06
VI	15	5	0.06
VII	16	7	0.09
VIII	21	4	0.12
IX	14	3	0.05
X	9	3	0.09
XI	23	1	1.50
XII	17	2	0.06

As a final test of the model, we tried to reproduce the shape of the typical moderately kinked TM domain VIII in LeuT_{Aa}, which may be functionally homologous to TM domain IX in PutP, from synthetic DEER data by the same fit procedure (see flow chart in Fig. 4.2.) that is applied to TM domain IX of PutP (*vide infra*). For this validation of the fit procedure, rotamer distributions (Jeschke and Polyhach, 2007; Hilger *et al.*, 2007) for 21 labeling positions were predicted based on experimental backbone and side-group coordinates from the X-ray structure (Yamashita *et al.*, 2005). Distance distributions for 20 label pairs were obtained from these rotamer distributions and converted to synthetic DEER form factors with subroutines of the DeerAnalysis2006 package (Jeschke *et al.*, 2006). An ideal helix backbone (all $\phi_i = -57^\circ$ and $\psi_i = -48^\circ$) was generated and side groups were added with SCWRL3 (Canutescu *et al.*, 2003). By starting from an ideal helix model, we were able to ascertain that a kink (loop segment) is introduced only for those TM domains that cannot be modeled with sufficient precision by an ideal helix. A rotamer distribution was computed for this initial model and the NO midpoint coordinates of all rotamers at all labeling positions were stored in a local frame defined by the backbone atoms N, C, and C $^\alpha$ of the labeled residue. For all possible choices of loop position p and loop length $3 \leq l \leq 7$, the dihedral angles (ϕ_i, ψ_i) in the loop region were fitted by minimizing the global RMSD with regard to all synthetic DEER form factors. A loop length of $l = 6$ was deemed satisfactory, and of all structures with this length, the one with the lowest RMSD was subjected to SCWRL3 side-group packing to adapt side-group conformations to the backbone deformation. A new

rotamer distribution was computed and all fits were repeated. The 11 structure models with the lowest RMSD were accepted.

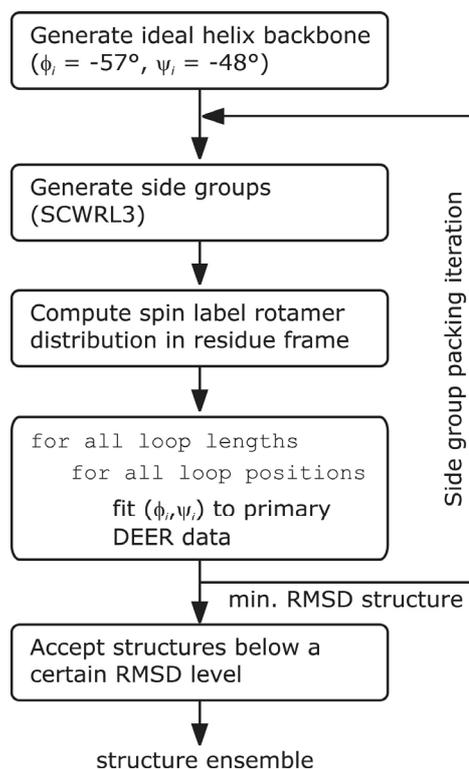


Figure 4.2.: Flow chart for generating an ensemble of helix-loop-helix models that fit DEER data by matching the label-to-label distance distributions inherent in these data.

The backbone coordinates of this ensemble are shown in Fig. 4.3.a. The backbone RMSD for the whole ensemble is 0.68 \AA , whereas the backbone RMSD with respect to the experimental structure (Fig. 4.3.b) is 1.99 \AA . The larger RMSD with regard to the experimental structure indicates a systematic deviation of the structural model. This systematic deviation could be traced back mainly to imperfections in SCWRL3 side group packing and corresponding deviations in the rotamer distribution. The dependence of the RMSD on loop length l and position p and the acceptance level are plotted in suppl. Fig. 4.1., and the agreement of simulated time-domain data and distance distributions between the original structure, an ideal helix model, and the best-fit helix-loop-helix model are shown in suppl. Fig. 4.2.. Corresponding data for the discontinuous TM domain VI of LeuT_{Aa} are given in suppl. Figs. 4.3. and 4.4..

4.4.2. TM domain IX of PutP is kinked

An additional complication arises with experimental DEER data, as these data correspond to a convolution of the form factor due to spin labels in the same protein molecule with a background function due to spin labels in neighboring protein molecules. At short distances separation of the two components is straightforward (Jeschke *et al.*, 2006). We find that the background is best fitted by a homogeneous two-dimensional spatial

distribution of protein molecules as is expected for membrane proteins in liposomes. However, these background fits are not perfect and distance distributions obtained by Tikhonov regularization contain a minor spurious peak at a distance of ~ 4.5 - 5.5 nm.

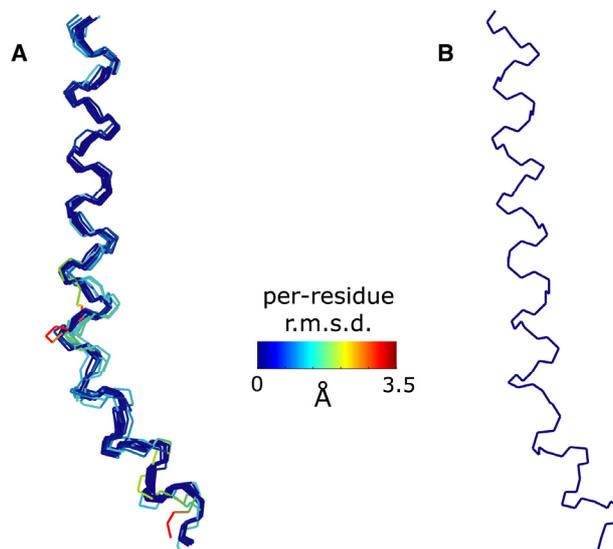


Figure 4.3.: Validation of the fit procedure on synthetic DEER data generated from the crystal structure (Yamashita *et al.*, 2005) of LeuT_{Aa} (PDB identifier 2A65) for TM domain VIII. (A) Backbone traces for the structural ensemble (11 structures) color-coded by RMSD per residue with respect to the mean structure. (B) Backbone trace from the crystal structure.

Table 4.2.: Comparison of experimental mean label-to-label distances $\langle r_{\text{exp}} \rangle$ with mean distances $\langle r_{\text{ideal}} \rangle$ in an ideal helix model and $\langle r_{\text{best}} \rangle$ in the best-fit model for TM domain IX of PutP.

Label positions	$\langle r_{\text{exp}} \rangle$ (Å)	$\langle r_{\text{ideal}} \rangle$ (Å)	$\langle r_{\text{best}} \rangle$ (Å)	Spins/molecule
328-330	17.0	17.2	17.3	1.27
328-332	16.8	16.7	16.6	1.42
328-336	18.2	18.9	18.4	1.34
328-339	22.9	19.4	22.3	1.37
328-340	25.5	22.3	22.6	1.43
328-341	28.1	23.4	26.6	1.49
328-342	22.6	22.3	22.6	1.52
328-343	27.7	23.6	27.2	1.54
328-344	32.0	27.1	32.1	1.37
328-345	30.2	27.4	28.2	1.35
328-346	26.2	27.1	27.2	1.20
328-347	35.7	29.8	33.8	1.50

The number of spins per molecules characterizes the degree of labeling.

For any given helix-loop-helix model in our structure fit and a given double mutant the form factor $F(t)$ is uniquely determined by the set of dihedral angles (ϕ_i, ψ_i) . To fit the primary data set of this double mutant we have to vary the modulation depth Δ and background density factor k as defined in (Jeschke *et al.*, 2006). At short distances this is again straightforward, but preliminary test runs showed that for the four longest distances, corresponding to double mutants 328/348, 328/349, 325/344, and 325/349, the background

fits become unstable and may result in unrealistically small or large modulation depths Δ . We thus decided to exclude these data sets from the fit procedure shown in Fig. 4.2.. Note that background fits performed separately with DeerAnalysis2006 rather than during structure fitting are stable for these mutants and give modulation depths that correspond to numbers of spins per molecule of 1.45, 1.28, 1.27, and 1.42, respectively, similar to the values found for the data sets used (Table 4.2.). Indeed, for the maximum dipolar evolution times of 2.0 μs (328/348 and 328/349) or 2.3 μs (325/344 and 325/349) achieved in the experiments model computations with simulated data sets show that the expected distances of less than 5 nm can still be determined whereas the width of the distributions becomes uncertain above 4 nm and the shape of the distributions above 3 nm.

Thus, these data sets can still provide lower limits for the mean distance between the labels. Structural models that gave a good fit of the remaining 12 data sets, but violated one or more of these lower limit constraints, were rejected. The experimental distance distributions, obtained by Tikhonov regularization using the program DeerAnalysis2006 (Jeschke *et al.*, 2006), are shown in Fig. 4.4. (*black lines*) together with distributions for the best-fit helix-loop-helix model (*red lines*) and the best possible fit by assuming an ideal helix and varying only background parameters Δ and k (*blue lines*). The corresponding fits of the primary data are given in suppl. Fig. 4.5.. For double mutants 328/330-336 fits by the ideal helix and the helix-loop-helix models nearly coincide and are both good, except for double mutant 328/336. The larger deviation for the latter mutant might be caused by deficiencies in SCWRL3 side group prediction. More likely, our rotamer library overestimates conformational freedom of the spin labeled side group for site 336 with neighboring leucine and alanine side groups, which is one of the least constrained labeling sites used in this study. For double mutant 328/339 the fit by the helix-loop-helix model is surprisingly worse than the one by the ideal helix, although this position is still in the first α -helical section. We attribute this to perturbations by side groups in the loop section, whose conformations may not be correctly predicted.

Starting with double mutant 328/340, but except for double mutants 328/342, and 328/345, fits by the helix-loop-helix model are significantly better than fits by an ideal helix. In only one further case, for double mutant 328/347 the helix-loop-helix model exhibits a significant deviation from experimental data, corresponding to a slightly longer distance in the experiment than in the structural model. This may again be due to deficiencies in SCWRL3 side group prediction and modeling of the spin label by the rotamer library.

Experimental mean distances are compared to mean distances for an ideal helix model and for the best-fit helix-loop-helix model in Table 4.2.. Again with the exception of double mutant 328/339, agreement is better for the helix-loop-helix model that has a total RMSD for the mean distances of 1.00 Å compared to 3.28 Å for the ideal helix model.

Of the 12 initially accepted structures (Fig. 4.5.) four had to be rejected as they violated lower distance limits derived from the four data sets not used in fitting. The ensemble of the remaining eight structures corresponds to a well defined, moderately kinked backbone of this TM domain with a RMSD of 1.9 Å (Fig. 4.6.a). Not surprisingly, variability of the coordinates is larger in the loop region than in the helix regions.

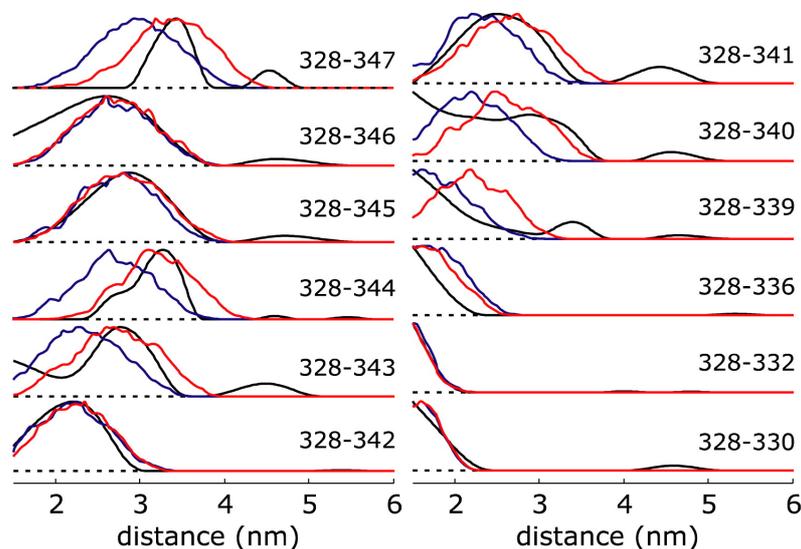


Figure 4.4.: Experimental label-to-label distance distributions for TM domain IX of PutP (black traces), fits for the best helix-loop-helix model (red lines), and fits assuming an ideal helix (blue lines).

We have tested stability of the fit by a similar approach as used in (Hilger *et al.*, 2007). In a first series of test fits, we doubled noise in the input data by adding pseudorandom numbers. The resulting ensemble of eight structures (Fig. 4.6.b) has a backbone RMSD of 2.7 Å, yet the general shape of the TM domain and the location of the loop region are preserved. This double-noise test demonstrates that the measurement time per sample cannot significantly be shortened without compromising the quality of the structural model.

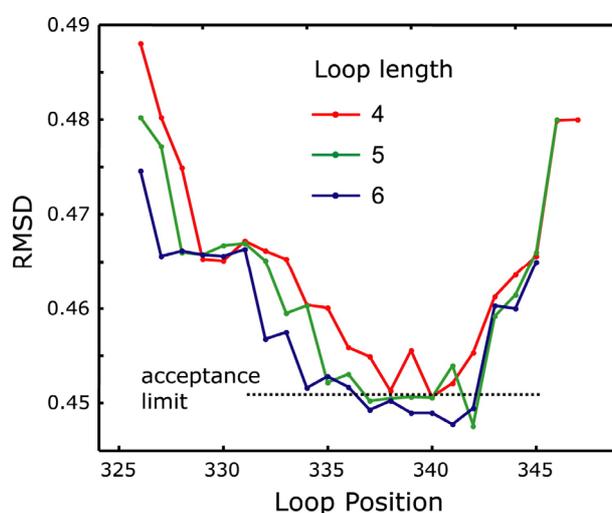


Figure 4.5.: Dependence of the RMSD between simulated and experimental DEER traces for best-fit helix-loop-helix models on loop length and position for TM domain IX of PutP. The acceptance limit for including structures in the final ensemble is shown as a dotted line.

In a second series of test fits we restricted ourselves to loop length 5 and performed 12 structure determinations, each one excluding one of the originally used 12 primary DEER data sets. To keep the size of the total ensemble manageable, the number of accepted structures for each fit was reduced by a factor of two. This provides altogether 65 acceptable structures, of which 37 are consistent with the lower limit constraints from the remaining four data sets. To condense this ensemble further, we performed a cluster analysis and selected those 26 structures that have the smallest RMSD (Fig. 4.6.c). This ensemble has a backbone RMSD of 1.6 Å. Again, the general shape and location of the loop region are preserved. However, the significant increase in backbone RMSD for some of the accepted models indicates that at least 10 to 11 constraints are needed to obtain a structural model with sufficient precision.

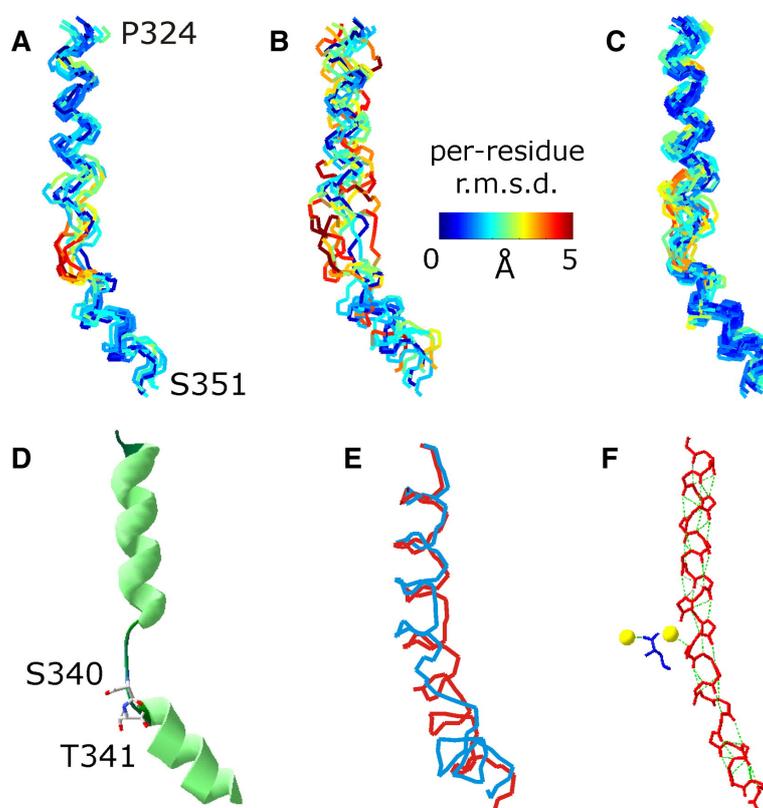


Figure 4.6.: Helix-loop-helix models for the backbone of TM domain IX in PutP. (A) Ensemble obtained by using the 12 DEER data sets with lower limit constraints from four additional data sets (eight structures). (B) Ensemble obtained by using the same data sets as in A with pseudorandom numbers added to double the noise level (eight structures). (C) Ensemble obtained from all combinations of only 11 out of the 12 DEER data sets with lower limit constraints from four additional data sets (26 structures). (D) Cartoon of the representative model of the ensemble with 26 structures with ligand-binding residues S340 and T341 shown as stick. (E) Comparison of TM domain IX in PutP (red) with TM domain VIII in a crystal structure (Singh *et al.*, 2007) of LeuT_{Aa} (blue, PDB identifier 2Q72). (F) Backbone model including carbonyl oxygens from the crystal structure (Singh *et al.*, 2007) of TM domain VIII in LeuT_{Aa} (red) with bound sodium ions (yellow) and substrate leucine (blue). Hydrogen bonds are visualized as dotted green lines (PDB identifier 2Q72).

Because we employed spin labels with a size similar to the one of the bulkiest natural amino acid side groups, the question arises as to whether the spin labels might have introduced the observed kink. For the spin labels at our reference points 325 and 328, this possibility can be safely excluded, since these points are remote from the kink. Because the

sites near the kink have been found to be highly accessible to water (Raba *et al.*, 2008) and labeling efficiencies are high, the labels at these sites are not expected to alter the structure. In most cases studied so far by SDSL techniques, spatial restrictions at a labeling site have led to poor labeling efficiency rather than to alterations in the structure. Furthermore, the kink is strongly sensed at at least three labeling sites: 343, 344, and 347. It appears rather unlikely that the labels at all three sites induce a consistent structural distortion.

Note also that the helix-loop-helix model used in structural modeling amounts to an assumption regarding the secondary structure. Alternative secondary structures may also be consistent with our data. Between residues 328 and 339 the observed incremental changes in distances and their periodicity strongly suggest an α -helical structure, whereas between residues 340 and 345 the data are not consistent with either an α -helix or a β -sheet. The prevalence of an α -helical structure between residues 324 and 327 at the one end and between residues 346 and 351 at the other end cannot be rigorously proved based on the experimental data alone. The assumption is supported by secondary structure prediction with PredictProtein (Rost *et al.*, 2004), which is displayed in suppl. Fig. 4.6.b.

In principle, such helix scans and structural modeling could be applied to each TM domain in an α -helical protein. This would provide coarse-grained backbone structures for these TM domains. An entire helical bundle with N TM domains could then be constructed by treating the predetermined backbone structures of the domains as rigid bodies and elucidating their relative translation and orientation, as demonstrated previously (Smirnova *et al.*, 2007; Hilger *et al.*, 2007). Because for each pair there are six degrees of freedom and there are $(N-1)$ independent pairs, this would require fits of another $6(N-1)$ parameters, which may require the preparation of and distance measurements on $\sim 10(N-1)$ further double mutants. Since some of the prepared mutants may not be suitable for labeling, the total number of required mutants may well exceed 300 for a protein of the size of PutP. Although a considerable effort is required, such an approach to a coarse-grained structural model is feasible with currently available methodology.

Recently, a different approach for deriving structural models from EPR data was suggested (Alexander *et al.*, 2008) that required fewer constraints by making extensive use of structure prediction by the program Rosetta (Simons *et al.*, 1997). This approach was shown to provide full atom models with 1.0 and 2.6 Å RMSD for the soluble proteins T4-lysozyme (107 residue helical domain) and α A-crystallin (88 residue β sandwich), respectively. For small and medium-sized soluble proteins it may be better suited than the approach suggested here. For membrane proteins, *de novo* structure predictions are currently less reliable. Furthermore, with current algorithms and computers, they are restricted to proteins with up to 200 residues, well below the size of the transporters.

Nevertheless, judicious use of constraints from such predictions or from molecular dynamics simulations may help to decrease the number of required experimental constraints.

4.4.3. Functional implications of the structural model

TM domain IX appears to line the translocation pathway, with residues S340 and T341 taking part in ligand binding and transport (Hilger *et al.*, 2008). Residues involved in ligand binding are frequently found at the apex of a hydrophilic cavity in the transporter, as described for lactose permease, glycerol-3-phosphate transporter (both members of the major facilitator family), and neurotransmitter transporter relatives (Abramson *et al.*, 2003; Huang *et al.*, 2003; Quick *et al.*, 2006; Zhang and Rudnick, 2006). In fact, a Cys accessibility analysis of TM domain IX of PutP suggests that the domain participates in the formation of an inwardly oriented cavity. Proline binding in the presence of Na⁺ blocks this accessibility (Raba *et al.*, 2008).

In all of our structural models, residue S340 is found in the loop region; in most models this also applies to residue T341. These residues are visualized as stick models in a representative structure from the ensemble of 26 structures (Fig. 4.6.d). This structure has the lowest RMSD with respect to all the other structures. Note that the orientation of the OH groups relative to the kink differs within the ensemble.

The structure shows that the inwardly oriented, ligand-sensitive cavity suggested by Raba *et al.* (2008) may be due to a kink in the helix that starts approximately at residue S340. However, the loop already starts around residue V338. Between residues V338 and S340 the domain continues in the direction of the first helical part. However, the backbone is twisted so that the side groups of these residues point to different directions than in an ideal helix. These residues may belong to the hinge that alters its conformation when residues S340 and T341 bind substrate. Possibly, a reorganization of this loop facilitates accessibility of the proline-binding site from the outside.

Of interest, a very similar kink is observed in the putatively homologous TM domain VIII of LeuT_{Aa} (Yamashita *et al.*, 2005; Sing *et al.*, 2007; Zhou *et al.*, 2007), although all X-ray structures correspond to substrate-bound states. If the two domains are aligned at T354/S355 of LeuT_{Aa} and S340/T341 of PutP, and are restricted to the same length (Fig. 4.6.e), our model (*red*) superimposes on the X-ray structure of TM domain VIII of LeuT_{Aa} (*blue*) with a backbone RMSD of 1.98 Å. The bend of the cytoplasmic end with respect to the periplasmic end in LeuT_{Aa} appears to open access to the substrate-binding site. Some of the residues in this bend do not exhibit the hydrogen bond pattern typical of α -helical secondary structure (Fig. 4.6.f). In contrast, the homologous TM domain IX in vSGLT in the substrate-bound state (Faham *et al.*, 2008) can be well described by the model of an ideal α -helix (E.

M. Wright, University of California, Los Angeles, personal communication, 2008). If TM domain IX in PutP on substrate binding indeed performs a hinge motion that closes an inwardly oriented cavity, it would also be expected to be approximately straight in the substrate-bound state.

4.5. Conclusions

Most kinked helices in membrane proteins can be described with a backbone RMSD of less than 0.5 Å by helix-loop-helix models with loop lengths l of seven residues or less. Such models have free $2l+1$ parameters, namely, two dihedral angles per loop residue and the loop position. These parameters can be determined by means of pulsed EPR distance distribution measurements on doubly spin-labeled mutants. For TM domain IX of the Na⁺/proline transporter PutP of *E. coli*, we obtain structural ensembles whose backbone coordinate variation is characterized by an internal RMSD of less than 2 Å. Taking into account systematic errors of about the same magnitude due to imperfect prediction of conformations for unlabeled side groups, and the conformational distribution of spin-labeled side groups, the backbone RMSD of the representative structure of the ensemble from the mean native structure is expected to be less than 3 Å. At this resolution, kinks in helices can clearly be discerned and their functional relevance can be discussed. The kink in TM domain IX of PutP appears to be a hinge that closes the inwardly oriented cavity on substrate binding and opens it for substrate release to the interior of the cell.

The newly introduced methodology for determining the shape of a single TM domain in a membrane protein could be combined with a previously introduced methodology for establishing the relative position and orientation of two rigid structures (Bhatnagar *et al.*, 2007; Hilger *et al.*, 2007) to provide structural models of the whole TM domain bundle. For a protein of the size of PutP, determining 13 TM domain shapes and 12 relative positions and orientations of TM domain pairs would require measurements on ~300 double mutants. It has to be taken into account that some of the intended labeling sites may not be accessible due to tight packing, although the flexibility and size of MTSSL spin labels allows for a surprisingly high success rate in labeling. To enhance the feasibility of such an approach for deriving coarse-grained structural models of whole membrane proteins, the number of required mutants should thus be decreased by systematically taking into account internal restraints and accessibility information. Work along these lines is now in progress.

4.6. Acknowledgments

We thank Christian Bauer for technical support, and Hans Wolfgang Spiess for providing access to a pulsed EPR spectrometer. The financial support of Deutsche Forschungsgemeinschaft (Exc114-1, Je246/3-2, Ju333/3-2, and Ju333/4-2) is gratefully acknowledged.

4.7. References

- Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R., and Iwata, S.** (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* 301: 610-615.
- Alexander, N., Bortolus, M., Al-Mestarihi, A., Mchaourab, H., and Meiler, J.** (2008) *De novo* high-resolution protein structure determination from sparse spin-labeling EPR data. *Structure* 16: 181-195.
- Amann, E., Ochs, B., and Abel, K. J.** (1988) Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* 69: 301-315.
- Bhatnagar, J., Freed, J. H., and Crane, B. R.** (2007) Rigid body refinement of protein complexes with long-range distance restraints from pulsed dipolar ESR. *Methods Enzymol.* 423: 117-133.
- Borbat, P. P., Surendhran, K., Bortolus, M., Zou, P., Freed, J. H., and Mchaourab, H. S.** (2007) Conformational motion of the ABC transporter MsbA induced by ATP hydrolysis. *PLOS Biology* 5: 2211-2219.
- Bowie, J. U.** (2005) Solving the membrane protein folding problem. *Nature* 438: 581-589.
- Canutescu, A. A., Shelenkov, A. A., and Dunbrack, R. L. Jr.** (2003) A graph theory algorithm for protein side-chain prediction. *Protein Science* 12: 2001-2014.
- De la Vieja, A., Reed, M. D., Ginter, C. S., and Carrasco, N.** (2007) Amino acid residues in transmembrane segment IX of the Na⁺/I⁻ symporter play a role in its Na⁺ dependence and are critical for transport activity. *J. Biol. Chem.* 282: 25290-25298.
- Dohán, O., De, I., V., and Carrasco, N.** (2006) Hydrocortisone and purinergic signaling stimulate sodium/iodide symporter (NIS)-mediated iodide transport in breast cancer cells. *Mol. Endocrinol.* 20: 1121-1137.
- Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T., and MacKinnon, R.** (2002) X-ray structure of a Cl⁻ channel at 3.0 angstrom reveals the molecular basis of anion selectivity. *Nature* 415: 287-294.
- Faham, S., Watanabe, A., Besserer, G. M., Cascio, D., Specht, A., Hirayama, B. A., Wright, E. M., and Abramson, J.** (2008) The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na⁺/sugar symport. *Science* 321: 810-814.

- Hilger, D., Jung, H., Padan, C., Wegener, K. P., Steinhoff, H.-J., and Jeschke, G.** (2005) Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na⁺/H⁺ antiporter of *E. coli*. *Biophys. J.* 89: 1328-1338.
- Hilger, D., Polyhach, Y., Padan, E., Jung, H., and Jeschke, G.** (2007) High-resolution structure of a Na⁺/H⁺ antiporter dimer obtained by pulsed EPR distance measurements. *Biophys. J.* 93: 3675-3683.
- Hilger, D., Böhm, M., Hackmann, A., and Jung, H.** (2008) Role of Ser340 and Thr341 in transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli* in ligand binding and transport. *J. Biol. Chem.* 283: 4921-4929.
- Huang, Y., Lemieux, M. J., Song, J., Auer, M., and Wang, D. N.** (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* 301: 616-620.
- Hubbell, W. L., Cafiso, D. S., and Altenbach, C.** (2000) Identifying conformational changes with site-directed spin labeling. *Nat. Struct. Biol.* 7: 735-739.
- Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E., and Michel, H.** (2005) Structure of a Na⁺/H⁺ antiporter and insights into mechanism of action and regulation by pH. *Nature* 435: 1197-1202.
- Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H.** (2006) DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Reson.* 30: 473-498.
- Jeschke, G., and Polyhach, Y.** (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* 9: 1895-1910.
- Jorgensen, W. L., and Tirado-Rives, J.** (1988) The OPLS potential functions for proteins - Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110: 1657-1666.
- Jung, H., Tebbe, S., Schmid, R., and Jung, K.** (1998) Unidirectional reconstitution and characterization of purified Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 37: 11083-11088.
- Jung, H.** (2002) The sodium/substrate symporter family: structural and functional features. *FEBS Lett.* 529: 73-77.
- Lacapere, J. J., Pebay-Peyroula, E., Neumann, J. M., and Etchebest, C.** (2007) Determining membrane protein structures: still a challenge! *Trends Biochem. Sci.* 32: 259-270.
- Liu, Y. S., Sompornpisut, P., and Perozo, E.** (2001) Structure of the KcsA channel intracellular gate in the open state. *Nat. Struct. Biol.* 8: 883-887.
- Opella, S. J., and Marassi, F. M.** (2004) Structure determination of membrane proteins by NMR spectroscopy. *Chem. Rev.* 104: 3587-3606.

- Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W.** (2000) Dead-time free measurement of dipole-dipole interactions between electron spins. *J. Magn. Reson.* 142: 331-340.
- Quick M., Yano H., Goldberg N. R., Duan L., Beuming T., Shi L., Weinstein H., and Javitch, J. A.** (2006) State-dependent conformations of the translocation pathway in the tyrosine transporter Tyt1, a novel neurotransmitter:sodium symporter from *Fusobacterium nucleatum*. *J. Biol. Chem.* 281: 26444-26454.
- Raba, M., Baumgartner, T., Hilger, D., Klempahn, K., Härtel, T., Jung, K., and Jung, H.** (2008) Function of transmembrane domain IX in the Na⁺/proline transporter PutP. *J. Mol. Biol.* 382: 884-893.
- Rost, B., Yachdav, G., and Liu, J.** (2004) The PredictProtein server. *Nucl. Acids Res.* 32: W321-W326.
- Sale, K., Song, L., Liu, Y. S., Peroro, E., and Fajer, P.** (2005). Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *J. Am. Chem. Soc.* 127: 9334-9335.
- Schiemann, O., and Prisner, T. F.** (2007) Long-range distance determinations in biomacromolecules by EPR spectroscopy. *Quater. Rev. Biophys.* 40: 1-53.
- Schwan, W. R., Lehmann, L., and McCormick, J.** (2006) Transcriptional activation of the *Staphylococcus aureus putP* gene by low-proline-high osmotic conditions and during infection of murine and human tissues. *Infect. Immun.* 74: 399-409.
- Screpanti, E., and Hunte, C.** (2007) Discontinuous membrane helices in transport proteins and their correlation with function. *J. Struct. Biol.* 159: 261-267.
- Simons, K. T., Kopperberg, C., Huang, E., and Baker, D.** (1997) Assembly of protein tertiary structures from fragments with similar local sequence using simulated annealing and bayesian scoring functions. *J. Mol. Biol.* 268: 209-225.
- Singh, S. K., Yamashita, A., and Gouaux, E.** (2007) Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. *Nature* 448: 952-956.
- Smirnova, I., Kasho, V., Choe, J. Y., Altenbach, C., Hubbell, W. L., and Kaback, H. R.** (2007) Sugar binding induces an outward facing conformation of LacY. *Proc. Natl. Acad. Sci. U. S. A.* 104: 16504-16509.
- Stalmach, M. E., Grothe, S., and Wood, J. M.** (1983) Two proline porters in *Escherichia coli* K-12. *J. Bacteriol.* 156: 481-486.
- Tabor, S., and Richardson, C. C.** (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U. S. A.* 82: 1074-1078.
- Tamm, L. K., and Liang, B.** (2006) NMR of membrane proteins in solution. *Progr. Nucl. Magn. Reson.* 48: 201-210.
- te Velde, G., Bickelhaupt, F. M., van Gisbergen, S. J. A., Fonseca Guerra, C., Baerends, E. J., Snijders, J. G., and Ziegler, T.** (2001) Chemistry with ADF. *J. Comput. Chem.* 22: 931-967.

- Tombolato, F., Ferrarini, A., and Freed, J. H.** (2006) Dynamics of the nitroxide side chain in spin-labeled proteins. *J. Phys. Chem. B* 110: 26248-26259.
- Wright, E. M., Hirayama, B. A., and Loo, D. F.** (2007) Active sugar transport in health and disease. *J. Intern. Med.* 261: 32-43.
- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E.** (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* 437: 215-223.
- Yohannan, S., Faham, S., Yang, D., Whitelegge, J. P., and Bowie, J. U.** (2004) The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* 101: 959-963.
- Zhang, Y. W., and Rudnick, G.** (2006) The cytoplasmic substrate permeation pathway of serotonin transporter. *J. Biol. Chem.* 281: 36213-36220.
- Zhou, Z., Zhen, J., Karpowich, N. K., Goetz, R. M., Law, C. J., Reith, M. E., and Wang, D. N.** (2007) LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. *Science* 317: 1390-1393.

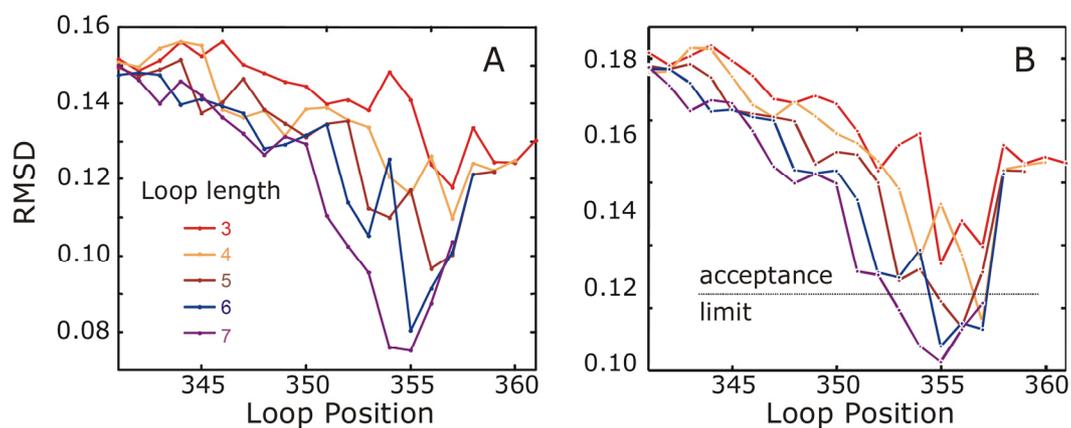
4.8. Supplementary Material

Supplementary Table 4.1.: Performance of a helix-loop-helix model in describing the X-ray structure (Hunte *et al.*, 2005) of TM domains in the Na⁺/H⁺ antiporter NhaA of *E. coli* (PDB identifier 1ZCD).

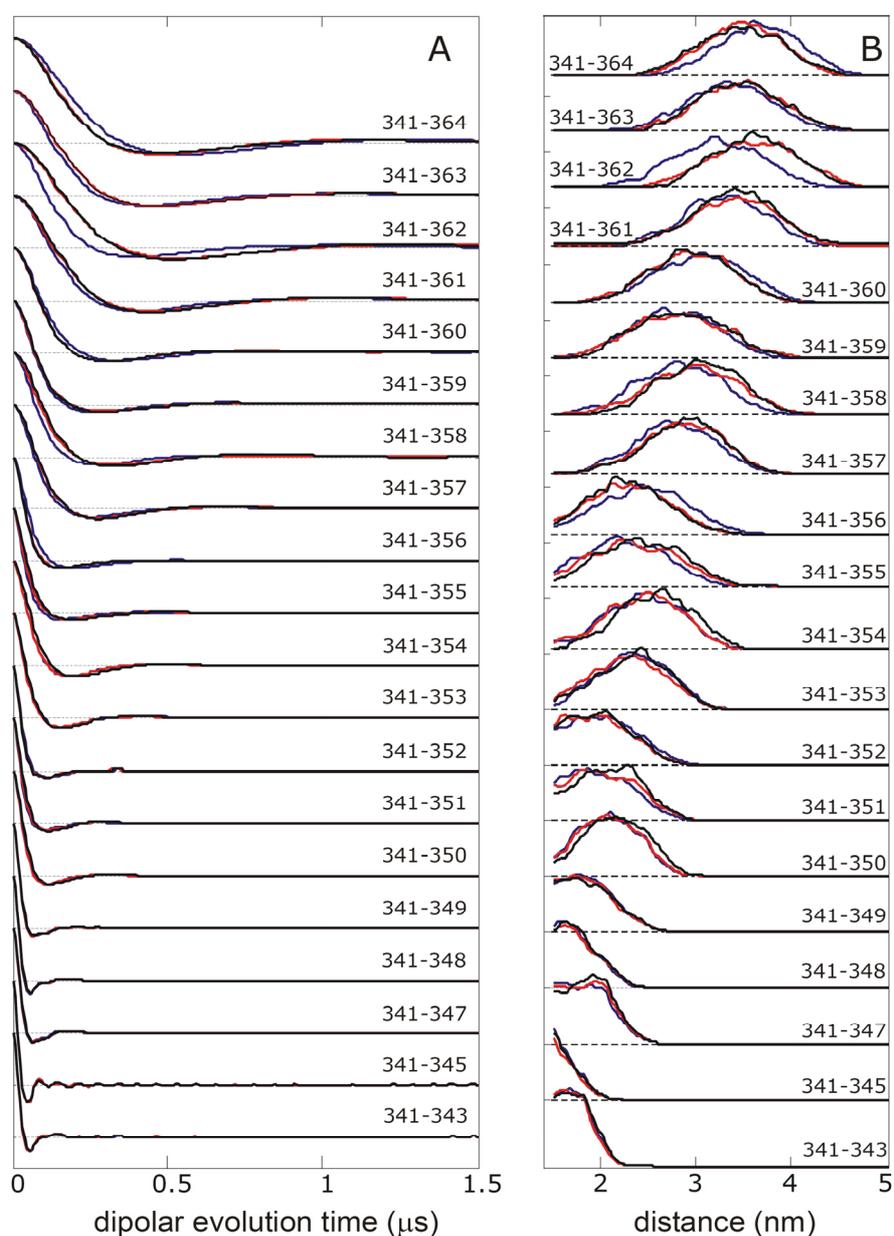
TM domain	Loop position p	Loop length l	Backbone RMSD (Å)
I	7	3	0.06
II	11	3	0.09
III	11	2	0.13
IV	8	6	0.26
V	15	2	0.14
VI	11	2	0.04
VII	3	5	0.02
VIII	6	3	0.03
IX	10	6	0.14
X	13	2	0.67
XI	4	10	0.21
XII	17	4	0.11

Supplementary Table 4.2.: Performance of a helix-loop-helix model in describing the X-ray structure (Duzler *et al.*, 2002) of TM domains in the H⁺/Cl⁻ antiporter CIC of *E. coli* (PDB identifier 1KPK).

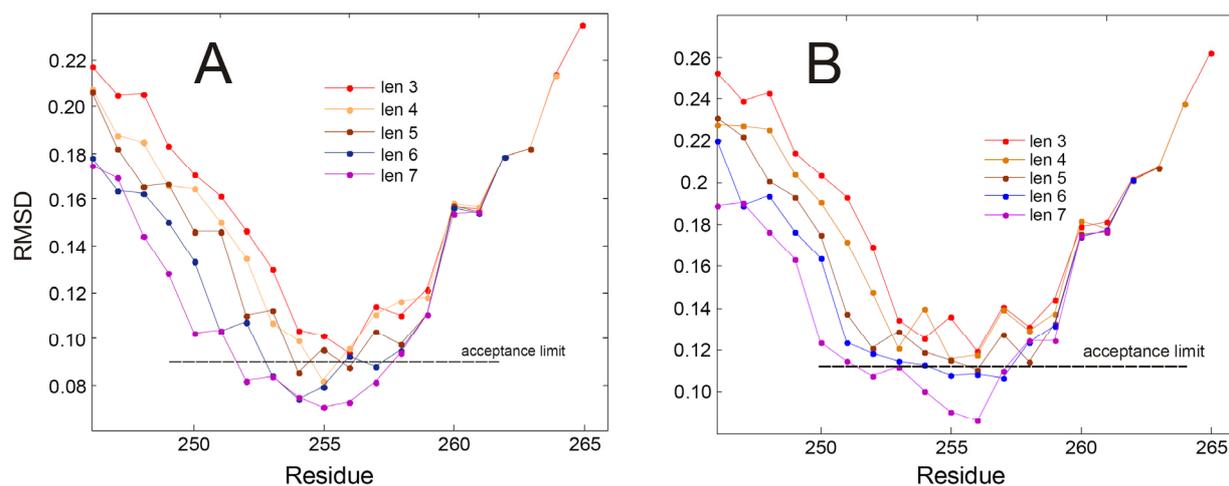
TM domain	Loop position p	Loop length l	Backbone RMSD (Å)
I	33	1	0.19
II	12	2	0.12
III	20	5	0.64
IV	4	3	0.06
V	7	3	0.06
VI	9	1	0.13
VII	13	2	0.31
VIII	13	2	0.05
IX	23	2	0.48
X	9	5	0.12
XI	13	1	0.23
XII	9	2	0.08
XIII	2	1	0.11



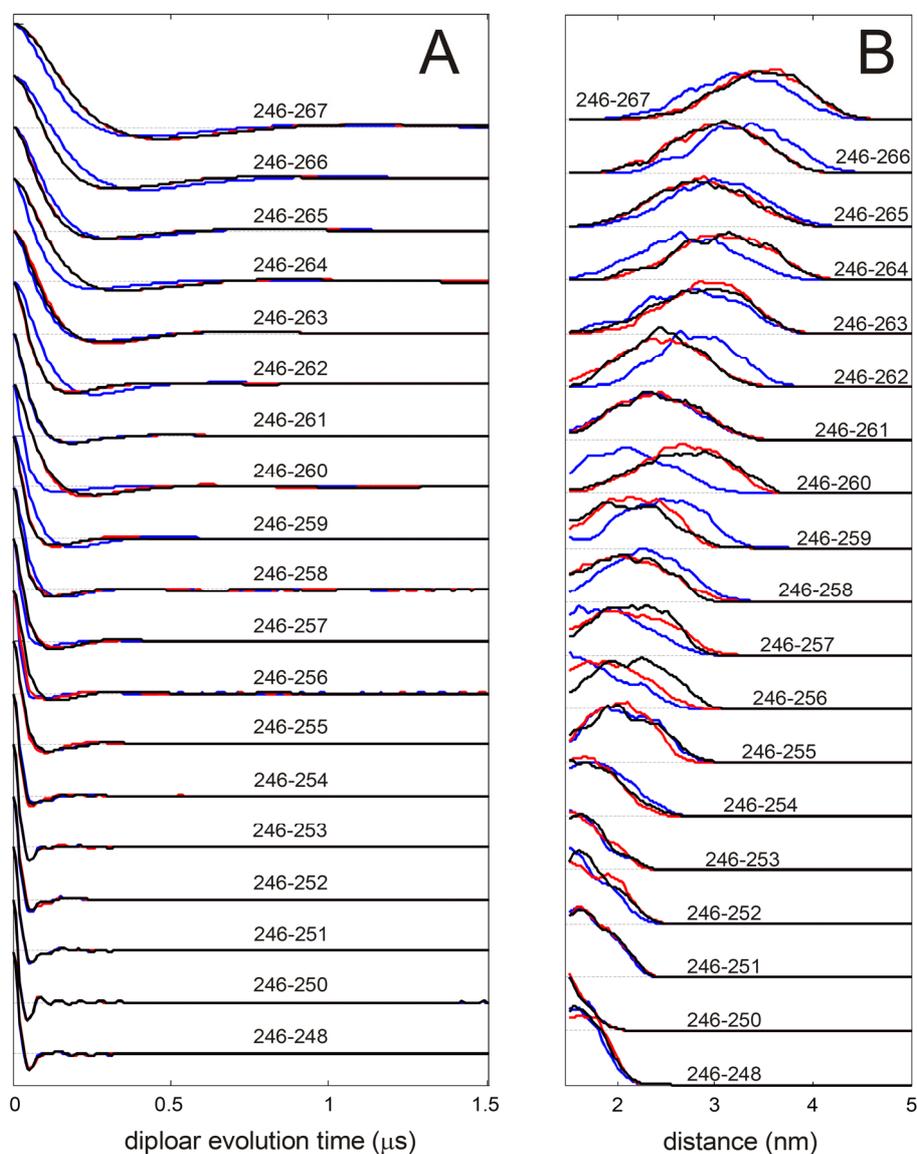
Supplementary Figure 4.1.: Dependence of the RMSD between DEER traces computed for the X-ray structure and DEER traces computed for best-fit helix-loop-helix models on loop length and position for TM domain VIII of LeuT_{Aa}. (A) Initial fits with SCWRL3 modeling of side groups for an ideal helix. (B) Refined fit with SCWRL3 modeling of side groups for the representative structure of the initial fit. The acceptance limit for including structures in the final ensemble is shown as a dotted line.



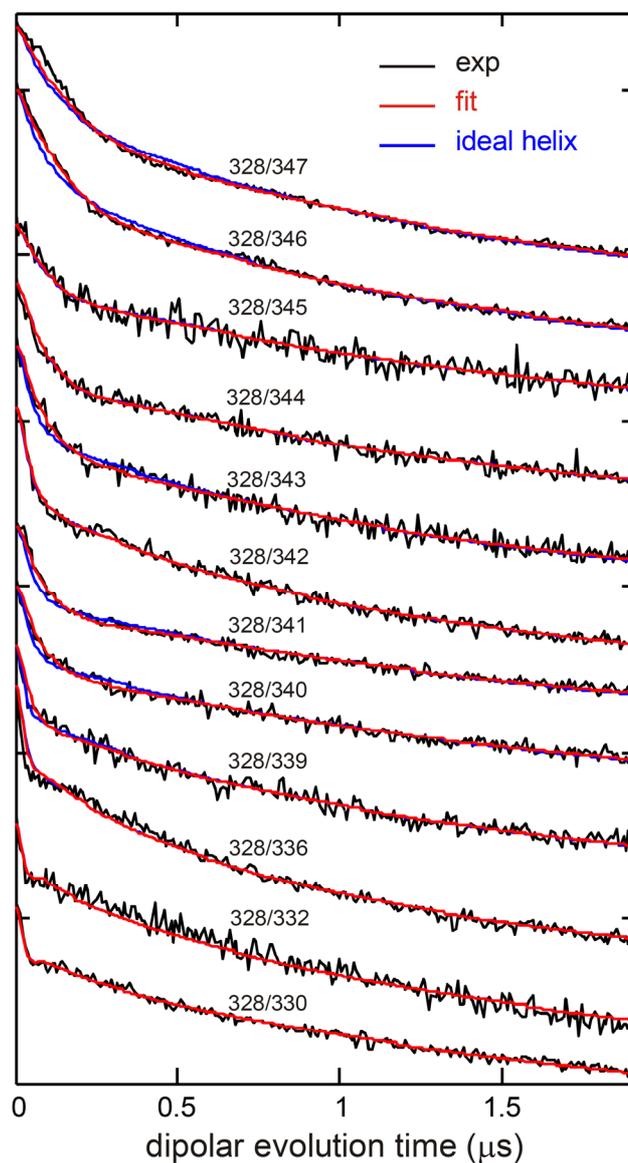
Supplementary Figure 4.2.: Comparison of primary DEER data (A) and distance distributions (B) for different structural models of TM domain VIII of LeuT_{Aa}. Black traces correspond to the X-ray structure (Yamashita *et al.*, 2005), red traces to best fits for the representative helix-loop-helix model, blue lines to best fits assuming an ideal helix.



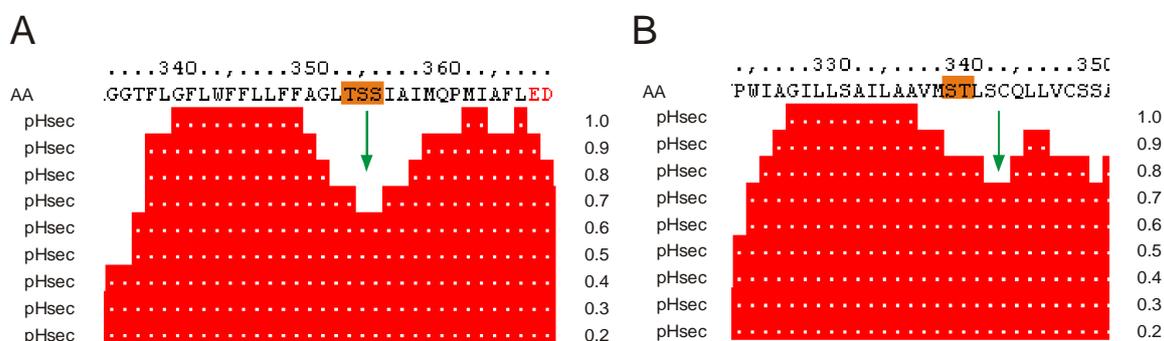
Supplementary Figure 4.3: Dependence of the RMSD between DEER traces computed for the X-ray structure and DEER traces computed for best-fit helix-loop-helix models on loop length and position for TM domain VI of LeuT_{Aa}. (A) Initial fits with SCWRL3 modeling of side groups for an ideal helix. (B) Refined fit with SCWRL3 modeling of side groups for the representative structure of the initial fit. The acceptance limit for including structures in the final ensemble is shown as a dotted line.



Supplementary Figure 4.4: Comparison of primary DEER data (A) and distance distributions (B) for different structural models of TM domain VI of LeuT_{Aa}. Black traces correspond to the X-ray structure (Yamashita *et al.*, 2005), red traces to best fits for the representative helix-loop-helix model, blue lines to best assuming an ideal helix.



Supplementary Figure 4.5.: Set of primary DEER data traces used in the fit procedure for TM domain IX of PutP (black traces), best fits for a typical accepted helix-loop-helix model (red lines), and best fits assuming an ideal helix (blue lines).



Supplementary Figure 4.6.: Secondary structure predictions by PredictProtein (Rost *et al.*, 2004) for TM domain VIII of LeuT_{Aa} (A) and TMD domain IX of PutP (B). The orange boxes mark the observed or putative substrate binding sites and the green arrows the positions in the sequence, where PredictProtein suggests a kink.

4.9. References

- Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T., and MacKinnon, R.** (2002) X-ray structure of a CIC chloride channel at 3.0 angstrom reveals the molecular basis of anion selectivity. *Nature* 415: 287-294.
- Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E., and Michel, H.** (2005) Structure of a Na⁺/H⁺ antiporter and insights into mechanism of action and regulation by pH. *Nature* 435: 1197-1202.
- Rost, B., Yachdav, G., and Liu, J.** (2004). The PredictProtein server. *Nuc. Acids Res.* 32: W321-W326.
- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E.** (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* 437: 215-223.

CHAPTER 5***Modeling of the helix bundle of the Na⁺/proline transporter PutP of Escherichia coli based on EPR distance restraints and structural homology*****5.1. Abstract**

Sodium-driven secondary transporters are particularly important for basic cellular processes as they catalyze the uptake of a diversity of substrates like amino acids, sugars, neurotransmitters, and vitamins. To understand the principles of Na⁺/substrate transport, knowledge about the tertiary structure of the transporters is indispensable. However, despite the increasing number of structures solved during the past few years, crystallizing membrane proteins remains a challenging endeavor and progress is rather slow. Although the structure of the related Na⁺/galactose symporter vSGLT of *Vibrio parahaemolyticus* has recently been solved, attempts to crystallize the Na⁺/proline symporter PutP of *Escherichia coli* have failed so far. Surprisingly, the fold of the core of ten transmembrane helices of the galactose symporter vSGLT from the sodium/substrate symporter (SSS) family was also observed in the previously reported structure of LeuT_{Aa}, a Na⁺/leucine transporter of the neurotransmitter/sodium symporter (NSS) family. Here, we modeled the helix bundle of PutP based on 45 double electron-electron resonance (DEER) distance measurements between spin labels attached to helix ends and template restraints derived from the ten-helix core of the vSGLT crystal structure. Two ensembles of structural models with a root mean-square deviation of about 12 Å are obtained. These models show core structures with similar fold to that of the vSGLT template. Furthermore, arrangement of the three non-core helices of PutP with respect to the core is established. Finally, comparison of the core structure of PutP with that of vSGLT and LeuT_{Aa} leads to a model of putative ligand-binding sites for PutP, which can guide functional studies to identify further residues essential for transporter function.

5.2. Introduction

Knowledge of the three-dimensional structure of membrane proteins is a basic prerequisite for understanding their molecular mechanism of function. However, structural information about them is sparse even though they are encoded by 20-30% of all open reading frames of eubacterial, archaeal and eukaryotic genomes (Stevens and Arkin, 2000; Krogh *et al.*, 2001). This slow progress in structure determination of membrane proteins is due to difficulties in obtaining well-ordered three-dimensional crystals for high-resolution X-ray structural analysis. Furthermore, NMR techniques are currently applicable only to small membrane proteins or β -barrel proteins, whereas the architecture of many transporters is

based on a bundle of more than a few α -helices (Saier, 2003). The modeling of such bundles on the basis of experimental information is thus of considerable interest.

Valuable experimental information for protein modeling strategies are intramolecular distances that can be obtained by e.g. fluorescence resonance energy transfer (FRET) measurements or site-directed spin labeling (SDSL) electron paramagnetic resonance (EPR) spectroscopy (Majumdar *et al.*, 2007; Liu *et al.*, 2001; Smirnova *et al.*, 2007; Sale *et al.*, 2004). These techniques are applicable to non-crystalline α -helical membrane proteins with no significant limitation to the size of the molecule. However, a drawback of these techniques is the relatively small number of distance restraints that can be obtained. Such scarcity of restraints prevents modeling of the structure with atomistic resolution. Nevertheless, knowledge of the relative arrangement of the helices in the bundle can potentially be obtained and this information can guide further functional studies. In the present work we explore the use of probe techniques, in particular SDSL EPR spectroscopy, for modeling a helix bundle. Compared to FRET, SDSL EPR has the advantage that spin labels are of relatively small size, since they resemble the dimensions of large amino acid side chains. Furthermore, the labeling strategy is simple and provides the measurement of distances between two spin labels with high precision. By applying SDSL in combination with pulse EPR, distance distributions between about 1.8 and 5 nm can be obtained that match the range of distances between helix ends in bundles of α -helices (Jeschke and Polyhach, 2007).

As a model system we use the 13 transmembrane Na^+ /proline transporter PutP of *E. coli* (Jung *et al.*, 1998a; Wegener *et al.*, 2000). PutP is a member of the sodium/solute symporter (SSS) family (TC 2.A.21, SLC5) that is an evolutionary related collection of currently around 1,000 proteins found in all kingdoms of life (Jung, 2002; Wright and Turk, 2004). Transporters of this family utilize a sodium motive force to drive uphill transport of substrates like sugars, amino acids, vitamins, ions, *myo*-inositol, phenyl acetate, and urea (Reizer *et al.*, 1994; Jung, 2001). Dysfunctions of two SSS family members, the Na^+ /glucose transporter (SGLT1) and Na^+ /iodide symporter (NIS), are implicated in the human diseases glucose-galactose malabsorption, and iodide transport defect (Wright *et al.*, 2007; Reed-Tsur *et al.*, 2008). Both proteins also play important roles in medical therapy (Wright *et al.*, 2007; Dohán *et al.*, 2006). Furthermore, bacterial transporters such as PutP of *Helicobacter pylori* and *Staphylococcus aureus* are essential for bacterial virulence (Kavermann *et al.*, 2003; Schwan *et al.*, 2006).

Very recently, the first crystal structure of one protein of this family, the Na^+ /galactose symporter vSGLT of *V. parahaemolyticus*, was obtained (Faham *et al.*, 2008). The structure contains 14 transmembrane helices (TMs) with a core of two inverted repeat domains of five TMs each. Surprisingly, the core of the α -helix bundle displays a similar helix arrangement to

that of the leucine transporter LeuT_{Aa} of *Aquifex aeolicus*, which belongs to the neurotransmitter sodium symporter (NSS) family (TC 2.A.22, SLC6) (Faham *et al.*, 2008; Yamashita *et al.*, 2005). It may be anticipated that such structural homology of the core of the α -helix bundle also applies to the more closely related PutP of *E. coli*. Since there is currently no crystal structure of PutP available, we modeled the helix bundle in its substrate-free state on the basis of SDSL EPR distance distribution measurements and tested for structural homology between the two cores of vSGLT and PutP.

Assuming close-to-regular shapes of the helices, the fold of a helix bundle is defined by the relative positions and orientations of the N α -helices. Such a model is fully determined by the Cartesian coordinates of the $2N$ helix end points that can be derived from experimentally measured distances between labels attached near the helix ends. As the measurements are performed between the labels rather than helix end points, the labels have to be modeled explicitly to obtain a precise structural model with such an approach (Sale *et al.*, 2005; Bhatnagar *et al.*, 2007; Hilger *et al.*, 2007; Hilger *et al.*, 2009). Each helix is thus described by a four-point model, including the midpoints of the N-O bond of the cytoplasmic and periplasmic label and helix backbone end points (Fig. 5.1.). For the complete determination of a model with $N \geq 6$ helices, distances between $N(2N-1)$ pairs of labels are required to fix the $F = 12N-6$ degrees of freedom. For the case of PutP with 13 transmembrane domains this implies 325 distance measurements, i.e. 325 experimental restraints.

This considerable effort can be reduced by including additional structural information. As the core of the bundle is expected to be structurally homologous to the core of vSGLT, it is possible to derive approximate distances from the vSGLT template structure (Faham *et al.*, 2008). The deviation between the core structures of the newly modeled protein PutP and the template vSGLT can be considered by defining lower and upper bounds for the template restraints. Furthermore, a helix bundle is internally constrained by general features of protein structure, e.g. the approximate length of a helix with n residues or the maximum length of a loop with m residues that connects two helix ends. Moreover, the distance of a spin label from the helix axis and the angle between the linker and the helix axis are constrained by the structure of the label. All these internal constraints can also be expressed as distance restraints. Since the restraints are uncertain to some extent, each of them is given in terms of a lower and an upper bound. Additional information about the structure is available, but cannot easily be encoded in terms of distance restraints. For instance, helices can be modeled as cylinders with a radius of 0.3 nm, i.e. two helix axes may nowhere approach closer than 0.6 nm. Together with several other structural characteristics found for membrane protein helix bundles (Bowie, 1997), these can be used to define an 'energy

function' of the bundle that can then be utilized to distinguish between poor and good models.

Distance restraints with lower and upper bounds are best processed by the distance matrix geometry approach (Crippen and Havel, 1988) that was also used in the early days of protein structure determination by nuclear magnetic resonance (NMR) techniques (Wüthrich, 1990). Our approach for generating structural models thus consists of setting up matrices of lower l_{ij} and upper bounds u_{ij} for all distances d_{ij} in the model, bound smoothing based on triangle inequalities, embedding, and directed refinement. For PutP, we tested this procedure in modeling the helix bundle in its substrate-free state and obtained two ensembles of helix bundle structures.

The manuscript is structured as follows: First, we analyzed the crystal structures of LeuT_{Aa} (PDB code: 2A65) and vSGLT (PDB code: 3DH4) to derive internal constraints. Then we use the structural homology between vSGLT and LeuT_{Aa} to test our modeling procedure. Finally, we model the helix bundle of PutP and discuss its functional implications.

5.3. Materials and Methods

5.3.1. Sample preparation

The *putP* alleles encoding double-Cys PutP molecules used in this study were generated by site-directed mutagenesis using plasmid pT7-5/*putP*(Δ Cys) as a template and synthetic mutagenic oligonucleotides in one- or two-step PCR reactions using *Taq*-DNA polymerase. Plasmid pT7-5/*putP*(Δ Cys) is a derivative of pT7-5 (Tabor and Richardson, 1985) and contains the *lac* promoter/operator for expression of the *putP* gene and an engineered cassette version of this gene that is devoid of all five native Cys residues. PCR fragments were digested with appropriate restriction endonucleases and ligated with similarly treated plasmid pT7-5/*putP*(Δ Cys). For overexpression, the *putP* alleles were cloned into plasmid pTrc99a (Amann *et al.*, 1988) using restriction endonucleases *Nco*I and *Hind*III. The resulting plasmids were transformed into *E. coli* WG170 [F^- *trp lacZ rpsL thi* Δ (*putPA*)101 *proP219*] (Stalmach *et al.*, 1983). Cells were grown, membranes were prepared, and PutP was solubilized and purified by Ni-nitrilotriacetic acid affinity chromatography as previously described (Jung *et al.*, 1998b). The double-Cys PutP was labeled with (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate (MTSSL; Toronto Research Chemicals, Toronto, Canada) on the column. For this purpose, 1 mM MTSSL in buffer W (50 mM KP_i , pH 8.0, 300 mM KCl, 10 mM imidazole, 10% glycerol (v/v), 0.04% β -D-dodecylmaltoside (w/v)) was applied to the column and incubated at 4°C for 3 h. The unbound label was removed by washing the column with buffer W, and the labeled protein was eluted with 200 mM imidazole in buffer W. After elution, the protein was reconstituted under nonreducing conditions into

liposomes composed of *E. coli* lipids (67% phosphatidylethanolamine, 23.2% phosphatidylglycerol, and 9.8% cardiolipin; Avanti Polar Lipids, Alabaster, AL) at a lipid/protein ratio of 20:1 (w/w) as previously described (Jung *et al.*, 1998b). Finally, the proteoliposomes were washed twice with 50 mM KP_i, pH 7.5, and resuspended in the same buffer to yield a PutP concentration of 100-250 μ M. The proteoliposomes were frozen and stored in liquid nitrogen until use.

Unless the native Cys residue Cys349 was used for labeling, it was mutated to a Ser residue. The following spin-labeled double derivatives were prepared: S41R1 combined with A25R1, M62R1, K91R1, R126R1, I128R1, G179R1, Q190R1, H253R1, T276R1, C349R1, L371R1, S423R1, A430R1, Q446R1, and S470R1; L371 combined with A25R1, M62R1, K91R1, R126R1, I128R1, G179R1, Q190R1, H253R1, T276R1, C349R1, S423R1, A430R1, Q446R1, and S470R1; M62R1 combined with T5R1, I80R1, A149R1, L164R1, G210R1, L233R1, A294R1, I326R1, A391R1, L400R1, Q446R1, and L450R1 as well as Q446R1 combined with T5R1, I80R1, A149R1, L164R1, G210R1, L233R1, A294R1, I326R1, A391R1, L400R1, and L450R1.

The effect of double-Cys replacements on activity of PutP was tested by transport assays in intact cells. Active transport was measured in *E. coli* WG170 (PutP^A) harboring derivatives of plasmids pTrc99a/*putP*(Δ Cys) encoding PutP with given amino acid replacements. The cells were grown aerobically in Luria-Bertani (LB) medium (Miller, 1992) containing 100 μ g/ml ampicillin at 37°C. Overnight cultures were diluted 25-fold and were allowed to grow up to an optical density at 420 nm (A_{420}) of 1.0 followed by induction with 0.5 mM isopropyl 1-thio- β -D-galactopyranoside for 1 h. Cells were harvested, washed, and transport assays were performed under standard conditions (70 mM Na⁺, 10 μ M proline) as previously described (Hilger *et al.*, 2008; Raba *et al.*, 2008). Under these conditions, the majority of double-Cys mutants exhibited an initial rate of transport of minimum 15% of PutP(Δ Cys), which in turn has 50% of the initial rate of transport and 100% of the steady-state level of proline accumulation. Initial rates of transport were calculated from the initial linear portion of the time course, and steady-state levels of proline accumulation of the wild-type. Exceptions were derivatives S41C/G179C, L371C/G179C, M62C/G210C, Q446C/G210C, M62C/L400C, and Q446C/L400C, with initial rates of ~1.8%, 1.1%, 8.9%, 1.7%, 9.1%, and 9.1% of PutP(Δ Cys). The latter inhibitory effects on transport activity are possibly due to a functional relevance of these residues. The glycines at positions 179 and 210 are conserved within the members of the SSS family, whereas the leucine at position 400 is conserved in PutP homologues and panthotenate transporters of different organisms (suppl. Fig. 5.1.).

5.3.2. DEER measurements

Four-pulse DEER measurements were performed with Bruker Elexsys 580 spectrometers (Bruker Biospin GmbH, Karlsruhe, Germany) equipped with a 3 mm split-ring resonator under conditions of strong overcoupling ($Q \approx 100$) at a temperature of 50 K. Before insertion into the probe head, the samples were shock-frozen in liquid nitrogen to avoid crystallization of water. The four-pulse DEER sequence $(\pi/2)_{\nu_1} - \tau_1 - (\pi)_{\nu_1} - t' - (\pi)_{\nu_2} - \tau_1 + \tau_2 - t' - (\pi)_{\nu_1} - \tau_2 - echo$ was used (Pannier *et al.*, 2000). The $\pi/2$ and π pulses at the observer frequency ν_1 had either pulse lengths of 16 ns ($\pi/2$) and 32 ns (π) or equal pulse lengths of 32 ns dependent on the used spectrometer, while the π pulse at the pump frequency ν_2 had a length of 12 ns to maximize modulation depth while still keeping the two excitation bands separate. The long interpulse delay τ_2 , defining the dipolar evolution time, was typically 1500-2500 ns depending on the relaxation properties of the sample and on distance. An initial value of $t' = 80$ ns and an increment $\Delta t' = 8$ ns were used to acquire the time trace. To suppress proton modulation, data were added for eight equidistant values of τ_1 between 200 and 256 ns. A phase cycle $[+(+x), -(-x)]$ was applied to the first pulse. The pump frequency ν_2 (typically 9.33 GHz) was set to the center of the resonator mode and to coincide with the global maximum of the nitroxide spectrum. The observer frequency ν_1 was set to the local maximum at the low-field edge of the spectrum ($\nu_1 - \nu_2 = 65$ MHz). Accumulation times for the data sets varied between 8 and 14 h. Data were analyzed for dipolar evolution times $t = t' - \tau_1 \geq 0$. Analysis of the data in terms of distance distributions and mean distances was performed with the program DeerAnalysis2006 (Jeschke *et al.*, 2006).

5.3.3. Determination of number of spins per molecules

The number of spins per molecules was determined based on the calibration performed previously with a series of biradicals and a triradical (Hilger *et al.*, 2005) except for derivatives S41R1/H253R1, S41R1/A430R1, L371R1/R126R1, L371R1/H253R1, L371R1/A430R1, and T5R1/M62R1. For these derivatives, continuous-wave (CW) EPR spectra were recorded using a Miniscope MS200 X-band spectrometer (microwave frequency ≈ 9.4 GHz) equipped with a rectangular TE102 resonator (Magnettech GmbH, Berlin, Germany). All measurements were performed at 298.15 K adjusted with the temperature control unit TC H02 (Magnettech GmbH). The microwave power was set to 1 mW and the B-Field modulation amplitude was 0.12 mT. Doubly integrated EPR signal intensities were compared with double integrals of spectra of 100 μ M MTSSL to determine

spin concentrations in the samples. The spin numbers per molecule were calculated by dividing the nitroxyl spin concentration by the molar protein concentration.

5.3.4. Generation of template restraints

Template restraints for the endpoints of the target core helices of LeuT_{Aa} and PutP were derived from the vSGLT template structure (PDB code: 3DH4). To this end, the crystal structure of vSGLT was projected on a helix bundle model by determining the helix axes as the lines from which the sum of squares of the distances to the C^α atoms is a minimum. The helix lengths were calculated from the C^α coordinates of the residues at the respective helix ends. This helix bundle model was then adapted to the helix length of the target proteins. For this purpose, the helix axes of the target protein were superimposed on the helix axis of the template model. As the number of residues of the target helices generally differed from the corresponding template helices of vSGLT, the midpoint of each helix was determined and the distance between the midpoints and the respective helix ends was scaled by the number of the intermediate residues. Finally, the distances r_i between the endpoints of the ten target core helices were used to define upper and lower bounds

$$u_i = r_i + \Delta \quad (5.1.a)$$

$$l_i = r_i - \Delta \quad (5.1.b)$$

for the template restraints by considering putative deviations Δ between the template and the target structure.

5.3.5. Structure determination

Two matrices D_l and D_u with lower bounds l_{ij} and upper bounds u_{ij} for any pair of points P_i , P_j were generated that are consistent with the internal, template, and distance restraints. Intrinsic relations among these initial set of bounds were ensured by bound smoothing using the triangle inequalities (Crippen and Havel, 1988),

$$u_{ij} \leq u_{ik} + u_{jk} \quad (5.2.a)$$

$$l_{ij} \geq l_{ik} - u_{jk} \quad (5.2.b)$$

These inequalities were applied to the upper and lower bounds on the distance between any triples of points P_i , P_j , and P_k . If $u_{ij} > u_{ik} + u_{jk}$ was found then u_{ij} was replaced by its upper limit $u_{ik} + u_{jk}$. Similarly, if the lower bound l_{ij} was smaller than $l_{ik} - u_{jk}$, it was increased to its lower limit $l_{ik} - u_{jk}$. From the refined matrices D_l and D_u trial distance matrices were generated by selection of uniformly distributed random numbers within the permitted intervals using metrization (Havel, 1990). After each choice of random distance, triangle inequality bound smoothing was repeated. Before the smoothing, the upper and lower bounds for the selected distance were set to the random distance.

From the trial distance matrix D a metric matrix G was derived by embedding (Havel, 2003). To this end, the vector of squared distances of all points from the centroid of the model was determined. The elements of the metric matrix G_{ij} are then given by the scalar product of the two coordinate vectors \mathbf{r}_i and \mathbf{r}_j from the centroid to each of the points P_i and P_j . From the three largest positive eigenvalues λ_x , λ_y , and λ_z of this real symmetric matrix G with corresponding eigenvectors \mathbf{e}_x , \mathbf{e}_y , and \mathbf{e}_z of length K , the Cartesian coordinates of the K points of the helix bundle model were generated (Havel, 2003),

$$x_i = \sqrt{\lambda_x} \mathbf{e}_{xi} \quad (5.3.a)$$

$$y_i = \sqrt{\lambda_y} \mathbf{e}_{yi} \quad (5.3.b)$$

$$z_i = \sqrt{\lambda_z} \mathbf{e}_{zi}. \quad (5.3.c)$$

If one or more of the largest eigenvalues of the metric matrix G were found to be negative, the metric matrix G and the corresponding trial distance matrix D were rejected.

5.3.6. Structure refinement

The trial helix bundle models were tested for lower Δ_l and upper bound violations Δ_u for each pair of points P_i and P_j by

$$\Delta_l = l_{ij} - d_{ij} \quad (5.4.a)$$

$$\Delta_u = d_{ij} - u_{ij}, \quad (5.4.b)$$

whereas negative values of Δ_l and Δ_u are replaced by zero. The normalized bound violation

$$v_{ij} = \frac{\Delta_l - \Delta_u}{d_{ij}} \quad (5.5.)$$

was then used together with the distance vector \mathbf{d}_{ij} from point P_i to point P_j to compute the improved coordinates of the two points by

$$\mathbf{P}_i' = \mathbf{P}_i - \lambda v_{ij} \mathbf{d}_{ij} \quad (5.6.a)$$

and

$$\mathbf{P}_j' = \mathbf{P}_j + \lambda v_{ij} \mathbf{d}_{ij}, \quad (5.6.b)$$

where parameter λ is a scaling factor. By summation of corrections with respect to all points P_j a new coordinate of point P_i was obtained that takes into account all constraint violations in which this point is involved:

$$\mathbf{P}_i' = \mathbf{P}_i - \lambda \sum_j v_{ij} \mathbf{d}_{ij}. \quad (5.7.)$$

To complete one iteration step this correction was performed for all points P_i ($i = 1 \dots K$). In consecutive iterations the set of coordinates is expected to improve so that the v_{ij} decrease and λ can be increased. We find that

$$\lambda = \frac{k}{k_{\max}} \quad (5.8.)$$

works well, with k being the iteration number and k_{\max} the maximum permissible iteration number. Parameter k_{\max} has to be adapted to the number of points and the extent of violation of the constraints. We find good convergence behavior and acceptable convergence rate for

$$k_{\max} = \sum_i \sum_j v_{ij}^2, \quad (5.9.)$$

if the right-hand side of Eq (5.9.) is larger than 1000 and $k_{\max} = 1000$ otherwise.

Collisions between helices were resolved by determining the normalized bound violation for the helix midpoints M_i by

$$v_{ij} = \frac{6 \text{ \AA} - r_{ij}}{r_{ij}}, \quad (5.10.)$$

where r_{ij} is the distance of closest approach of the two helices. If v_{ij} was positive, the coordinate shift

$$\Delta P_k = -\lambda \sum_j v_{ij} \mathbf{m}_{ij} \quad (5.11.)$$

was applied to all points P_k in the helix with index i , where \mathbf{m}_{ij} is the midpoint-midpoint vector. The definition of the normalized bound violation in Eq. (5.10.) diverges if two helices in the preliminary model intersect. Therefore, a limit of $v_{ij} < 100$ was imposed. This limit also guards against too strong shifts in the rather arbitrary directions between helix midpoints that may occur for very closely approaching helices.

5.4. Results and Discussion

5.4.1. Derivation of internal restraints

Our coarse-grained model of a helix bundle, where each helix is described by four points, corresponding to the mean coordinate of the cytoplasmic label (A), the cytoplasmic (B) and periplasmic (C) end of the helix, and the periplasmic label (D), has some internal constraints on distances and angles (Fig. 5.1.).

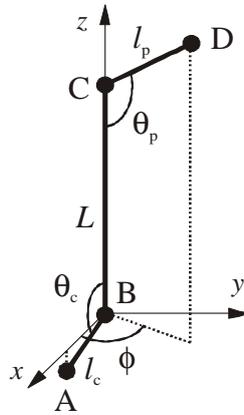


Figure 5.1.: Coarse-grained model of a TM helix that carries a spin label on its cytoplasmic and periplasmic end. The helix model is described by four points, corresponding to the mean coordinate of the cytoplasmic label (A), the cytoplasmic (B) and periplasmic (C) end of the helix, and the periplasmic label (D). Only the $2N$ points B_i and C_i ($i = 1 \dots N$) are relevant for describing the fold. Label coordinates correspond to the midpoint of the N-O bond in the nitroxide. L , helix length; l_c and l_p , length between the N-O midpoint of the cytoplasmic or periplasmic label and the corresponding end of the helix, respectively; θ_c and θ_p , angle between the helix axis and the vector \mathbf{l}_c or \mathbf{l}_p , respectively; ϕ , angle between the vectors \mathbf{l}_c and \mathbf{l}_p .

The helix length L is constrained by the number n of residues and by the observation that the length per residue of an α -helix varies only slightly. By projecting the X-ray structure of LeuT_{Aa} (PDB identifier 2A65) onto a bundle of straight helices, we obtained a minimum length per residue of 0.140 nm and a maximum length per residue of 0.153 nm. Therefore, we assumed the following internal restraints for the helix length:

$$[L_{\min}, L_{\max}] = n [0.135, 0.155] \text{ nm.} \quad (5.12.)$$

As an exception we found the discontinuous helix VI of LeuT_{Aa} that has a length per residue of only 0.130 nm. The constraints defined in Eq. (5.12.) lead to a bound violation by 0.123 nm for the total length of this helix. Given the expected precision of our bundle model, this violation is acceptable.

We also analyzed the distance of the mean N-O midpoint of the label from the corresponding end of the helix, i.e. lengths l_p and l_c in Fig. 5.1.. For this purpose the conformational distribution of the labels attached to helix ends in LeuT_{Aa} was simulated by a rotamer library approach (Jeschke and Polyhach, 2007). These lengths vary between 0.37 and 1.17 nm for LeuT_{Aa} and between 0.42 and 0.98 nm for vSGLT. Therefore, we defined the internal restraints of l_p and l_c as

$$[l_{c,\min}, l_{c,\max}] = [l_{p,\min}, l_{p,\max}] = [0.30, 1.30] \text{ nm.} \quad (5.13.)$$

Upper and lower limits for angles θ_p and θ_c can also be derived from the possible conformations of the label. Analysis for LeuT_{Aa} provides a range from 80.2° to 131.7°. Here, we assumed restraints of 70° and 140°, respectively, which can be converted to restraints on the label-to-label distance within the same helix. Angle ϕ appears to be unrestrained, as helix periodicity is far from ideal.

The minimum and maximum possible distance between two labels can then be computed as follows:

$$d_{cp,\min} = L_{\min} - 2 \cos(\theta_{\min}) l_{\max} = n \cdot 0.141 \text{ nm} - 0.89 \text{ nm}, \quad (5.14.a)$$

$$\begin{aligned} d_{cp,\max} &= 2[(L_{\max}/2 - \cos(\theta_{\max}) l_{\max})^2 + (\sin(\theta_{\max}) l_{\max})^2]^{1/2} \\ &= 2[(n/2 \cdot 0.141 + 0.995 \text{ \AA})^2 + (0.84)^2]^{1/2}, \end{aligned} \quad (5.14.b)$$

where we have considered that the minimum distance is obtained for $\phi = 0^\circ$ and the maximum distance for $\phi = 180^\circ$.

Distances between a label on one side of the membrane and the helix end on the other side (points A and C as well as B and D) can be restrained in analogy to geminal distances for a fixed bond angle and bond length (Crippen and Havel, 1988)

$$\begin{aligned} d_{AC,\min} &= [l_{c,\min}^2 + L_{\min}^2 - 2 l_{c,\min} L_{\min} \cos \theta_{c,\min}]^{1/2} \\ &= [0.09 + n^2 \cdot 0.019881 - 2 \cdot n \cdot 0.3 \cdot 0.141 \cdot 0.3420]^{1/2} \text{ nm} \\ &= [0.09 + n^2 \cdot 0.01988 - n \cdot 0.02893]^{1/2} \text{ nm}, \end{aligned} \quad (5.15.a)$$

$$d_{AC,\max} = [l_{c,\max}^2 + L_{\max}^2 - 2 l_{c,\max} L_{\max} \cos \theta_{c,\max}]^{1/2}$$

$$\begin{aligned}
&= [1.69 + n^2 \cdot 0.02161 + 2 \cdot n \cdot 1.3 \cdot 0.147 \cdot 0.7660]^{1/2} \text{ nm} \\
&= [1.69 + n^2 \cdot 0.02161 + n \cdot 0.2928]^{1/2} \text{ nm}. \tag{5.15.b}
\end{aligned}$$

Because of symmetry we have $d_{\text{BD},\text{min}} = d_{\text{AC},\text{min}}$ and $d_{\text{BD},\text{max}} = d_{\text{AC},\text{max}}$. Thus, all distances within the same helix are internally constrained.

5.4.2. Internal constraints on interhelical distances

Two consecutive helices are joined by a loop with m residues. This loop has a maximum length that corresponds to the contour length of a peptide chain. In crystal structures, the contour length is usually assumed as 0.36 nm/residue, corresponding to the length of extended antiparallel β -sheets (Pauling and Corey, 1951). Although values of up to 0.40 nm/residue have been found when extending peptide chains in atomic force microscopy measurements (Ainavarapu *et al.*, 2007), we stick with the former value that is in agreement with known equilibrium structures of proteins.

Two helices can be joined by a cytoplasmic loop between points B_i and B_{i+1} or by a periplasmic loop between points C_i and C_{i+1} . In both cases there is an upper bound

$$d_{\text{loop,max}} = m \cdot 0.36 \text{ nm} + 0.6 \text{ nm}. \tag{5.16.}$$

The second term on the right-hand side of Eq. (5.16.) is twice the helix radius. It is required as the bundle model specifies distances between helix end points situated on the helix axis rather than between C^α atoms of the labeled residues.

For loops that connect consecutive helices we use the lower and upper bounds given in Sale *et al.* (2004). The upper bound for short loops (up to 7 residues) was increased by 0.3 nm to account for the fact that Sale *et al.* (2004) appear to relate distances to the C^α atoms in the helix backbone while we relate them to the helix axis. For loops with more than 20 residues we increase the upper bound to $(m-20) \cdot 0.36 \text{ nm} + 3.2 \text{ nm}$.

Between any two helix ends on the same side of the membrane there exists a lower distance bound as the helices may not overlap. A study on helix-helix packing in proteins provides a lower bound for the radius of an α -helix of 0.3 nm (Walther *et al.*, 1996), thus we have

$$d_{\text{interhelix,min}} = 0.6 \text{ nm}. \tag{5.17.}$$

In the X-ray structure of LeuT_{Aa} helix ends on the same side of the membrane do not approach closer than 0.9 nm. The mean distance is 2.88 nm, well within the range accessible by DEER.

Helix ends on opposite sides of the membrane have a larger minimum distance. In the structure of LeuT_{Aa} we find that they can approach as closely as 2.5 nm, while their mean distance is 4.5 nm, and their maximum distance 7.4 nm. To be safe, we assume

$$d_{\text{TM,min}} = 2.0 \text{ nm}. \tag{5.18.}$$

The maximum distance between any two helix ends in LeuT_{Aa} is 7.4 nm. Of course, this value depends on protein size. In many cases, one will have an estimate of the maximum dimensions of the protein molecule. As a conservative estimate for PutP we use

$$d_{\text{interhelix,max}} = 10.0 \text{ nm.} \quad (5.19.a)$$

For helix ends on the same side of the membrane in LeuT_{Aa} we find maximum distances of 5.35 nm (cytoplasmic) and 5.50 nm (periplasmic). In this case we thus use a tighter upper bound

$$d_{\text{IH,cyto,max}} = d_{\text{IH,peri,max}} = 7.0 \text{ nm.} \quad (5.19.b)$$

Yet another, possibly tighter upper bound can be defined for the far ends of two consecutive helices (N and C terminus of a TM-helix-loop-TM-helix substructure). This is based on the fact that tilt angles of TM helices with respect to each other are restricted. In the structure of LeuT_{Aa} we find that the tilt angle θ_{ij} for any helix i with respect to any other helix j does not exceed 81° . We assume a maximum value $\theta_{ij,\text{max}} = 90^\circ$. We may then estimate the upper bound for this case as

$$d_{\text{consecutive,max}} = (L_i + L_j) \sin(\theta_{ij,\text{max}}/2) + d_{\text{loop,max}}, \quad (5.20.)$$

where L_i and L_j are the lengths of the two helices and $d_{\text{loop,max}}$ is defined by Eq. (5.16.).

Another internal constraint arises as two helices may not overlap anywhere. This constraint, however, cannot be easily implemented in the distance matrix D . Instead, it is considered in the error function for refinement of the initial model.

5.4.3. Test of the approach on LeuT_{Aa}

Analysis of the X-ray structure of vSGLT revealed an unexpected structural homology of the ten-helix core with the corresponding helices in LeuT_{Aa}, a member of a different transporter gene family (Faham *et al.*, 2008). The performance of our approach can thus be tested by modeling the LeuT_{Aa} helix bundle based on a template derived from the vSGLT structure. For that purpose, synthetic EPR restraints were generated in the following way. In analogy to the selection of labeling sites made for PutP, one residue near the cytoplasmic and one residue near the periplasmic helix end of each of the 12 TM helices of LeuT_{Aa} (PDB code: 2A65) were selected for labeling. The conformational distribution of the spin label at each labeled site was modeled by attaching 98 MTSSL rotamers from a precomputed rotamer library, determining their relative interaction energy with the proteins, and computing their populations according to a Boltzmann distribution at the expected glass transition temperature of a lipid bilayer (175 K) as described previously (Hilger *et al.*, 2009). One reference helix was selected and the 11 distance distributions between the label on the cytoplasmic end of this helix and the cytoplasmic ends of all other helices were computed from the respective distribution of label conformations. The same was done for the 11 pairs

of labels on the periplasmic side. Additionally, a second reference site was selected on each side of the membrane that was as far away as possible from the first reference point. This second reference point was used to compute ten further distance distributions between the labels at the cytoplasmic and periplasmic helix ends, respectively. Furthermore, the four transmembrane distance distributions between the reference sites were generated, so that altogether 46 synthetical EPR distance restraints were available.

To derive these distance restraints in the same way as for PutP, we first transformed the distance distributions to DEER form factors and multiplied them with experimental background functions. These background functions were generated by random linear combinations of two DEER data sets from singly spin-labeled PutP mutants. The thus simulated 46 primary DEER data sets for LeuT_{Aa} were encoded with six-digit random numbers to avoid subjective bias in their analysis. The coding table was stored but not looked at before the analysis was finished. The data sets were then analyzed in the same way as the experimental data for PutP to derive distance restraints. At last, the simulated EPR-based restraints were assigned to the label pairs using the coding table.

In the first modeling step the 46 simulated EPR-based restraints were used together with internal restraints and template restraints to set up matrices of lower bounds l_{ij} and upper bounds u_{ij} for the 1128 distances in the 48-point model of the spin-labeled helix bundle of LeuT_{Aa}. These matrices were subjected to triangle inequality bound smoothing to check the consistency of the initial set of distance limits. By applying the triangle inequalities, the matrices were found to be free of erroneous bounds $l_{ij} > u_{ij}$ and the removal of triangle inequality violations led to smaller distance intervals and thus to a more restricted model space. For computing different sets of model coordinates that satisfy the constraints, trial distance matrices were generated from the distance bounds via metrization (Havel, 1990). This was done by choosing each pair distance from uniformly distributed random numbers within the permitted intervals. After each distance choice the bound smoothing step was repeated to guarantee that the distance estimates will obey the triangle inequalities. The trial distance matrices were then translated into individual sets of Cartesian coordinates of all N points by embedding the constraints. Finally, bound violations and helix-helix clashes in the embedded structures were resolved by directed refinement as described in the Materials and Methods section. For LeuT_{Aa} we decided to generate two ensembles of 200 structural models. Typically 250 embeddings had to be performed to obtain 200 helix bundle structures. In the remaining embeddings directed refinement failed.

The RMSD Q_i of the helix endpoints between the i^{th} EPR-based structural model and the superimposed X-ray based model can be used as a criterion for the quality of the EPR-based model. However, for PutP and the majority of membrane proteins, no X-ray structure is available. Therefore, other criteria have to be found that correlate with Q_i and are thus

suitable for selecting between good and poor structural models. For this purpose, we analyzed the ensemble of structures in terms of structural characteristics of helical bundles of membrane proteins derived mainly from structures of G-protein-coupled receptors (Sale *et al.*, 2004; Bowie, 1997). Generally, only poor correlation between these characteristics and the RMSD Q_i were found with correlation coefficients R between -0.2 and 0.2. However, three characteristics represent good predictors of structural quality: a characteristic H_i based on the relative tilt of helices, a characteristic T_i based on the tilt of the helices with respect to the approximated membrane plane, and a characteristic M_i based on the distance of helix midpoints to the approximate membrane plane. The correlation with Q_i were found to be maximal for the linear combination $P_i = 0.0143 H_i + 0.9768 T_i + 0.0089 M_i$ with $R = 0.5294$. While this correlation coefficient is too small to use the penalty function P_i as the only criterion for selecting good structures, it is sufficiently large to recognize very poor structures.

Another criterion for structure selection can be derived solely from the internal properties of the ensemble. For that we define a matrix S_{ij} ($i, j = 1 \dots 200$) of pairwise RMSDs between the helix end points of structures in the ensemble. The distance of a given structure i from the center of the ensemble can then be characterized by the sum of the corresponding row of this matrix $V_i = \sum_j S_{ij}$. We find that for LeuT_{Aa} V_i nicely correlates with Q_i and, consequently, V_i also correlates with P_i . This indicates that the center of the ensemble of structures is the best representative of the true structure. The center can be computed by averaging of the end point coordinates. However, we find that such averaging usually leads to an average structure that features helix-helix collisions. Thus, we selected the central structure of the ensemble as the representative structure. The central structure is the structure with the minimum row sum V_i . It turns out that the quality of the central structure improves when the ensemble is first reduced to half its size by rejecting the 100 structures with the largest V_i .

The central structures thus obtained for the two ensembles of LeuT_{Aa} structures have penalty function values P_c of 8.0 and 16.1, respectively, that nicely conform to the value determined for the LeuT_{Aa} X-ray structure ($P_s = 13.5$). The RMSDs Q_c of the structures with respect to the X-ray structure are 7.17 Å and 8.65 Å, while the original template derived from the vSGLT structure has a Q_o of 11.65 Å. The EPR-based distance restraints thus lead to an improvement of the model (Fig. 5.2.). In particular, the relative arrangement of the two non-core TMs XI and XII with respect to the core, which is poorly predicted by the vSGLT-based template, is correctly elucidated by our approach, when compared to the X-ray structure of LeuT_{Aa}.

The situation is less clear for the core region. The RMSDs of helix end points for only the ten core helices are 6.08 Å and 6.36 Å for the two central structures and 6.49 Å for the

original structure based on only the vSGLT X-ray structure. Thus, for the core region the refinement is only marginal (Fig. 5.3.), i.e., the resolution of the EPR-based model is not sufficient to discuss differences in the core between vSGLT and LeuT_{Aa}.

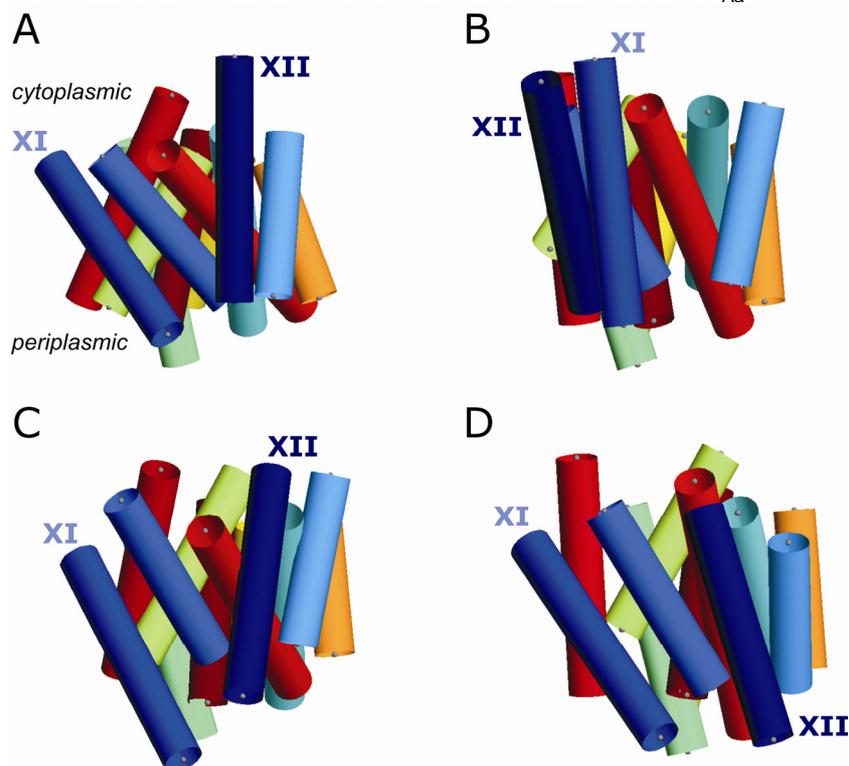


Figure 5.2.: Comparison of helix bundle models for LeuT_{Aa}. Helices are indicated by rigid cylinders and selected helices are numbered with Roman numerals. Helix ends are represented by grey dots. (A) Helix bundle model of the LeuT_{Aa} X-ray structure (PDB code: 2A65). (B) vSGLT-based LeuT_{Aa} template of the helix bundle. The crystal structure of vSGLT (Faham *et al.*, 2008) was projected on a helix bundle model by determining the helix axes as the lines from which the sum of squares of the distance to the C^α atoms is a minimum. The helix lengths were calculated from the C^α coordinates of the residues at the respective helix ends. This helix bundle model was then adapted to the helix length of the target protein. For this purpose, the helix axes of the target protein were superimposed on the helix axes of the template model. As the number of residues of the target helices generally differed from the corresponding template helices of vSGLT, the midpoint of each helix was determined and the distance between the midpoints and the respective helix ends was scaled by the number of the intermediate residues. (C and D) Central structures of two independently calculated EPR-based ensembles of LeuT_{Aa}. View perpendicular to the membrane normal.

The RMSD of the structural model can be estimated from the standard deviation of the coordinates in the ensemble. We find that the standard deviation of a subensemble of the 40 structures that are closest to the center of the full ensemble of 200 structures is a good approximation of the RMSD of the central structure with respect to the X-ray structure. The standard deviation of the full ensemble overestimates the RMSD by about a factor of 1.5. This is probably due to the use of a uniform rather than normal distribution of random numbers in the selection of distances in the intervals between the lower and upper bounds during metrization.

We also tested whether our approach is a critical test for the quality of the core template. For such tests we generated core decoys by random permutation of the helix assignment in the bundle model derived from the X-ray structure of vSGLT. This corresponds to a situation where the relative arrangement of helices in the bundle is known, but not their assignment. For a given relative arrangement of the 12 helices, 479 million distinct helix

assignments are possible. Among 100.000 randomly generated assignments, we found only 24 that allowed for the generation of error-free distance matrices from the lower and upper bounds derived from all internal, template, and experimental restraints. For two of these 24 decoys we tested embedding. Only 20 out of 103 embeddings succeeded.

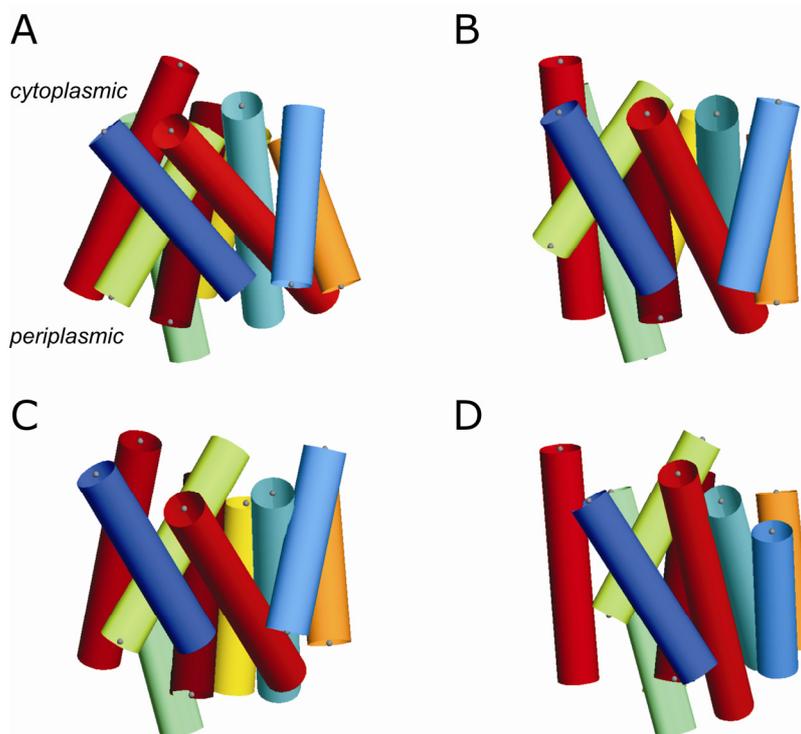


Figure 5.3.: Comparison of helix bundle core models for LeuT_{Aa}. Helices are indicated by rigid cylinders. Helix ends are represented by grey dots. (A) Helix bundle core model of the LeuT_{Aa} X-ray structure (PDB code: 2A65) (B) vSGLT-based LeuT_{Aa} template of the helix bundle core, generated as described in the legend of Figure 5.2.. (C and D) Central core structures of two independently calculated EPR-based ensembles of LeuT_{Aa}. View perpendicular to the membrane normal.

5.4.4. Helix bundle model for PutP

In the case of PutP that has one more helix than LeuT_{Aa}, analogous selection of labeling sites, including two reference points on each side of the membrane (41 and 371 at the cytoplasmic ends of TMs II and X; 62 and 446 at the periplasmic ends of TMs II and XII) and one position at each helix end of the other TMs (Fig. 5.4.), provides 50 distance restraints.

Of these, only 45 were obtained, as five double mutants do not show significant dipolar modulation after background correction (Table 5.1.). For double mutants 41/253, 41/430, 371/253, and 371/430 this is due to poor labeling efficiency. The average number of spins per molecule of these PutP derivatives was found to be approximately one (Table 5.1.), indicating that the cystein side chains introduced at positions 253 and 430 are not accessible to the spin label. For the fifth double mutant 62/5 we observed a significantly higher spin number of 1.35. In this case, the interspin distance probably lies below 1.2 nm and therefore out of the range of sensitivity of DEER measurements. Furthermore, fast transversal relaxation in the sample 371/126 only allowed the use of a dipolar evolution time of at most

0.9 μ s. As such short observation times prevent the reliable extraction of long distances (Jeschke and Polyhach, 2007), we decided to use the double mutant 371/128 that shows a significant lower transversal relaxation.

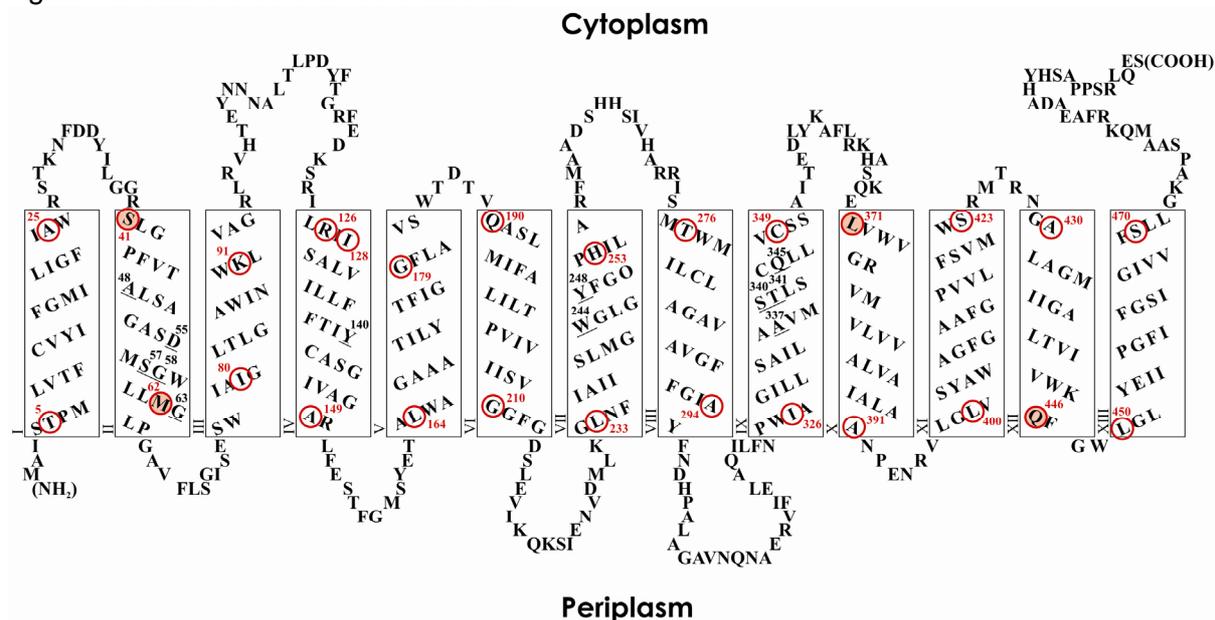


Figure 5.4.: 13-Helix secondary structure model of the Na^+ /proline transporter PutP of *E. coli* according to Jung *et al.* (1998a). Transmembrane domains are represented as rectangles and numbered with Roman numerals. Amino acid positions where spin labels were attached in this work are encircled (red) and labeled with red Arabic numerals. The four reference points used for construction of double-cysteine mutants are indicated by filled circles. Amino acids proposed to be involved in ligand binding are underlined and numbered with black Arabic numerals.

Because of the five missing distance restraints, the bundle model is slightly less restrained than for LeuT_{Aa}. In addition, the spread between the lower and upper distance bound tends to be somewhat larger for PutP than for LeuT_{Aa}. This can be traced back to larger uncertainties in background correction in the experimental data sets for PutP compared to the synthetically generated data sets for LeuT_{Aa}. Furthermore, the experimental distance distributions obtained for the PutP derivatives are significantly broader than the synthetically generated ones for LeuT_{Aa}, since they most likely represent different conformations of the PutP protein, while the data set for LeuT_{Aa} is based on the crystal structure of one single conformational state. Consequently the standard deviation of the structural ensemble is somewhat larger. Compared to standard deviations of 9.05 Å and 9.21 Å for the 20% of structures closest to the central structure in the two LeuT_{Aa} ensembles, we find standard deviations of 11.52 Å and 11.69 Å for the two PutP ensembles. To obtain better statistics for PutP, we generated ensembles of 300 structures. The mean deviation of the ten-helix core from the core template is about 6.5 Å for both PutP and LeuT_{Aa}. This indicates that the arrangement of the core helices in PutP does not deviate more strongly from vSGLT than LeuT_{Aa}. Again we find that the EPR distance restraints are a critical test for agreement of the core structure with the template. Out of 10.000 decoys generated by random permutations of helix order only 45 could be metrized with a distance uncertainty of $\Delta = 8$ Å. Due to the smaller ratio of experimental restraints to degrees of freedom and the larger

uncertainty of the experimental restraints, this fraction of feasible decoys is larger than for LeuT_{Aa}. However, it is still very unlikely that the set of restraints is consistent with an erroneous structure of the core.

Table 5.1.: Mean distances between spin labels in double mutants of PutP (in nanometers).

A) Double mutants with cytoplasmic reference points (cRP) 1 and 2.

cRP 1	Residue 2	Mean distance	Width	Spins/molecule	cRP 2	Residue 2	Mean distance	Width	Spins/molecule
41	25	2.28	0.49	1.41	371	25	3.20	0.36	1.63
41	91	3.06	0.49	1.51	371	91	3.99	0.49	1.41
41	126	4.25	0.69	1.67	371	126	n.a.	n.a.	1.36
41	179	4.31	0.36	1.21	371	128	1.99	0.34	1.21
41	190	2.92	0.40	1.43	371	179	2.10	0.51	1.36
41	253	n.a.	n.a.	1.01	371	190	2.86	0.34	1.36
41	276	2.47	0.55	1.31	371	253	n.a.	n.a.	0.97
41	349	3.79	0.80	1.81	371	276	3.94	0.38	2.80
41	371	5.04	0.74	1.53	371	349	1.96	0.50	1.60
41	423	4.46	0.51	1.47	371	423	3.85	0.19	1.58
41	430	n.a.	n.a.	1.03	371	430	n.a.	n.a.	0.94
41	470	3.22	0.80	1.47	371	470	3.22	0.83	1.58

B) Double mutants with periplasmic reference points (pRP) 1 and 2.

pRP 1	Residue 2	Mean distance	Width	Spins/molecule	pRP 2	Residue 2	Mean distance	Width	Spins/molecule
62	5	n.a.	n.a.	1.35	446	5	4.92	0.66	1.64
62	80	2.00	0.44	1.51	446	80	3.64	0.55	2.18
62	149	3.92	1.05	1.36	446	149	3.22	0.97	1.38
62	164	3.50	0.99	1.63	446	164	2.14	0.42	1.68
62	210	3.70	1.02	1.39	446	210	4.27	0.65	1.46
62	233	3.54	0.64	1.16	446	233	3.50	0.31	1.66
62	294	3.45	0.84	1.55	446	294	4.40	0.66	1.58
62	326	3.19	1.06	1.73	446	326	4.23	0.54	1.97
62	391	3.88	1.05	1.29	446	391	1.73	0.48	1.88
62	400	2.30	0.31	1.11	446	400	1.40	0.27	1.27
62	446	4.12	0.68	1.29	446	450	2.29	0.16	1.42
62	450	3.95	0.56	1.30					

C) Transmembrane distances between cRPs and pRPs.

Residue 1	Residue 2	Mean distance	Width	Spins/Molecule
41	62	3.42	0.80	1.64
41	446	4.67	0.47	1.40
62	371	4.26	0.62	1.70
446	371	3.29	0.36	1.77

The number of spins per molecules characterizes the degree of labeling.

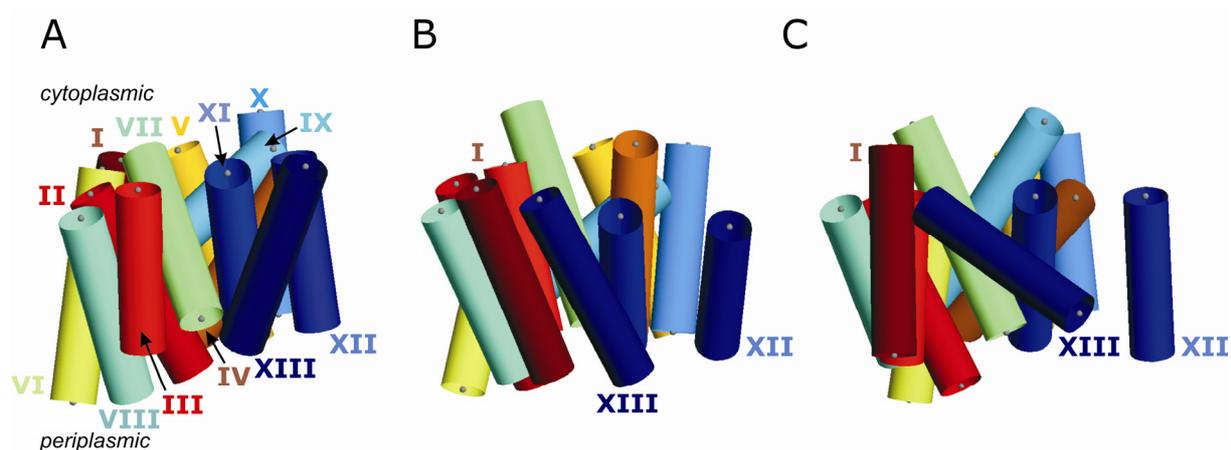


Figure 5.5.: Comparison of helix bundle models for PutP. Helices are indicated by rigid cylinders and selected helices are numbered with Roman numerals. Helix ends are represented by grey dots. (A) vSGLT based PutP template of the helix bundle, generated as described in the legend of Figure 5.2.. (B and C) Central core structures of two independently calculated EPR-based ensembles of PutP. View perpendicular to the membrane normal.

The central structures of the two ensembles (Fig. 5.5.) have penalty function values P_c of 1.0 and 10.8, i.e. they nicely conform to expectations for a helical bundle of a membrane protein.

As it was already found for LeuT_{Aa}, differences in the core are only marginal (Fig. 5.6.). Similar to vSGLT, the PutP core structures are composed of two inverted repeat domains (TMs II-VI and TMs VII-XI) that are related by a pseudo-two-fold axis in the plane of the membrane. These structural repeats are intertwined, forming two interior domains: a central four-helix bundle consisting of TMs II, III, VII, and VIII, which is surrounded by a scaffold defined by TMs IV, V, VI, IX, X, and XI (Fig. 5.6.). The most relevant differences with respect to the template based on the vSGLT structure concern the three non-core helices, whose positions are determined exclusively by internal and experimental restraints. Helix I, which is found in the back of the bundle for vSGLT is situated in the front in both experimental structures. For helix XIII the position of the periplasmic end is not significantly different, however, the tilt direction differs and thus the cytoplasmic end near the C terminus of the protein is found in a different position. TM XII is similarly positioned relative to the core in the PutP models and the vSGLT template, but it is more separated from the remaining helices in the PutP structures than in the vSGLT template (Fig. 5.5.). Although the positions of the helices in the two ensembles are fully consistent with the available restraints, such a separation of TM XII is presumably energetically unfavorable due to the absence of helix contacts between TM XII and other helices of the bundle. However, it must be taken into account that the calculated central structures of the ensembles represent only two out of many possibilities, which are consistent with the restraints. Therefore, it is feasible that in other structures TM XII is in close contact to the remaining helix bundle. To select those structures in which the individual helices are in contact with at least one neighboring helix, a penalty function can be incorporated into our approach. However, the contact penalty

function published by Sale *et al.* (2004) cannot be used and must be redefined, since it is not suitable for discrimination between good and poor structures in the case of LeuT_{Aa}. Work along these lines is now in progress. Furthermore, more distance restraints are needed, especially between the two reference points and the end of TM XII on the cytoplasmic side of the membrane, which cannot be obtained so far.

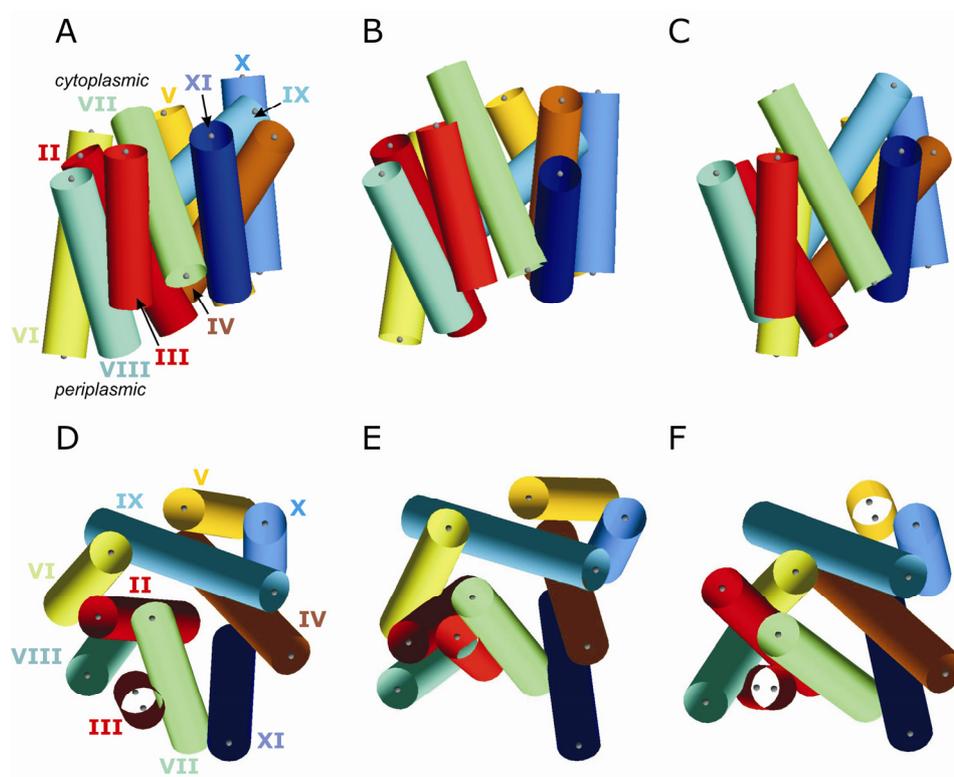


Figure 5.6.: Comparison of helix bundle core models for PutP. Helices are indicated by rigid cylinders and selected helices are numbered with Roman numerals. Helix ends are represented by grey dots. (A, D) vSGLT based PutP template of the helix bundle core, generated as described in the legend of Figure 5.2.. (B, E and C, F) Central structures of two independently calculated EPR-based ensemble of PutP structures. (A-C) View perpendicular to the membrane normal. (D-F) Top view of the core from the cytoplasm.

5.4.5. Functional implications of the helix bundle models

Structure-function analyses of PutP provided evidence that TM II and IX are involved in the formation of a ligand-binding site (Quick *et al.*, 1996; Quick and Jung, 1997; Pirch *et al.*, 2002; Hilger *et al.*, 2008; Raba *et al.*, 2008). TM II contains residue Asp55 that is proposed to directly participate in binding of the coupling ion (Fig. 5.4.). In addition, residues Ser57 and Gly58 of the same helix were found to be important for Na⁺ and/or proline binding, suggesting closely located binding sites for both ligands (Quick *et al.*, 1996; Quick and Jung, 1997; Pirch *et al.*, 2002). TM IX harbors residues Ser340 and Thr341, which are also required for high-affinity Na⁺ and proline binding (Hilger *et al.*, 2008) (Fig. 5.4.). Finally, cross-linking experiments and Cys accessibility studies revealed that TMs II and IX are in close proximity to each other in the tertiary structure of the protein and contribute to the formation of a translocation pathway (Pirch *et al.*, 2003; Hilger *et al.*, 2008; Raba *et al.*, 2008).

However, the complete set of helices that participate in the formation of the ion- and/or substrate-binding site in PutP is not known. Given the structural homology between the core of PutP and vSGLT, we suggest a common transport mechanism for these proteins, including principles in ligand binding. Indeed, the X-ray structures of the structural homologous vSGLT and LeuT_{Aa} show similarly located substrate-binding sites at the centre of the proteins, even though they belong to different transporter gene families and transport different substrates (Faham *et al.*, 2008; Yamashita *et al.*, 2005). The helix arrangement in the PutP structures and its comparison with the vSGLT and LeuT_{Aa} can thus be used to draw conclusions about the TMs that are involved in ligand binding in PutP. This information can then guide further functional studies to identify residues that are particularly important for the transport activity of PutP.

In the two helix bundle structures obtained for PutP, TM II and TM IX are located in the core domain (Fig. 5.6.e, f). TM II forms part of the central four-helix bundle, whereas TM IX is situated in the surrounding scaffold. The position of both segments relative to each other is found to be somewhat different in the two bundle structures. In the first one, TM II faces TM IX (Fig. 5.6.e), while in the second structure the two transmembrane domains are separated by TM VI (Fig. 5.6.f). Although both models are consistent with the available restraints, effective cross-linking between positions in TM II and TM IX, even by a zero length cross-linker, favors an arrangement similar to that observed in the first structure (Hilger *et al.*, 2008) (Fig. 5.6.b, e). Such an arrangement is also found in the vSGLT X-ray structure in which the corresponding helices are in close vicinity to each other, forming a sodium-binding site in approximately the middle of the membrane (Faham *et al.*, 2008) (Fig. 5.6.a, d). Thereby, the coordination of the coupling ion is predicted to be achieved by one hydroxyl oxygen of Ser365 of TM IX and three carbonyl oxygens of Ala62, Ile65 (TM II), and Ala361 (TM IX). This sodium-binding scenario is in part resembled by the situation suggested for PutP. In fact the hydroxyl moiety of residue Thr341, which is conserved throughout the SSS family, is at equivalent position to Ser365 in TM IX of vSGLT (suppl. Fig. 5.1.). Functional important main-chain carbonyl groups are more difficult to identify as the side chains of the corresponding amino acids are not necessarily highly conserved and are not required to possess specific properties (e.g., charged or polar groups). Nevertheless, replacement of conserved Ala48 and Gly63 in TM II and Ala337 (corresponds to Ala361 of vSGLT) (suppl. Fig. 5.1.) in TM IX of PutP has an impact on sodium-dependent transport (Raba *et al.*, 2008; Pirch *et al.*, 2003; Hilger *et al.*, 2008) (Fig. 5.4.). Thus, it is tempting to speculate that the main-chain carbonyl groups of some of these residues contribute to sodium binding. In addition, functional studies on Ser340 suggested that this residue may participate in ion binding via its main-chain carbonyl group (Hilger *et al.*, 2008) (Fig. 5.4.). Besides the hydroxyl and carbonyl oxygens, the carboxyl group of Asp55 in TM II of PutP is proposed to

be essential for sodium dependence (Quick and Jung, 1997). However, the general importance of this residue is less clear as it is not conserved within the SSS family as e.g. Thr341 (suppl. Fig. 5.1.). Possibly, Asp55 stabilizes the empty and/or occupied sodium-binding site in PutP as it is most likely deprotonated at physiological pH by additional hydrogen bonds. Furthermore, because TM II is supposed to be involved in both ion and substrate binding, the properties of the substrate (e.g. charged proline versus polar sugar) may also have influenced the precise mechanism of Na⁺ binding. Taken together, the obtained helix bundle structures in conjunction with structure-function analyses and comparison with the vSGLT template strongly support the idea of a Na⁺-binding site that is located at the interface between TM II and TM IX. However, additional distance measurements are required for more reliable determination of the relative position of both helices in the bundle structures of PutP. Nevertheless, with exception of Asp55 of PutP, the coordination pattern seems to be conserved in both transporters as it is formed in each case by hydroxyl groups from serine or threonine of TM IX and carbonyl oxygens of TM II and TM IX. Intriguingly, this sodium-binding scenario is also present in LeuT_{Aa} that has a different Na⁺:substrate stoichiometry of 2:1 in comparison to the 1:1 stoichiometry of PutP and vSGLT and thus possesses two sodium-binding sites (Na₁⁺ and Na₂⁺) (Yamashita *et al.*, 2005). Thereby, Na₂⁺ (Ala351, Ser355, and Thr354 of TM VIII and Gly20 and Val23 of TM I) corresponds to the likely sodium-binding sites of PutP and vSGLT, whereas for Na₁⁺ no equivalent site in PutP or vSGLT exists. This indicates that Na₂⁺ probably plays an essential role for sodium-coupled substrate translocation in members of this structural family.

As mentioned above, TM II and TM IX of PutP are not only proposed to be important for sodium coordination, but also for binding of the substrate. In fact, site-specific amino acid substitution analyses indicate that the conserved residues Ser57 and Gly58 of TM II and Thr341 of TM IX contribute to the formation of a substrate-binding site that is situated in close vicinity to the ion-binding site (Quick *et al.*, 1996; Pirch *et al.*, 2002; Hilger *et al.*, 2008). However, information about the complete ensemble of helices that participate in substrate binding is not available yet. In the obtained bundle structures of PutP, five transmembrane domains (TMs IV, VI, and XI of the scaffold and TMs III and VII of the four-helix bundle) enclose the proposed proline-binding site in addition to TM II and TM IX and thus represent appropriate candidates for substrate coordination (Fig. 5.6.b, e and c, f). Two of these helices, TM IV and TM VII, have recently been found to be of functional relevance (Raba, M., Pirch, T., and Jung, H., unpublished data). Thereby, residues Tyr140 in TM IV and Trp244 and Tyr248 in TM VII, positions highly conserved throughout the proline symporters of the SSS family (suppl. Fig. 5.1.), were found to be essential for sodium-dependent proline transport (Fig. 5.4.). Although helices III, VI, and XI have not been analyzed yet, the available data support the idea of a proline-binding site formed by TMs II, IV, VII, and IX. Interestingly,

this arrangement resembles only in part the situation found in sugar transporter vSGLT, whereas it is fully consistent with the binding region in LeuT_{Aa} that catalyzes, like PutP, the uptake of amino acids. In particular, the galactose-binding site in vSGLT encompasses helices II, III, VII, and VIII of the four-helix bundle and TM XI of the surrounding scaffold and is located approximately 1 nm away from the ion-binding site (Faham *et al.*, 2008). The sugar in vSGLT is almost exclusively coordinated via side-chain interactions, especially by H-bonding with polar amino acids. In the X-ray structure of LeuT_{Aa}, the substrate-binding site is located at a slightly different position between TM I and TM VI of the four-helix bundle and TM III and TM VIII of the scaffold (Yamashita *et al.*, 2005), which correspond to TMs II, VII, IV, and IX of PutP. The α -substitutions of leucine are positioned near the unwound regions of the discontinuous helices I and VI of LeuT_{Aa} in a distance of ~ 0.6 nm to the sodium-binding site Na₂⁺. The α -amino and α -carboxy groups are predominantly exposed to main-chain atoms of the extended peptides, thereby exploiting helix dipole moments at the end of the helical segments. Only two amino acid side chains are involved in the coordination of the α -amino and α -carboxyl group of the leucine molecule, Tyr108 in TM III and Ser256 in TM VI. In PutP, Tyr140 of TM IV and the hydroxyl group of Tyr248 of TM VII are at equivalent position to Tyr108 and Ser256 of LeuT_{Aa} and thus may fulfill similar function in coordinating the carboxyl group of the substrate. Furthermore, the observed functional relevance of Gly58 in TM II of PutP is in line with the main-chain interaction between the α -carboxy group of leucine and the equivalent residue Gly26 in TM I of LeuT_{Aa}. In addition to these residues, the side chain of Ser57 in TM II of PutP is proposed to be crucial for high-affinity proline uptake (Quick *et al.*, 1996). However, no corresponding residue is involved in leucine binding in TM I of LeuT_{Aa}. This is most likely due to the different substrate specificities of LeuT_{Aa} and PutP, which is determined by variation of the specific interaction sites. In contrast to the α -substituents that are proposed to be bound in a more polar binding region, the aliphatic side chain of leucine is located within a hydrophobic pocket formed by residues of TMs III, VI, and VIII of LeuT_{Aa}. One of these residues, Ser355 of TM VIII is engaged in substrate and Na⁺ binding as it is also suggested for the corresponding residue Thr341 of TM IX in PutP. It is therefore imaginable that Thr341 also contributes to the formation of a hydrophobic cavity that accommodates the cyclic aliphatic side chain of proline. In addition to Thr341, Trp244 in TM VII and Gln345 in TM IX of PutP could also be part of this pocket, since they are conserved throughout the proline transporters of the SSS family and are crucial for sodium-dependent proline uptake (Raba *et al.*, 2008; Raba, M., Jung, H., unpublished results).

Taken together, the helix bundle models in combination with biochemical and mutagenesis studies on PutP indicate a proline-binding site in PutP, which is created by TMs II, IV, VII, and IX. This helix quartet only partially agrees with the substrate-binding site of vSGLT, but is fully concordant with the situation found in LeuT_{Aa}, suggesting that the position

of the binding site in this structural family slightly varies depending on the chemical nature of the substrate (sugar vs. amino acid). However, it must be stated that the complete set of residues engaged in substrate binding is not known for PutP and no functional data is available for TMs III, VI, and XI corresponding to the TMs taking part in substrate binding in vSGLT but not in LeuT_{Aa}. Clearly, more information on transporter function is thus required to verify this putative proline-binding site. Since the crystal structures of vSGLT and LeuT_{Aa} reside in two different conformations, vSGLT in the inward-facing conformation vs. LeuT_{Aa} in the outward-facing conformation, it is also possible that TMs II, III, VII, VIII, and XI take part in the initial coordination of the substrate from the cell outside, whereas TMs II, IV, VII and IX participate in substrate binding after conversion from the outward-facing to the inward-facing conformation, before the substrate is released into the cell inside. However, this is highly speculative and needs to be proven by crystallizing one individual transporter with bound substrate in these two different conformations.

5.5. Conclusions

It has been demonstrated that the helix bundle structure of the 13 transmembrane domain membrane protein PutP can be modeled by a distance geometry approach using sparse DEER distance data together with template restraints derived from the ten-helix core of the closely related transporter vSGLT and internal restraints of the coarse-grained helix bundle model. The newly introduced methodology distinguishes with high certainty between structurally homologous and structurally unrelated core templates. Hence, the fold of the core can be reliably recognized by this approach.

For PutP, the fold of the ten-helix core is found to be homologous to the core structure of vSGLT. Since this architecture has been seen before in LeuT_{Aa}, the core domain appears to be the elementary machinery of membrane transport proteins of this structural family (Faham *et al.*, 2008; Yamashita *et al.*, 2005; Lolkema and Slotboom, 2008). The similarity in fold among PutP, vSGLT, and LeuT_{Aa} implicates commonalities in transport mechanism, including principles in ligand binding. The most relevant difference between the helix structures and the template based on the vSGLT structure concerns the three auxiliary helices, whose positions are established exclusively by internal and experimental restraints. However, it has to be noticed that the accuracy of the PutP models is still low (RMSD value of the helix end points about 12 Å), and additional distance measurements are required to reveal more subtle differences between the transporter structures. Along this line, the helix bundle models can be fine-grained by determining the backbone structures of the transmembrane domains as it was previously shown for TM IX of PutP (Hilger *et al.*, 2009) and modeling of missing loops using the “Modeller” software (Fiser *et al.*, 2000). This allows us in principle to establish the structure of membrane proteins at the level of the backbone

fold. In contrast to most established techniques, this approach is applicable to membrane proteins in their native environment of a lipid bilayer and is thus also accessible to molecules that evade crystallization.

5.6. Acknowledgments

The authors thank Christian Bauer for technical support and Hans Wolfgang Spiess and Heinz-Jürgen Steinhoff for providing access to pulsed EPR spectrometers. Financial support of Deutsche Forschungsgemeinschaft is gratefully acknowledged.

5.7. References

- Ainavarapu, S. R., Brujic, J., Huang, H. H., Wiita, A. P., Lu, H., Li, L., Walther, K. A., Carrion-Vazquez, M., Li, H., and Fernandez, J. M.** (2007) Contour length and refolding rate of a small protein controlled by engineered disulfide bonds. *Biophys. J.* 92: 225-233.
- Amann, E., Ochs, B., and Abel, K. J.** (1988) Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* 69: 301-315.
- Bhatnagar, J., Freed, J. H., and Crane, B. R.** (2007) Rigid body refinement of protein complexes with long-range distance restraints from pulsed dipolar ESR. *Methods Enzymol.* 423: 117-133.
- Bowie, J. U.** (1997) Helix packing in membrane proteins. *J. Mol. Biol.* 272: 780-789.
- Crippen, G. M., and Havel, T. F.** (1988) *Distance geometry and molecular conformation*, Research Studies Press Ltd., Taunton.
- Dohán, O., De la Vieja, A., and Carrasco, N.** (2006) Hydrocortisone and purinergic signaling stimulate sodium/iodide symporter (NIS)-mediated iodide transport in breast cancer cells. *Mol. Endocrinol.* 20: 1121-1137.
- Faham, S., Watanabe, A., Besserer, G. M., Cascio, D., Specht, A., Hirayama, B. A., Wright, E. M., and Abramson, J.** (2008) The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na⁺/sugar symport. *Science* 321: 810-814.
- Fiser, A., Do, R. K., and Sali, A.** (2000) Modeling of loops in protein structures. *Protein Sci.* 9: 1753-1773.
- Havel, T. F.** (1990) The sampling properties of some distance geometry algorithms applied to polypeptide chains: a study of 1830 independently computed conformations. *Biopolymers* 29: 1565-1585.
- Havel, T. F.** (2003) Metric matrix embedding in protein structure calculations, NMR spectra analysis, and relaxation theory. *Magn. Reson. Chem.* 41: 37-50.

- Hilger, D., Jung, H., Padan, E., Wegener, C., Vogel, K.-P., Steinhoff, H.-J., and Jeschke, G. (2005) Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na⁺/H⁺ antiporter of *E. coli*. *Biophys. J.* 89: 1328-1338.
- Hilger, D., Polyhach, Y., Padan, E., Jung, H., and Jeschke, G. (2007) High-resolution structure of a Na⁺/H⁺ antiporter dimer obtained by pulsed EPR distance measurements. *Biophys. J.* 93: 3675-3683.
- Hilger, D., Böhm, M., Hackmann, A., and Jung, H. (2008) Role of Ser-340 and Thr-341 in transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli* in ligand binding and transport. *J. Biol. Chem.* 283: 4921-4929.
- Hilger, D., Polyhach, Y., Jung, H., and Jeschke, G. (2009) Backbone structure of transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli*. *Biophys. J.* 96: 217-225.
- Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H. (2006) DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Reson.* 30: 473-498.
- Jeschke G., and Polyhach, Y. (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* 9: 1895-1910.
- Jung, H., Rübenhagen, R., Tebbe, S., Leifker, K., Tholema, N., Quick, M., and Schmid, R. (1998a) Topology of the Na⁺/proline transporter of *Escherichia coli*. *J. Biol. Chem.* 273: 26400-26407.
- Jung, H., Tebbe, S., Schmid, R., and Jung, K. (1998b) Unidirectional reconstitution and characterization of purified Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 37: 11083-11088.
- Jung, H. (2001) Towards the molecular mechanism of Na⁺/solute symport in prokaryotes. *Biochim. Biophys. Acta* 1505: 131-143.
- Jung, H. (2002) The sodium/substrate symporter family: structural and functional features. *FEBS Lett.* 529: 73-77.
- Kavermann, H., Burns, B. P., Angermüller, K., Odenbreit, S., Fischer, W., Melchers, K., and Haas, R. (2003) Identification and characterization of *Helicobacter pylori* genes essential for gastric colonization. *J. Exp. Med.* 197: 813-822.
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305: 567-580.
- Liu, Y. S., Sompornpisut, P., and Perozo, E. (2001) Structure of the KcsA channel intracellular gate in the open state. *Nat. Struct. Biol.* 8: 883-887.
- Lolkema, J. S., and Slotboom, D. J. (2008) The major amino acid transporter superfamily has a similar core structure as Na⁺-galactose and Na⁺-leucine transporters. *Mol. Membr. Biol.* 25: 567-570.

- Majumdar, D. S., Smirnova, I., Kasho, V., Nir, E., Kong, X., Weiss, S., and Kaback, H. R.** (2007) Single-molecule FRET reveals sugar-induced conformational dynamics in LacY. *Proc. Natl. Acad. Sci. U. S. A.* 104: 12640-12645.
- Miller, J. H.** (1992) *A short course in bacterial genetics. A laboratory manual and handbook for Escherichia coli and related bacteria*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W.** (2000) Dead-time free measurement of dipole-dipole interactions between electron spins. *J. Magn. Reson.* 142: 331-340.
- Pauling, L., and Corey, R. B.** (1951) The pleated sheet, a new layer configuration of polypeptide chains. *Proc. Natl. Acad. Sci. U. S. A.* 37: 251-256.
- Pirch, T., Quick, M., Nietschke, M., Langkamp, M., and Jung, H.** (2002) Sites important for Na⁺ and substrate binding in the Na⁺/proline transporter of *Escherichia coli*, a member of the Na⁺/solute symporter family. *J. Biol. Chem.* 277: 8790-8796.
- Pirch, T., Landmeier, S., and Jung, H.** (2003) Transmembrane domain II of the Na⁺/proline transporter PutP of *Escherichia coli* forms part of a conformationally flexible, cytoplasmic exposed aqueous cavity within the membrane. *J. Biol. Chem.* 278: 42942-42949.
- Quick, M., Tebbe, S., and Jung, H.** (1996) Ser57 in the Na⁺/proline permease of *Escherichia coli* is critical for high-affinity proline uptake. *Eur. J. Biochem.* 239: 732-736.
- Quick, M., and Jung, H.** (1997) Aspartate 55 in the Na⁺/proline permease of *Escherichia coli* is essential for Na⁺-coupled proline uptake. *Biochemistry* 36: 4631-4636.
- Raba, M., Baumgartner, T., Hilger, D., Klempahn, K., Härtel, T., Jung, K., and Jung, H.** (2008) Function of transmembrane domain IX in the Na⁺/proline transporter PutP. *J. Mol. Biol.* 382: 884-893.
- Reed-Tsur, M. D., De la Vieja, A., Ginter, C. S., and Carrasco, N.** (2008) Molecular characterization of V59E NIS, a Na⁺/I⁻ symporter mutant that causes congenital I⁻ transport defect. *Endocrinology* 149: 3077-3084.
- Reizer, J., Reizer, A., and Saier, M. H. Jr.** (1994) A functional superfamily of sodium/solute symporters. *Biochim. Biophys. Acta* 1197: 133-166.
- Saier, M. H. Jr.** (2003) Tracing pathways of transport protein evolution. *Mol. Microbiol.* 48: 1145-1156.
- Sale, K., Faulon, J. L., Gray, G. A., Schoeniger, J. S., and Young, M. M.** (2004) Optimal bundling of transmembrane helices using sparse distance constraints. *Protein Sci.* 13: 2613-2627.
- Sale, K., Song, L., Liu, Y. S., Perozo, E., and Fajer, P.** (2005) Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *J. Am. Chem. Soc.* 127: 9334-9335.

- Schwan, W. R., Lehmann, L., and McCormick, J.** (2006) Transcriptional activation of the *Staphylococcus aureus putP* gene by low-proline-high osmotic conditions and during infection of murine and human tissues. *Infect. Immun.* 74: 399-409.
- Smirnova, I., Kasho, V., Choe, J. Y., Altenbach, C., Hubbell, W. L., and Kaback, H. R.** (2007) Sugar binding induces an outward facing conformation of LacY. *Proc. Natl. Acad. Sci. U. S. A.* 104: 16504-16509.
- Stalmach, M. E., Grothe, S., and Wood, J. M.** (1983) Two proline porters in *Escherichia coli* K-12. *J. Bacteriol.* 156: 481-486.
- Stevens, T. J., and Arkin, I. T.** (2000) Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins* 39: 417-420.
- Tabor, S., and Richardson, C. C.** (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U. S. A.* 82: 1074-1078.
- Walther, D., Eisenhaber, F., and Argos, P.** (1996) Principles of helix-helix packing in proteins: the helical lattice superposition model. *J. Mol. Biol.* 255: 536-553.
- Wegener, C., Tebbe, S., Steinhoff, H.-J., and Jung, H.** (2000) Spin labeling analysis of structure and dynamics of the Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 39: 4831-4837.
- Wright, E. M., and Turk, E.** (2004) The sodium/glucose cotransport family SLC5. *Pfluegers Arch. Eur. J. Physiol.* 447: 510-518.
- Wright, E. M., Hirayama, B. A., and Loo, D. F.** (2007) Active sugar transport in health and disease. *J. Intern. Med.* 261: 32-43.
- Wüthrich, K.** (1990) Protein structure determination in solution by NMR spectroscopy. *J. Biol. Chem.* 265: 22059-22062.
- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E.** (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* 437: 215-223.

5.8. Supplementary Material

	1	10	20	30	40	50	60	70	80	90	TM III																			
sp_P07117_PUTP_ECOLI	MAIS	TPMLV	TCVYIF	FGMILL	IGF	LAWRSTK	DDYII	GRRS	GPFFV	TALS	SAGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA				
sp_P10502_PUTP_SALTY	MAIS	TPMLV	TCVYIF	FGMILL	IGF	LAWRSTK	DDYII	GRRS	GPFFV	TALS	SAGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA				
tr_O7C108_PUTP_YERPE	MTMN	TPMLV	TCVYIF	FGMILL	IGF	LAWRSTK	DDYII	GRRS	GPFFV	TALS	SAGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA				
tr_O915P5_PUTP_PSEAB	MSVN	TPTLI	TEVYIA	AMVLI	IGL	AAVRSN	NFDDYII	GRRS	GSFV	TALS	SAGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA				
tr_O7DD05_PUTP_NEIME	MNPM	XTLI	TEVYIA	AMVLI	IGL	AAVRSN	NFDDYII	GRRS	GSFV	TALS	SAGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA				
tr_O24896_PUTP_HELPY	MGNY	ATLIT	TEVYIA	AMVLI	IGL	AAVRSN	NFDDYII	GRRS	GSFV	TALS	SAGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA				
tr_A3GXV9_PUTP_VIBCH	MGHV	STPVT	MFI	IVS	LLM	YIGF	YKQNE	EDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA		
tr_O3FWY7_PUTP_STAB8	MENF	ANQIT	FI	IVY	LLM	IAIGF	YKQRTK	NSDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA		
tr_O64933_OPUE_BACSU	MLTM	GTALS	QQVAN	QWYI	IVY	LLM	IAIGF	YKQRTK	NSDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA	
tr_O79VH1_PUTP_CORGL	MSIE	II	SLGH	FI	IVY	LLM	IAIGF	YKQRTK	NSDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA	
tr_O59086_PUTP_PYRAB	MSDN	TW	FI	IVY	LLM	IAIGF	YKQRTK	NSDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA		
tr_O592P3_PUTP_PYRAB	MNGL	GILG	FL	IVY	LLM	IAIGF	YKQRTK	NSDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA		
sp_P16256_PANF_ECOLI	MSNA	YEM	KLNP	VA	FL	IVY	LLM	IAIGF	YKQRTK	NSDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA
tr_O8D1P2_PANF_YERPE	MQ	TDV	PL	IVY	LLM	IAIGF	YKQRTK	NSDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA		
sp_P14963_PANF_HAEIN	MNGL	GILG	FL	IVY	LLM	IAIGF	YKQRTK	NSDXEL	GR	MPV	VS	SALS	AGS	DMSG	WLL	MG	LP	GA	VFL	LS	ES	WIA	IG	LT	IG	AW	INK	LVA		
sp_P13866_SGLT1_HUMAN	MS	TS	WSP	K	IT	AE	TP	PE	PL	SR	NA	AD	LS	IV	Y	LV	V	NA	V	NA	V	NA	V	NA	V	NA	V	NA		
sp_P09NY91_SGLT1_HUMAN	MA	ST	VS	PT	AE	TP	PE	PL	SR	NA	AD	LS	IV	Y	LV	V	NA	V	NA											
sp_P316239_SGLT2_HUMAN	ME	HT	EA	S	AP	EM	G				
sp_P31289_S156_HUMAN	MS	VG	VS	AP	LS	PT	SG	TS	V	M				
sp_P092911_NIS_HUMAN	MS	NE	H	GL	S	F	I	D	M	V				
tr_O27894_PUTP_VIBBA	MS	NI	E	H	G	L	S	F	I	D	M	V				
tr_O9HP63_PANF_HAEIN	MA	Q	V	E	A	P				
tr_O29293_PUTP_ARCFU	MA	Q	V	E	A	P				
tr_O29063_PANF_ARCFU	MA	Q	V	E	A	P				
sp_P33413_DUR3_YEAST	MG	E	K	P	L	P	Q	G	A	Y	A	V	L	G	G	F	A	G	M	V	L	T	T	T	T	T	T			
consensus>60					

5.9. References

- Jung, H., Rübenhagen, R., Tebbe, S., Leifker, K., Tholema, N., Quick, M., and Schmid, R.** (1998) Topology of the Na⁺/proline transporter of *Escherichia coli*. *J. Biol. Chem.* 273: 26400-26407.
- Pearson, W. R. and Lipman, D. J.** (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U. S. A.* 85: 2444-2448.
- Pearson, W. R.** (1990) Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods. Enzymol.* 183: 63-98.

CHAPTER 6***Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na⁺/H⁺ antiporter of E. coli*****6.1. Abstract**

The pH dependence of the structure of the main Na⁺/H⁺ antiporter NhaA of *Escherichia coli* is studied by continuous-wave (CW) and pulse electron paramagnetic resonance (EPR) techniques on singly spin-labeled mutants. Residues 225 and 254 were selected for site-directed spin labeling, as previous work suggested that they are situated in domains undergoing pH-dependent structural changes. A well-defined distance of 4.4 nm between residues H225R1 in neighboring molecules is detected by a modulation in double electron-electron resonance data. This indicates that NhaA exists as a dimer, as previously suggested by a low-resolution electron density map and cross-linking experiments. The modulation depth decreases reversibly when pH is decreased from 8.0 to 5.8. A quantitative analysis suggests a dimerization equilibrium, which depends moderately on pH. Furthermore, the mobility and polarity of the environment of a spin label attached to residue 225 change only slightly with changing pH, while no other changes are detected by CW EPR. As antiporter activity of NhaA changes drastically in the studied pH range, residues 225 and 254 are probably located not in the sensor or ion-translocation sites themselves but in domains that convey the signal from the pH sensor to the translocation site.

6.2. Introduction

A better understanding of the formation of well-defined homo- and heterooligomers as well as larger aggregates of membrane proteins could enhance our insight into regulatory processes (Veenhoff *et al.*, 2002) and lead to improved strategies for membrane protein crystallization (Hunte and Michel, 2000). Such understanding is currently hindered by a relative lack of characterization techniques that can quantify partial aggregation and that give detailed structural information on the aggregates. A combination of site-directed spin labeling (for reviews see, e.g., Hubbell and Altenbach, 1994; Hubbell, *et al.* 2000; Steinhoff, 2002) with the measurement of small magnetic dipole-dipole interactions between electron spins at distances of 2-6 nm by pulse electron paramagnetic resonance (EPR) (Milov *et al.*, 1998; Borbat and Freed, 2000; Jeschke, 2002) should be well suited for studies of protein aggregates as it can characterize local structure without depending on long-range order. We have recently demonstrated that the four-pulse double electron-electron resonance (DEER) experiment (Pannier *et al.*, 2000) can provide structural information on integral membrane

proteins (Jeschke *et al.*, 2004a). Combined with triangulation techniques (Borbat *et al.*, 2002), the experiment was used to assess structural differences between monomers and trimers of plant light harvesting complex IIb (Jeschke *et al.*, 2005) - a case where both the monomer and trimer are stable in detergent micelles and can be separated from each other. In many cases, however, oligomerization of membrane proteins has to be treated as a thermodynamic equilibrium, and the dependence of that equilibrium on external parameters is of interest. This situation implies that not only the structure of the oligomer but also the average number of protein molecules in the oligomer have to be determined.

Pulse electron double resonance techniques such as four-pulse DEER are suited for this purpose since the modulation depth in the signal depends on the number of remote spins coupled to the observer spin (Milov *et al.*, 1984). We have recently given the corresponding equations for the case where spin-to-spin distances are distributed and a contribution from spins in neighboring oligomers or clusters has to be considered (Jeschke *et al.*, 2004b). In this study, we estimate the precision of determining the average number of coupled spins by measurements on a number of authentic biradicals and a triradical (Godt *et al.*, 2000) and apply the approach to pH-dependent oligomerization of the sodium/proton antiporter NhaA of *Escherichia coli*.

NhaA is the main one of two antiporters of *E. coli* that accomplishes specific exchange of Na⁺ or Li⁺ for H⁺ (Schuldiner and Padan, 1993). This electrogenic antiporter, which exhibits a stoichiometry of 2H⁺/Na⁺ (Taglicht *et al.*, 1993), is indispensable for ion homeostasis in *E. coli* and has homologues in many other enterobacteria (Padan *et al.*, 2001). Among all Na⁺/H⁺ exchangers, NhaA of *E. coli* is the best characterized. Two-dimensional crystals diffracting at 4 Å resolution could be obtained, and analysis of electron density projection maps (Williams *et al.*, 1999) and of a reconstructed three-dimensional map (Williams, 2000) revealed that NhaA forms dimers in crystals, with each monomer containing 12 tilted transmembrane helices. Coexpression of pairs of conditional lethal mutants demonstrated functional complementation, which, together with cross-linking data, strongly suggests that NhaA also functions as a dimer (Gerchman *et al.*, 2001). Like many other Na⁺/H⁺ antiporters, NhaA is regulated by pH (Padan *et al.*, 2001; Padan *et al.*, 2004). It is essentially inactive below pH 7, and the Na⁺ efflux rate increases by three orders of magnitude between pH 6.5 and 8.5 (Taglicht *et al.*, 1991). Cross-linking between loops 9, which are located in the interface between the individual NhaA molecules in the dimer, causes a dramatic change in the pH response for a short and rigid linker and no effect for a long and flexible linker (Gerchman *et al.*, 2001). This suggests that a functional interaction between NhaA molecules is involved in the pH response.

From a mechanistic point of view, the pH response could be explained by pH-dependent dimerization, by a pH-dependent change of the structure in the interface between

the two molecules (monomers), or by a conformational change within one of the two monomers that depends on the presence of certain residues in the nearby second monomer. By measuring the degree of dimerization as a function of pH and distances between residues in different monomers, we attempt to decide between these possible mechanisms. We concentrate on two residues that appear to be involved in the pH response. The first of these residues, His225, is located at the periplasmic end of transmembrane domain VIII. The mutation H225R was shown to change, but not eliminate, pH dependence of NhaA antiporter activity (Gerchman *et al.*, 1993). The second residue, Val254 in loop 9, was selected as its mutation to Cys causes an acidic shift of the pH response. Furthermore, nearby Lys249 in the same loop is a site of pH-dependent proteolytic cleavage by trypsin (Gerchman *et al.*, 1999). Since susceptibility to this cleavage increases in the pH range of activation of NhaA, one may assume that loop 9 becomes more accessible to trypsin by a conformational change that is related to the pH response of the antiporter.

This article is organized as follows: First, continuous-wave (CW) EPR is used to study possible pH-dependent changes of these residues in their mobility and accessibility to water- and lipid-soluble paramagnetic quenchers. Second, calibration of measurements of the average number of spins within the same oligomer or aggregate is discussed briefly. Third, four-pulse DEER is applied to test for pH-dependent changes in the distances H225R1/H225R1' and V254R1/V254R1' in dimers and in the degree of dimerization. Finally, the results are discussed in terms of the most likely mechanism of pH response.

6.3. Materials and Methods

6.3.1. Preparation and labeling of NhaA

The generation of *nhaA* alleles encoding single-Cys NhaA molecules used in this study has already been described (Olami *et al.*, 1997; Gerchman *et al.*, 2001). For overexpression the resulting single-Cys NhaA encoding plasmids were transformed into *E. coli* TA16. TA16 is *nhaA*⁺ *nhaB*⁺ *lacIQ* and otherwise isogenic to TA15 (Taglicht *et al.*, 1991). Cells were grown, membranes were prepared, and NhaA was solubilized and purified by Ni-nitrilotriacetic acid affinity chromatography as described (Rimon *et al.*, 1998). Before elution of the single-Cys NhaA molecules (in 20 mM Tris-HCl, pH 7.9, containing 500 mM NaCl, 300 mM imidazole, 10% glycerol (v/v), 0.02% β -D-dodecylmaltoside (w/v)) the protein was reacted with MTSSL (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate (Toronto Research Chemicals, Toronto). For this purpose, washing buffer (20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5 mM imidazole, 10% glycerol (v/v), 0.1% β -D-dodecylmaltoside (w/v)) containing 1 mM MTSSL was applied to the column followed by overnight incubation at 4°C (hereafter we will refer to the spin-labeled residue as R1). Afterwards, unbound label was

removed by washing the column with washing buffer. The bond of MTSSL to the protein was stable over the few hours required for sample preparation, and measurements throughout the pH range from 5.8 to 8.0. No free label was observed in the CW EPR spectra. After elution the protein was reconstituted into proteoliposomes with a composition of 67% phosphatidylethanolamine, 23.2% phosphatidylglycerol, and 9.8% cardiolipin at a lipid/protein ratio of 20:1 (w/w) as described (Jung *et al.*, 1998). Finally, the proteoliposomes were washed twice with 50 mM KP_i , pH 7.5, and resuspended in 10 mM KP_i , pH 5.8 or 8.0, 100 mM KCl and 5 mM $MgCl_2$ to yield a NhaA concentration of 50-100 μM . Proteoliposomes were frozen and stored in liquid nitrogen until use.

6.3.2. Four-pulse DEER measurements

Four-pulse DEER measurements were performed with a Bruker Elexsys 580 spectrometer (Bruker Biospin GmbH, Karlsruhe, Germany) equipped with a 3 mm split-ring resonator under conditions of strong overcoupling ($Q \approx 100$) at a temperature of 50 K. Before insertion into the probehead, the samples were shock-frozen in liquid nitrogen to avoid crystallization of water. The pulse sequence $(\pi/2)_{\nu_1} - \tau_1 - (\pi)_{\nu_1} - t' - (\pi)_{\nu_2} - \tau_1 + \tau_2 - t' - (\pi)_{\nu_1} - \tau_2$ -echo was used. The $\pi/2$ - and π -pulses at the observer frequency ν_1 had equal pulse lengths of 32 ns to ensure equal excitation bandwidths, while the π -pulse at the pump frequency ν_2 had a length of 12 ns to maximize modulation depth while still keeping the two excitation bands separate. These conditions maximize sensitivity of the experiment (Jeschke *et al.*, 2004c). The long interpulse delay was $\tau_2 = 2000$ ns with an initial value of $t' = 80$ ns and an increment $\Delta t' = 8$ ns. To suppress proton modulation, data were added for eight equidistant values of τ_1 between 200 and 256 ns. A phase cycle $[+(+x), -(-x)]$ was applied to the first pulse, and the two signals were subtracted. The pump frequency ν_2 (typically 9.33 GHz) was set to the center of the resonator mode and to coincide with the global maximum of the nitroxide spectrum. The observer frequency ν_1 was set to the local maximum at the low-field edge of the spectrum ($\nu_1 - \nu_2 = 65$ MHz). Accumulation times for the data sets varied between 8 and 14 h. Data were analyzed for dipolar evolution times $t = t' - \tau_1 \geq 0$.

6.3.3. Four-pulse DEER data analysis

Distance distributions and average numbers of spins were obtained from the dipolar time evolution data by the home-written Matlab (The MathWorks, Natick, MA) program package DeerAnalysis 2004 (freely available at <http://www.mpip-mainz.mpg.de/~jeschke/distance.html>), which is based on algorithms explained in Jeschke *et*

al. (2004b) and Weese (1992). The background contribution from spin labels in other dimers was fitted either by a three-dimensional distribution (spatial background correction) corresponding to a functional dependence $\exp(-kt)$ in time domain or by a two-dimensional distribution (planar background correction) corresponding to $\exp(-kt^{2/3})$. Measurements on single mutants of a monomeric protein had shown that for membrane proteins inserted into liposomes, such a two-dimensional distribution fits data significantly better than a three-dimensional one. Realistic resolution of the distance distributions was estimated from the dependence of the root mean-square deviation of simulated from experimental data for Hermite interpolation with different numbers of sampling points. The smallest number of sampling points for which there was no significant deviation of fitted from experimental data was selected. Final distance distributions $P(r)$ were then obtained by Tikhonov regularization in distance domain, incorporating the constraint $P(r) > 0$. The regularization parameter was adjusted to obtain the realistic resolution determined before. Modulation depths were determined from the level of the fitted dipolar evolution function. As fitting is performed for background-corrected data, these values pertain to the average number of spins in the same NhaA molecule or dimer, excluding spins in remote molecules or dimers.

6.3.4. CW EPR measurements

CW EPR spectra were recorded with home made X-Band EPR spectrometers equipped with a dielectric cavity (Bruker, Ettlingen, Germany) for room-temperature ($T \approx 293$ K) measurements or with an AEG TE-103 cavity for low-temperature ($T \approx 170$ K) measurements. The magnetic field was monitored using a Bruker B-NM 12 B-field meter; microwave power was determined using a HP 430C bolometer. The NhaA samples were loaded into quartz capillaries with an inner diameter of 1 mm. At room temperature, 10 scans with a scan width of 12.8 mT (microwave power <0.6 mW, modulation amplitude <0.33 mT) were averaged. At 170 K, 10 or 20 scans (12.8 mT scan width, microwave power <0.4 mW, modulation amplitude 0.23 mT) were accumulated.

For power saturation experiments in the presence of oxygen or chromium oxalate (CROX), the samples were loaded into gas-permeable TPX capillaries. The samples were deoxygenated by passing nitrogen around the sample capillary. For oxygen accessibility experiments, nitrogen was replaced by air. Saturation curves were determined from the peak-to-peak amplitudes of the center line measured at seven different incident microwave power levels in the range from 0.5 to 80 mW. The saturation behavior of the samples was parametrized by the quantity $P_{1/2}$, which is defined as the power level of the incident radiation at which the amplitude of the saturated line is half of the amplitude in the absence of saturation. Values for this parameter were calculated from fitting the function

$$A(P) = I\sqrt{P}\left[1 + (2^{1/\varepsilon} - 1)P/P_{1/2}\right]^{-\varepsilon} \quad (6.1.)$$

to the experimental amplitudes $A(P)$ according to the method of Altenbach *et al.* (1994). The scaling factor I and the measure of the saturation homogeneity ε were adjustable parameters. The quantity $\Delta P_{1/2}$ was calculated from the difference in $P_{1/2}$ values in the presence and absence of the relaxing agent. The $\Delta P_{1/2}$ values were divided by the peak-peak line width and normalized by the same quantity of a 2,2-diphenyl-1-picrylhydrazyl (DPPH) standard sample to obtain the dimensionless accessibility parameter II .

6.3.5. Simulation of CW EPR spectra

CW EPR powder spectra were simulated based on the method described previously (Steinhoff *et al.*, 1997). The motion of proteins is strongly restricted for temperatures <200 K. Below that temperature the apparent hyperfine splitting was found to deviate no more than 2% from the rigid limit value determined at 80 K, and the EPR line shape resembles that of a powder spectrum. The values of the g - and hyperfine-tensor components were taken from a sample for which no dipolar broadening is expected: $g_{xx} = 2.0084$, $g_{yy} = 2.0065$, $g_{zz} = 2.0026$, $A_{xx} = 0.51$ mT, $A_{yy} = 0.40$ mT. To account for unresolved hyperfine interactions and relaxational broadening, the calculated powder stick spectrum was convoluted by a field-independent line-shape function (48% Lorentzian and 52% Gaussian with widths of 0.351 and 0.387 mT, respectively). Dipolar interaction is considered by additional convolution with a Pake pattern (Steinhoff *et al.*, 1997; Radzwill *et al.*, 2001). The width of the distance distribution was fixed to 0.2 nm.

6.4. Results

6.4.1. Local mobility-CW EPR spectra at room temperature

The room-temperature spectra ($T = 293$ K) at pH 5.8 of the singly spin-labeled NhaA mutants are depicted in Fig. 6.1.. The spectrum of V254R1 (*dotted line*) indicates a more restricted motion of the bound spin label compared to H225R1 (*solid line*). This is revealed by a shift of the low field peak of V254R1 from ~ 342.0 mT to 340 mT and a reduced reciprocal line width (ΔH_0^{-1}) of $\sim 0.16 \times 10^{-4} \text{ T}^{-1}$ compared to $0.22 \times 10^{-4} \text{ T}^{-1}$ at position 225. These values indicate an exposed surface position of the nitroxide at position 225 and a more buried location of the nitroxide at position 254.

We recorded the room temperature EPR spectra of H225R1 and V254R1 at two additional pH values to determine whether changes of pH cause conformational changes. In contrast to the spectra of the V254R1 (Fig. 6.2.b), the data of H225R1 reveal slight but

significant alterations of the mobility of the spin label side chain (Fig. 6.2.a). The nitroxide becomes successively more immobile as the pH is shifted from 5.8 to 7.0 (*dashed line*) and 8.0 (*dotted line*). This is represented by a peak shift toward 342.0 mT and a corresponding slightly more pronounced dent in the high field region between 348.5 mT and 349.0 mT.

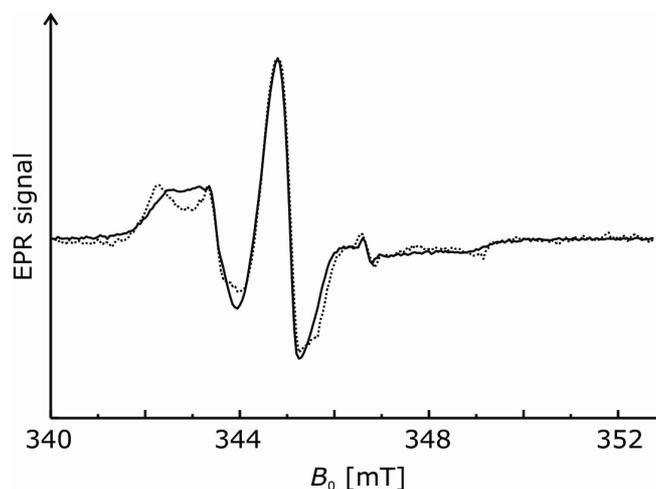


Figure 6.1.: Comparison of normalized room temperature CW EPR spectra ($T \approx 293$ K) of mutants H225R1 (solid line) and V254R1 (dotted line) recorded at pH 5.8. Differences in the S/N are due to different concentrations of the samples.

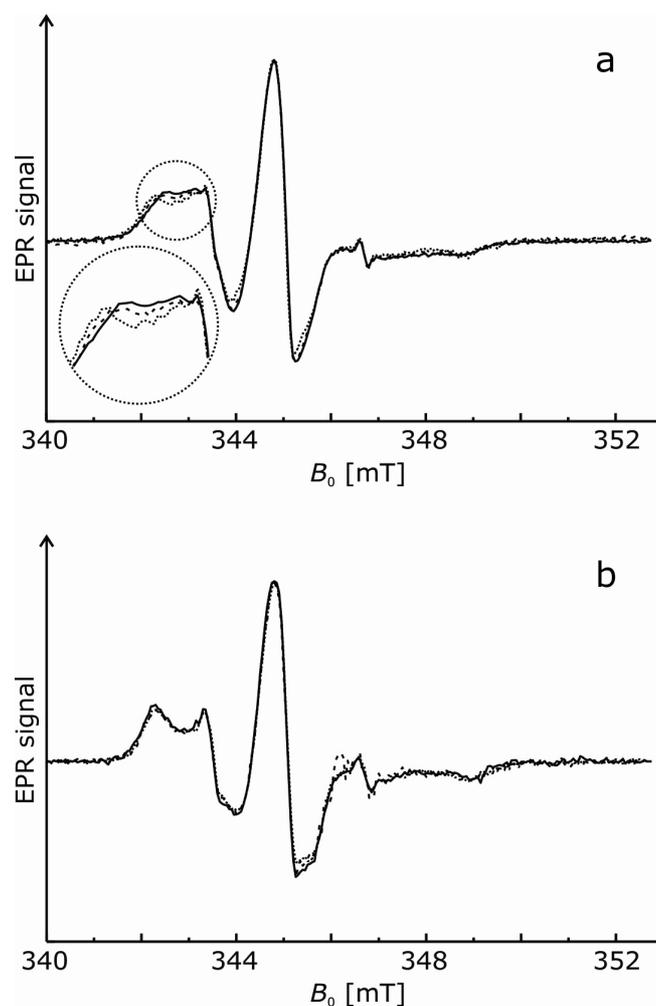


Figure 6.2.: Effect of pH on CW EPR spectra of mutants H225R1 and V254R1 ($T \approx 293$ K). Normalized spectra are shown for pH 5.8 (solid lines), pH 7.0 (dashed lines), and pH 8.0 (dotted lines). (a) H225R1. (b) V254R1.

6.4.2. Polarity- and proximity-CW EPR spectra at low temperature

The existence of a pH-induced conformational change affecting position 225 is confirmed by low temperature EPR spectroscopy ($T = 170$ K). Increasing the pH from 5.8 to 8.0 causes a shift of the low field peak minimum to lower field strengths and, consequently, an increase of the A_{zz} value by ~ 0.5 mT. This is evidence for a movement of the R1 side chain into a more polar environment (Fig. 6.3.). In contrast, the change of pH does not affect the A_{zz} value determined for V254R1, which indicates that the environmental polarity of the nitroxide attached to position 254 is not influenced by pH (data not shown).

Low temperature EPR spectra of H225R1 and V254R1 were also analysed for possible dipole-dipole interactions, which would be manifested in CW EPR spectra if dimerization leads to interspin distances H225R1/H225R1' or V254R1/V254R1' shorter than 2 nm. However, none of the low-temperature spectra of the singly spin-labeled NhaA mutants exhibits significant dipolar line broadening. Best fits of the spectra according to the method described by Steinhoff *et al.* (1997) are obtained with interspin distance values equal to or exceeding (2.0 ± 0.1) nm (data not shown), which corresponds to the upper detection limit of the CW EPR technique. Therefore, we conclude that there is no significant fraction of NhaA with interspin distance values of H225R1 and V254R1 shorter than 1.8 nm. No evidence for dimerization is obtained by CW EPR.

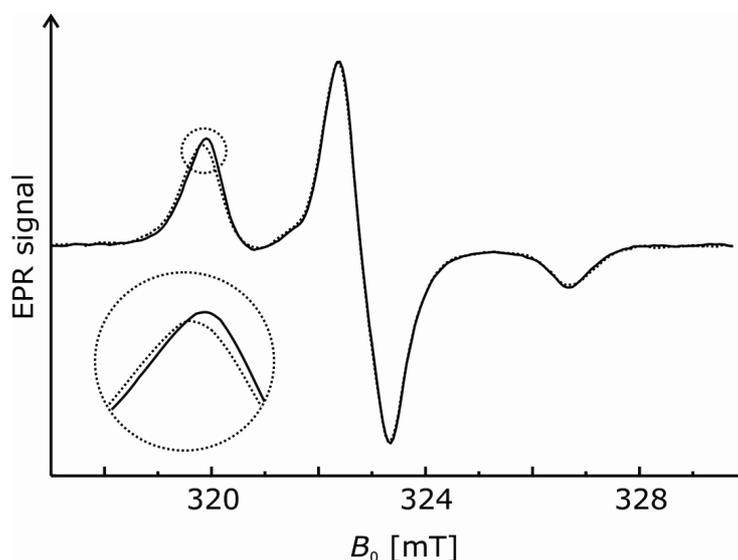


Figure 6.3: Effect of pH on the low temperature CW EPR spectrum ($T = 170$ K) of NhaA mutant H225R1. Normalized spectra are shown for pH 5.8 (solid lines) and pH 8.0 (dotted lines).

6.4.3. Accessibility measurements by CW EPR

To obtain semi-quantitative information on the location of residues H225R1 and V254R1 relative to the membrane, we use power saturation of CW EPR spectra to determine the accessibility parameter \mathcal{I} of the nitroxide side chains for freely diffusing paramagnetic

quenchers (Altenbach *et al.*, 1990; Altenbach *et al.*, 1994). Molecular oxygen and water-soluble chromium oxalate (CROX) are suitable quenchers because of their sizes and solubility properties. Molecular oxygen has a low concentration in the tightly packed protein interior, whereas both the solubility and the diffusion coefficient are high in the lipid bilayer. Therefore, for spin label side chains located in the protein interior, the collision frequency with oxygen and the accessibility parameter Π_{Oxygen} are low, whereas these are high if the spin label faces the bilayer. Likewise, measurement of the collision frequency with the water-soluble CROX allows determination of the location of the R1 side chain with respect to the aqueous phase (Altenbach *et al.*, 1994; Anthony-Cahill *et al.*, 1992). In principle, it has to be taken into account that the collision frequency of the R1 side chain with the negatively charged CROX is a function of both accessibility and local electrostatic potential. However, at the relatively high salt concentrations used here the electrostatic potential is strongly screened, so that accessibility has the dominating influence on the collision frequency.

Table 6.1.: Accessibility of nitroxide groups in NhaA mutants H225R1 and V254R1

Mutant	pH 5.8		pH 8.0	
	Π_{Oxygen}	Π_{CROX}	Π_{Oxygen}	Π_{CROX}
NhaA-H225R1	0.44	0.02	0.48	0.04
NhaA-V254R1	0.40	0.01	0.27	0.02

Π , orientation of the spin-labeled side chains with respect to the protein, lipid, or aqueous phase (see text) normalized by the quantities of a DPPH standard sample ($\Pi = (P_{1/2}/I_w - P_{1/2}(N_2)/I_w(N_2)) \times I_w(\text{DPPH})/P_{1/2}(\text{DPPH})$). $P_{1/2}$, power level of the incident radiation at which the amplitude of the saturated line is half of the amplitude in absence of saturation.

I_w , line width.

Π_{Oxygen} , measured in equilibrium with air and scaled to pure oxygen.

Π_{CROX} , measured in the presence of 50 mM CROX.

The results of the accessibility analysis are displayed in Table 6.1.. The low Π_{CROX} values of 0.01 and 0.02 at pH 5.8 imply low collision frequencies with CROX for nitroxides at both positions 225 and 254. This suggests that both nitroxide groups are either facing the phospholipid bilayer or are buried within the protein monomer or dimer. Also, for both residues, Π_{Oxygen} values of ~ 0.4 indicate intermediate oxygen accessibility at pH 5.8 suggesting that the R1 side chain is in contact with the apolar environment of the bilayer. Increasing pH from 5.8 to 8.0 does not significantly affect the accessibility of both nitroxides.

6.4.4. Calibration of spin counting by DEER measurements

For sufficiently dilute samples, primary DEER data can be separated into contributions due to spins within the same nanoobject as the observer spin and due to spins in other nanoobjects. This requires that the size of the nanoobject, in our case a membrane protein monomer or oligomer, must be within the distance range accessible to DEER. For membrane proteins, we find that this range extends from 1.75 to 6 nm. In this situation, the modulation depth Δ for the contribution from spins within the same nanoobject depends on

the average number $\langle n \rangle$ of spins in that nanoobject as

$$\langle n \rangle = 1 + \frac{\log(1 - \Delta)}{C}, \quad (6.2.)$$

where the parameter C depends on the excitation profile of the pump pulse and spectral line shape of the nitroxide (Milov *et al.*, 1984; Jeschke *et al.*, 2004b). For measurements performed with the same microwave resonator and pulses of the same lengths on nitroxides with similar A_{zz} values, C can be considered a constant. For calibration of that constant across the distance range of interest, we used dilutions in *o*-terphenyl of the rigid biradicals **3** from Jeschke *et al.* (2002) with an average distance $\langle r \rangle = 1.96$ nm and **1** ($\langle r \rangle = 2.84$ nm), **2a** ($\langle r \rangle = 3.65$ nm), and **2b** ($\langle r \rangle = 5.02$ nm) from Godt *et al.* (2000), the flexible biradical **4** ($\langle r \rangle = 1.96$ nm) from Jeschke *et al.* (2002) and the rigid symmetric triradical **3** ($\langle r \rangle = 3.40$ nm) from Godt *et al.* (2000). We find a correlation coefficient of 0.9665. The calibration was tested by independent measurements on rigid biradicals **2a** and **2b** (Godt *et al.*, 2000) and the ultralong biradical shown in Fig. 4a of Jeschke *et al.* (2004c) ($\langle r \rangle = 7.6$ nm), all diluted in perdeuterated *o*-terphenyl. In these test measurements, we find $\langle n \rangle = 2.08$, 1.99, and 2.03, respectively, in good agreement with the expectation of $\langle n \rangle = 2$.

6.4.5. Evidence for specific oligomerization by four-pulse DEER

Pulse EPR distance measurements using the four-pulse DEER sequence were performed for both mutants, H225R1 and V254R1. Primary time-domain DEER data at pH 8.0 are compared in Fig. 6.4.. For a homogeneous distribution of the spin labels in three dimensions or in a lipid membrane, one would expect a smoothly decaying signal as indicated by the dashed lines. Clearly, the experimental data contain additional contributions at short times. These contributions correspond to label-to-label distances within a nanoobject composed of several NhaA molecules that must be smaller than the upper limit of DEER distance measurements of 6 nm.

To quantify the label-to-label distance distribution, one has to separate the contribution of labels within the same nanoobject from the contribution of labels in neighboring objects. The latter contribution corresponds to the smooth decay and is of the form $\exp(-kt)$ for a homogeneous spatial distribution of spins in three dimensions and of the form $\exp(-kt^{2/3})$ for a homogeneous planar distribution as expected for nanoobjects distributed in a membrane. For singly labeled mutants of Na⁺/proline transporter PutP in liposomes, we have found that the decay conforms to a planar distribution (Hilger, D., Jung, H., Steinhoff, H.-J., and Jeschke, G., unpublished), while singly labeled mutants of plant light-harvesting complex IIb monomers solubilized in detergent micelles conform to a spatial distribution (Jeschke *et al.*,

2005). In the case at hand, a planar distribution is more likely but cannot be verified independently by control experiments. We therefore analyze our data for the cases of both planar and spatial distributions (Fig. 6.5.).

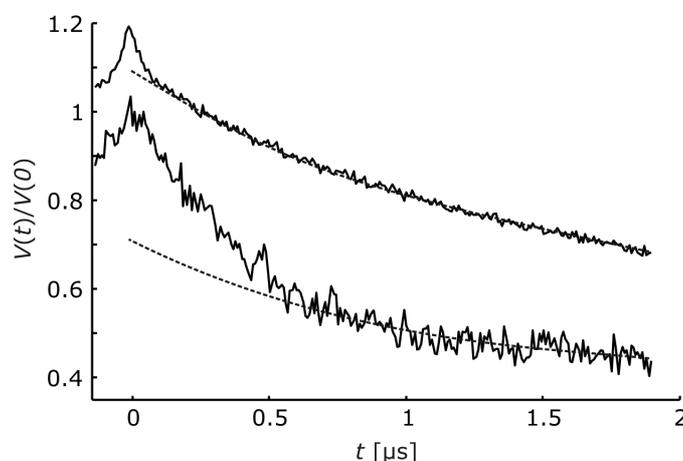


Figure 6.4.: Normalized DEER time-domain data for mutants H225R1 (*bottom*) and V254R1 (*top*, shifted by 0.2) at pH 8. The dashed lines indicate smooth decaying signals as expected for a homogeneous distribution.

For H225R1 the distance distribution does not strongly depend on the choice of the background distribution. Less noise-related baseline artifacts are observed with the planar distribution (Fig. 6.5.a) compared to the spatial distribution (Fig. 6.5.b), which suggests that the planar distribution is the better model. The dominant peak in the distribution is at (4.36 ± 0.04) nm and has a width of ~ 0.3 nm, as found by fitting the data with a distance distribution consisting of a single Gaussian peak. The occurrence of one relatively narrow peak strongly suggests formation of an oligomer with well-defined structure. This oligomer cannot be composed of more than four NhaA molecules, as this would imply more than one peak in the distance distribution H225R1/H225R1'. For a tetramer a single peak would only be expected if it had tetrahedral symmetry, which can be excluded for an integral membrane protein. The distance distribution is thus consistent only with a dimer or trimer.

For V254R1, the reliability of the distance distribution crucially depends on the proper choice of background function. A single peak at 2.00 ± 0.04 nm with a width of 0.20 ± 0.10 nm is found under the assumption of a planar distribution of aggregates (Fig. 6.5.c), whereas a broad distribution covering the whole range between at least 1.5 and 5 nm is found when assuming a three-dimensional spatial distribution (Fig. 6.5.d). We have checked that the result for the planar distribution is self-consistent: For the distance distribution in Fig. 6.5.c the time-domain data at $t > 600$ ns are exclusively due to labels in neighboring nanoobjects. The dimensionality D can thus be determined independently by fitting a function $\exp(-kt^{D/3})$ to the decay at $t > 600$ ns. We find $D = 2.009$ in good agreement with the expectation of $D = 2$ for a planar distribution. Furthermore, as helix IX is thought to extend up to residue Pro257 (Rothman *et al.*, 1996) one would hardly expect a very broad distance distribution for residue V254R1 as seen in Fig. 6.5.d, even if one accounts for the conformational freedom of the

MTSSL label. Finally the peak at 2 nm persists even with the three-dimensional background correction and is also found at pH 5.8 and 7.0 (data not shown). We thus assume that the distance distributions for H225R1 and V254R1 are both monomodal. In any case it is safe to conclude that in the observed oligomer of NhaA, labeled residues H225R1 are separated by more than 4 nm and there is at least a population of oligomers with a separation of <2.5 nm between labeled residues V254R1.

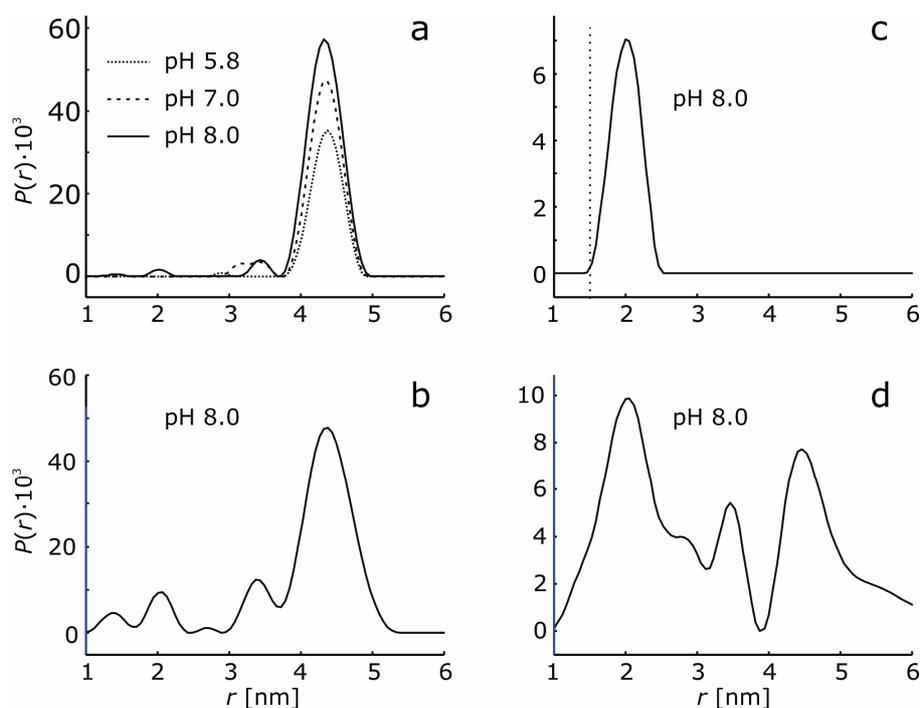


Figure 6.5.: DEER distance distributions $P(r)$ obtained from the time-domain data in Fig. 6.4. by Tikhonov regularization in the range from 1 to 8 nm with a regularization parameter of 100. (a) H225R1 with planar background correction. (b) H225R1 with spatial background correction. (c) V254R1 with planar background correction. The dotted line marks the lower limit of distances that are detectable by DEER with our setup. (d) V254R1 with spatial background correction.

The probabilities to find a second labeled NhaA molecule at a distance within the DEER range from 1.75 to 6 nm can be compared for H225R1 and V254R1 by comparing the vertical scales of Fig. 6.5.a and c, or, more directly, the deviations from the smooth dashed line in Fig. 6.4.. Clearly, this probability is lower for V254R1. This result could be explained by assuming that the mutation at position 225 and the attachment of the MTSSL label increase the tendency for oligomerization, or conversely, mutation and attachment at position 254 decrease this tendency. However, our experience with measurements of short distances in the Na^+ /proline transporter (Jeschke *et al.*, 2004a) and labeled peptides (J. Banham, G. Jeschke, and C. Timmel, unpublished) suggests that the low modulation depth for V254R1 could also be due to suppression of part of the distance distribution at the lower edge of the DEER distance range (see *dotted line* in Fig. 6.5.c), even if significant dipolar broadening cannot be detected in the CW EPR spectra. However, it cannot be excluded that the labeling efficiency at this position is less than that at position 225. Because of this ambiguity and the

stronger dependency of data analysis on the choice of background function at position 254, we restrict further discussion of the pH dependence to the mutant H225R1.

Conformational changes that lead to a movement of the periplasmic end of helix VIII should manifest in a peak shift in the distance distribution for residue H225R1, whereas changes in the degree of oligomerization should result in an amplitude scaling of $P(r)$, as only labels within oligomers contribute to $P(r)$. The data displayed in Fig. 6.5.a, which were obtained on the same sample with increasing pH, rule out any significant peak shift, but point to a sizeable amplitude change. Similar data were obtained on a second sample with decreasing pH (not shown). Quantitative analysis of the modulation depth can thus provide more information on the type and extent of oligomerization.

6.4.6. Dependence of the average degree of oligomerization on pH

As mentioned above, the average number $\langle n \rangle$ of labels per protein oligomer can be obtained from the modulation depth. To obtain precise values, the contribution due to labels in neighboring oligomers has to be removed by a background correction that is as accurate as possible. Ideally, the background is fitted at times where modulation due to labels within the same oligomer has decayed completely. For the distance of 4.36 nm in H225R1 oligomers of NhaA, such complete decay is not observed within experimentally feasible observation times. In this situation the background fit can still be optimized, since we know that for distances larger than the expected size of the nanoobject $P(r)$ must approach zero. For a given background function, this optimization is done by varying the starting time t_0 of the background fit and eliminating the background as well as computing $P(r)$ by approximate Pake transformation (Jeschke *et al.*, 2002) for each choice of t_0 . For all data sets of H225R1 samples, we find optimum values of t_0 between 360 and 376 ns. Further analysis was thus performed with $t_0 = 368$ ns as indicated by the dashed line in Fig. 6.6.a. The best planar background fit by a function $\exp(-kt^{2/3})$ is shown as a solid line. The dipolar evolution function corresponding to only labels within the same oligomer is now obtained by dividing the original data (*dots*) by the background function. The dipolar spectrum obtained by Fourier transformation of the corrected data (Fig. 6.6.b) has the expected shape of a broadened Pake pattern and is well fitted by the dipolar spectrum corresponding to the distance distribution in Fig. 6.5.a.

This high fit quality indicates that we can confidently extract modulation depths Δ (Fig. 6.6.c) and thus reliable estimates of the average number $\langle n \rangle$ of labels per oligomer from our data. Indeed, the data in Fig. 6.6.d demonstrate that the dependence of $\langle n \rangle$ on pH for H225R1 samples of NhaA agrees within the expected error for two different sample

preparations measured with increasing and decreasing pH, respectively. We find that $\langle n \rangle$ increases slightly but significantly upon shifting pH from 5.8 to 8.0. On the other hand, lowering pH from 8.0 to 5.8 decreases $\langle n \rangle$, indicating the pH effect is completely reversible (Fig. 6.6.d).

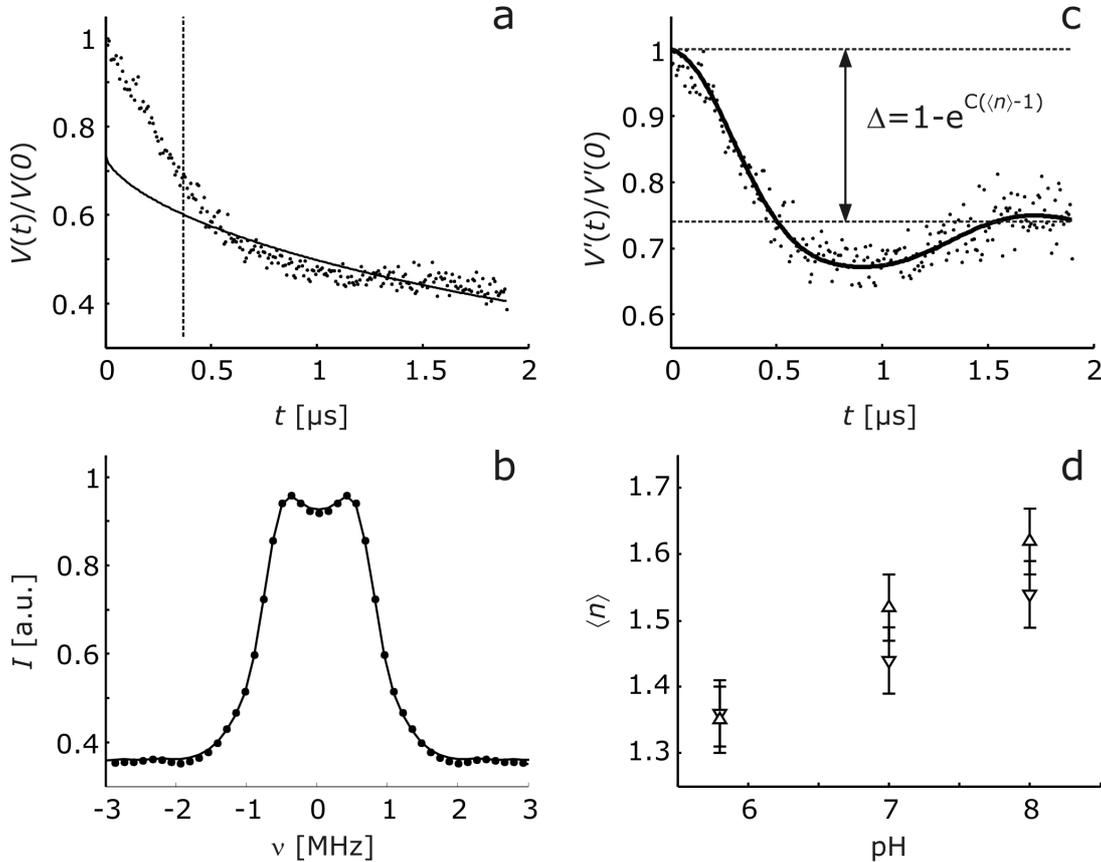


Figure 6.6.: Measurement of the pH dependence of the average oligomerization number $\langle n \rangle$ of the H225R1 spin-labeled mutant of NhaA. (a) Primary data (dots) at pH 8 normalized to the echo intensity at $t = 0$ with planar background fit (solid line). The fit range are times $t > t_0$ with t_0 indicated by the dashed line. (b) Dipolar spectrum obtained by Fourier transformation of background corrected data (dots) and fit by the spectrum corresponding to the distance distribution shown in Fig. 6.5.a. (c) Dipolar evolution function due to labels within the same oligomer (dots) obtained by the background correction and fit corresponding to the distance distribution shown in Fig. 6.5.a. Data are normalized to the background-corrected echo intensity at $t = 0$. The dashed lines and double arrow indicate the total modulation depth Δ . (d) pH dependences of $\langle n \rangle$ obtained in two series of measurements with increasing pH (Δ , bold error bars) and decreasing pH (∇ , slim lined error bars). Error bars are estimated from calibration experiments on a sample set of six chemically pure biradicals and one triradical.

To interpret $\langle n \rangle$ as an average number $\langle n_m \rangle$ of NhaA molecules per oligomer, we have to make an assumption about the fraction f of labeled protein molecules. Incomplete labeling of the protein corresponds to $f < 1$. We find

$$\langle n_m \rangle = 1 + \frac{\log(1 - \Delta)}{fC} = 1 + \frac{\langle n \rangle - 1}{f}. \quad (6.3.)$$

Independently of the value of f , $\langle n \rangle > 1$ implies $\langle n_m \rangle > 1$. Hence, our measurements show that even at the lowest pH value of 5.8, a significant amount of dimer is present. The value of $\langle n \rangle \approx 1.6$ at pH = 8.0 would correspond to complete dimerization for $f = 0.6$. Indeed, we found by determination of the spin concentration and protein concentration of a H225R1 sample after

the DEER measurement that at least 60% of the protein molecules are labeled ($f \geq 0.6$), hence $\langle n_m \rangle \leq 2$. Within experimental error, our results are thus consistent with complete dimerization at pH = 8.0 but do not exclude incomplete dimerization at that pH value. Note that the relative change in $\langle n_m \rangle - 1$ is independent of the uncertainty in f and is nicely reproducible. We can thus safely conclude that for H225R1 mutants, a fraction of oligomers is already present at pH 5.8 ($\langle n_m \rangle - 1 > 0$) and that the degree of oligomerization increases significantly when increasing pH from 5.8 to 8.0.

6.5. Discussion

6.5.1. Comparison with structural models for NhaA dimers

For mutant H225R1, we find that the average number $\langle n \rangle$ of spin labels at distances in the DEER range from 1.75 to 6 nm depends on pH (Fig. 6.6.d). Clearly, a fraction of NhaA undergoes a change in the pH range between 5.8 and 8.0. Furthermore, distances H225R1/H225R1' shorter than 1.8 nm are inconsistent with the CW EPR spectra at any pH. We may thus safely conclude that for a fraction of NhaA, a decrease of pH from 8.0 to 5.8 causes an increase in the distance between H225R1 and H225R1' from 4.36 nm to more than 6 nm. This observation is most easily explained by dissociation of a fraction of NhaA dimers at lower pH, i.e., by a pH-dependent dimerization equilibrium. In principle, our data could also be explained by pH-dependent dissociation of a symmetric trimer. However, as the low-resolution crystal structure (Williams *et al.*, 1999; Williams, 2000) also suggests dimerization, we feel that the latter explanation is unlikely. A symmetric trimer could also hardly be reconciled with the cross-linking experiments (Gerchman *et al.*, 2001), which show that dimers cross-linked between residues H254C are active and, for a long, flexible cross-linker, exhibit essentially the same pH profile of activity as NhaA that is not cross-linked. It seems implausible that such linked dimers could form a symmetric trimer in which residues H254R1 have a distance of only 2 nm. In the following, we therefore discuss our data in terms of a dimer structure.

Structural models of NhaA with different degree of detail have been suggested by Williams (2000) based on electron crystallography, by Ravna *et al.* (2001) based on the model by Williams (2000) and an analogy to LacY permease, and by Rimon *et al.* (2002) based on site-directed thiol cross-linking. The latter two models can essentially be considered as tentative assignments of the helices detected in the crystal structure. For discussion of such assignments, we may note that His225 is the last loop residue before the periplasmic end of helix VIII and Val254 is separated by two residues from the cytoplasmic end of helix IX (Padan *et al.*, 2001). We may thus relate our data and the helix assignments

to the horizontal slices through the electron density of the putative dimer reported by Williams (2000). The slices 0.8 nm 'above' and 'below' the membrane plane are shown in Fig. 6.7.a and b, respectively. It appears likely that the terms 'above' and 'below' the plane have to be assigned to the periplasmic and cytoplasmic side, respectively, as the cross-linking study (Rimon *et al.*, 2002) suggests that it is the periplasmic side where helices are more densely packed. Note that Ravna *et al.* (2001) make the same assignment. In such projection maps, the locations of the spin labels that are consistent with our distance measurements can be visualized as circles centered at the twofold symmetry axis with a diameter corresponding to the measured distance. Given the size of the spin label and the distance of the labeled residues from the helix, the assignment of helix VIII by Ravna *et al.* (2001) might be consistent with our data (label R:VIII in Fig. 6.7.a), but the assignment of helix IX clearly is not (label R:IX in Fig. 6.7.b). Whatever the loop and helix conformations are, it is hardly possible for this assignment to lead to a distance H225R1/H225R1' that is by 2 nm longer than the distance V254R1/V254R1'.

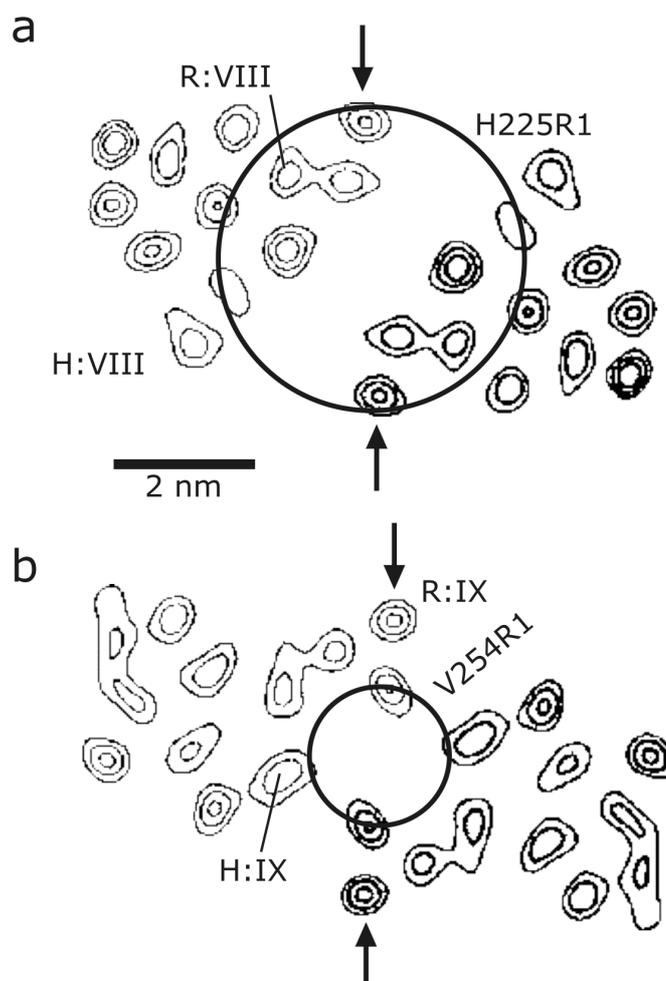


Figure 6.7.: Comparison of intradimer distances between residues H225R1/H225R1' and V254R1/V254R1' to electron density maps of NhaA taken from Williams (Williams, 2000) and to assignments of transmembrane helices VIII and IX in these maps. Arrows pinpoint a helix-helix contact in the electron density maps that was assigned as an interdimer contact by Williams (Williams, 2000). Labels R:VIII and R:IX designate assignments made by Ravna *et al.* (2001), whereas labels H:VIII and H:IX are assignments that are compatible with our measurements and conclusions by Rimon *et al.* (2002). (a) Horizontal slice 0.8 nm above the center of the membrane, assigned here to the periplasmic side. (b) Horizontal slice 0.8 nm below the center of the membrane, assigned here to the cytoplasmic side.

Rimon *et al.* (2002) did not explicitly map their helix assignment to the electron density map and made very clear that it is tentative. However, it seems to be clear from their results that helices VIII and IX reside on the outside rather than in the core of the helix bundle. If we assume that this is indeed the case, our data suggest a unique assignment of these two helices in the electron density map (labeled H:VIII and H:IX in Fig. 6.7.).

We also checked whether the helix assignment of Ravna *et al.* (2001) can be rescued by assuming a different choice of the dimer in the crystal structure. This idea is based on a close 'interdimer' contact between the helices labeled R:IX in Fig. 6.7. (see *arrow*) that was pointed out in the crystal structure by Williams (2000). If this is interpreted as an intradimer contact, one obtains an alternative dimer that interacts only via helices R:IX. In this case, a distance V254R1/V254R1' of 2 nm appears feasible, but the distance H225R1/H225R1' would be expected to be significantly shorter than 4.3 nm. In principle, it might also be possible that the functional dimer existing in a biomembrane or a liposome is different from the dimer formed in crystals since crystallization is performed at pH 4.0, significantly below the range where NhaA is active. However, as our measurements show that the extent of dimerization depends only moderately on pH (Fig. 6.6.d) and distance H225R1/H225R1' does not at all depend on pH, it appears likely that the dimer observed by us is also formed at pH 4.0 and sufficiently high concentrations of NhaA. Furthermore, it is possible that substitution of His225 or Val254 by Cys and the subsequent labeling reaction affect the monomer-dimer equilibrium. Therefore, it can not be excluded that the influence of pH on the oligomeric state of NhaA wild-type in the native membrane is much stronger than observed in our experiments. In any case, our assignment of helices VIII and IX in the electron density map is a prediction that can be tested when a crystal structure with atomistic resolution becomes available. Even if it would turn out to be wrong, the two distances H225R1/H225R1' and V254R1/V254R1' determined by us should allow for identification of the functional dimer from such a structure.

6.5.2. Mechanistic implications of the observed pH dependence

Previous work has demonstrated that the activity of NhaA changes drastically in the studied range between pH 5.8 and 8.0, not only for the wild type (Taglicht *et al.*, 1991) but also for mutants H225R (Gerchman *et al.*, 1993) and V254C (Gerchman *et al.*, 2001). Considering this, the pH dependence of mobility, accessibility, and conformation close to residues H225R1 and V254R1 is surprisingly weak. The CW EPR spectra of mutant H225R1 indicate that residue 225 is slightly less mobile and resides in a slightly more polar environment in the dimer as compared to the monomer. These slight changes suggest that the basic mechanism of pH-dependent activation of NhaA does not depend on this residue. Rather, His225 may serve only to fine tune the pH response of NhaA. In fact, it may not even

be involved in the pH sensor, but may belong to the part of the protein that conveys the signal from the site of the sensor to the site of the ion-translocation pathway.

Note also that the low accessibility of position 225 in the mutant H225R1 to water-soluble CROX is in apparent contradiction to the at least moderate accessibility of this residue to the alkylation reagent *N*-ethylmaleimide in mutant H225C, as this reagent requires ionization of cysteine (Olami *et al.*, 1997). This contradiction could be resolved by assuming that the thiol group of Cys225 is accessible, whereas the nitroxide group of the longer side chain of Cys-MTSSL 225 is buried. Alternatively, the residue might be buried in the protein but accessible to nearby water molecules that help to ionize the thiol group. We may not, however, exclude an alternative explanation based on the different time scales of the two experiments. In a dynamic environment, chemical experiments such as cross-linking or alkylation techniques sense proximity or exposure of residues, even if it is realized only during small fractions of the conformational trajectory, as reaction times are in the range of tens of minutes. EPR experiments, on the other hand, measure the time average of accessibility and, for ergodic systems, also the distribution of distances over the whole trajectory.

Just as for position 225, no significant changes of nitroxide accessibility were detected for V254R1. Furthermore, neither the mobility and polarity analyses nor the DEER measurements of distance V254R1/V254R1' in dimers of NhaA revealed a significant conformational change close to this residue. For the DEER measurement, this failure to detect any change may be partially due to the small modulation depth and relatively poor signal-to-noise ratio in the measurements on mutant V254R1. We can thus neither exclude nor prove a small change of up to 0.3 nm in this distance in the range between pH 5.8 and 8.0. Much larger changes, however, can be excluded. On the other hand, a significant change at residue Lys249 is implied by the strong pH dependence of its accessibility to trypsin (Gerchman *et al.*, 1999). Taken together, these results would suggest that loop 9 moves only slightly or does not move at all close to its point of attachment to helix IX, where residue 254 is located, but undergoes a conformational change of its central part that exposes residue Lys249.

On the basis of the much better DEER data sets for mutant H225R1, we can safely exclude any significant change of distance H225R1/H225R1'. Taken together, these findings indicate that the monomer-monomer interface in the NhaA dimer does not undergo any large-scale changes in the range between pH 5.8 and 8.0.

The increase in the degree of dimerization α of mutant H225R1 with increasing pH (Fig. 6.6.d) is the most significant change that we observed in this study. As $\langle n_m \rangle = 1 + \alpha$, Eq. 6.3. yields

$$\alpha = \frac{\langle n \rangle - 1}{f} . \quad (6.4.)$$

If labeling is complete ($f = 1$), our data would correspond to a change from approximately $\alpha = 0.35$ at pH 5.8 to $\alpha = 0.6$ at pH = 8.0. With $f = 0.6$, which could be assumed as a lower limit of the degree of labeling, we find $\alpha \approx 0.6$ at pH 5.8 and $\alpha \approx 1$ at pH = 8.0. In any case, dimerization is substantial at pH 5.8 where antiporter activity of NhaA mutant H225C is almost absent. It may thus be safe to conclude that dimerization is not a sufficient condition for antiporter activity of NhaA. Dimerization may still be a necessary condition, as is indicated by the fact that coexpression of pairs of conditional lethal mutants may restore activity that is missing in the individual mutants (Gerchman *et al.*, 2001).

The fact that the degree of dimerization and the structure at the monomer-monomer interface change only moderately with pH may suggest that the pH-dependent ion-translocation pathway is not located in this interface. This may also indirectly be deduced from the fact that many residues which are essential for NhaA function or selectivity are located in transmembrane helices IV, V, and XI (Padan *et al.*, 2001; Galili *et al.*, 2002; Galili *et al.*, 2004), which appear to be remote from helix IX (Rimon *et al.*, 2002), and thus from residue 254 and the monomer-monomer interface. It thus appears to be likely that the pH-induced conformational change of loop 9 translates the primary structural change at pH-sensing residues to the ion-translocation pathway through the membrane. In such a picture, dimerization serves to optimize the pH sensor. The pH dependence of the dimerization equilibrium would then be an indirect effect of the localization of pH sensing residues at or close to the monomer-monomer interface but not be directly related to control of the ion flow by pH.

6.6. Acknowledgments

We thank A. Godt and A. Koch for providing the biradicals and triradical for the calibration and C. Bauer for technical assistance. This work was financially supported by the Deutsche Forschungsgemeinschaft (SFB 431-P10 and 431-P18 (HJ and HJS)), Joint Lower saxony, German-Israeli Research Projects (HJ and EP) and the Israel Science Foundation (EP).

6.7. References

- Altenbach, C., Marti, T., Khorana, H. G., and Hubbell, W. L.** (1990) Transmembrane protein structure: spin labeling of bacteriorhodopsin mutants. *Science* 248: 1088-1092.
- Altenbach, C., Greenhalgh, D. A., Khorana, H. G., and Hubbell, W. L.** (1994) A collision gradient method to determine the immersion depth of nitroxides in lipid bilayers: application to spin-labeled mutants of bacteriorhodopsin. *Proc. Natl. Acad. Sci. U. S. A.* 91: 1667-1671.

- Anthony-Cahill, S. J., Benfield, P. A., Fairman, R., Wasserman, Z. R., Brenner, S. L., Stafford 3rd, W. F., Altenbach, C., Hubbell, W. L., and DeGrado, W. F.** (1992) Molecular characterization of helix-loop-helix peptides. *Science* 255: 979-983.
- Borbat, P. P., and Freed, J. H.** (2000) Double quantum ESR and distance measurements. In *Distance measurements in biological systems by EPR*. (Berliner, L., Eaton, S. S., and Eaton, G. R., eds.) pp. 383-459, Kluwer, New York.
- Borbat, P. P., Mchaourab, H. S., and Freed, J. H.** (2002) Protein structure determination using long-distance constraints from double-quantum coherence ESR: study of T4 lysozyme. *J. Am. Chem. Soc.* 124: 5304-5314.
- Galili, L., Rothman, A., Kozachkov, L., Rimon, A., and Padan, E.** (2002) Transmembrane domain IV is involved in ion transport activity and pH regulation of the NhaA- Na^+/H^+ antiporter of *Escherichia coli*. *Biochemistry* 41: 609-617.
- Galili, L., Herz, K., Dym, O., and Padan, E.** (2004) Unraveling functional and structural interactions between transmembrane domains IV and XI of NhaA Na^+/H^+ antiporter of *Escherichia coli*. *J. Biol. Chem.* 279: 23104-23113.
- Gerchman, Y., Olami, Y., Rimon, A., Taglicht, D., Schuldiner, S., and Padan, E.** (1993) Histidine-226 is part of the pH sensor of NhaA, a Na^+/H^+ antiporter in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 90: 1212-1216.
- Gerchman, Y., Rimon, A., and Padan, E.** (1999) A pH-dependent conformational change of NhaA Na^+/H^+ antiporter of *Escherichia coli* involves loop VIII-IX, plays a role in the pH response of the protein, and is maintained by the pure protein in dodecyl maltoside. *J. Biol. Chem.* 274: 24617-24624.
- Gerchman, Y., Rimon, A., Venturi, M., and Padan, E.** (2001) Oligomerization of NhaA, the Na^+/H^+ antiporter of *Escherichia coli* in the membrane and its functional and structural consequences. *Biochemistry* 40: 3403-3412.
- Godt, A., Franzen, C., Veit, S., Enkelmann, V., Pannier, M., and Jeschke, G.** (2000) EPR probes with well defined, long distances between two or three unpaired electrons. *J. Org. Chem.* 65: 7575-7582.
- Hubbell, W. L., and Altenbach, C.** (1994) Investigation of structure and dynamics in membrane proteins using site-directed spin labeling. *Curr. Opin. Struct. Biol.* 4: 566-573.
- Hubbell, W. L., Cafiso, D. S., and Altenbach, C.** (2000) Identifying conformational changes with site-directed spin labeling. *Nat. Struct. Biol.* 7: 735-739.
- Hunte, C., and Michel, H.** (2000) Membrane protein crystallization. In *Membrane protein purification and crystallization*, 2nd ed. (Hunte, C., Von Jagow, G., and Schaegger, H., eds.) pp. 143-160, Elsevier Science, San Diego, CA.
- Jeschke, G.** (2002) Distance measurements in the nanometer range by pulse EPR. *ChemPhysChem* 3: 927-932.
- Jeschke, G., Koch, A., Jonas, U., and Godt, A.** (2002) Direct conversion of EPR dipolar time evolution data to distance distributions. *J. Magn. Reson.* 155: 72-82.

- Jeschke, G., Wegener, C., Nietschke, M., Jung, H., and Steinhoff, H.-J.** (2004a) Inter-residual distance determination by four-pulse DEER in an integral membrane protein: the Na⁺/proline transporter PutP of *Escherichia coli*. *Biophys. J.* 86: 2551-2557.
- Jeschke, G., Panek, G., Godt, A., Bender, A., and Paulsen, H.** (2004b) Data analysis procedures for pulse ELDOR measurements of broad distance distributions. *Appl. Magn. Reson.* 26: 223-244.
- Jeschke, G., Bender, A., Paulsen, H., Zimmermann, H., and Godt, A.** (2004c) Sensitivity enhancement in pulse EPR distance measurements. *J. Magn. Reson.* 169: 1-12.
- Jeschke, G., Bender, A., Schweikardt, T., Panek, G., Decker, H., and Paulsen, H.** (2005) Localization of the N-terminal domain in light-harvesting chlorophyll a/b protein by EPR measurements. *J. Biol. Chem.* 280: 18623-18630.
- Jung, H., Rübenhagen, R., Tebbe, S., Leifker, K., Tholema, N., Quick, M., and Schmid, R.** (1998) Topology of the Na⁺/proline transporter of *Escherichia coli*. *J. Biol. Chem.* 273: 26400-26407.
- Milov, A. D., Ponomarev, A. B., and Tsvetkov, Y. D.** (1984) Electron-electron double resonance in electron spin echo: model biradical systems and the sensitized photolysis of decalin. *Chem. Phys. Lett.* 110: 67-72.
- Milov, A. D., Maryasov, A. G., and Tsvetkov, Y. D.** (1998) Pulsed electron double resonance (PELDOR) and its application in freeradicals research. *Appl. Magn. Reson.* 15: 107-143.
- Olami, Y., Rimon, A., Gerchman, Y., Rothman, A., and Padan, E.** (1997) Histidine 225, a residue of the NhaA-Na⁺/H⁺ antiporter of *Escherichia coli* is exposed and faces the cell exterior. *J. Biol. Chem.* 272: 1761-1768.
- Padan, E., Venturi, M., Gerchman, Y., and Dover, N.** (2001) Na⁺/H⁺ antiporters. *Biochim. Biophys. Acta* 1505: 144-157.
- Padan, E., Tzuberly, T., Herz, K., Kozachkov, L., Rimon, A., and Galili, L.** (2004) NhaA of *Escherichia coli*, as a model of a pH-regulated Na⁺/H⁺ antiporter. *Biochim. Biophys. Acta* 1658: 2-13.
- Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W.** (2000) Dead-time free measurement of dipole-dipole interactions between electron spins. *J. Magn. Reson.* 142: 331-340.
- Radzwill, N., Gerwert, K., and Steinhoff, H.-J.** (2001) Time-resolved detection of transient movement of helices F and G in doubly spinlabeled bacteriorhodopsin. *Biophys. J.* 80: 2856-2866.
- Ravna, A. W., Sylte, I., and Dahl, S. G.** (2001) Molecular model of the *Escherichia coli* Na⁺/H⁺ antiporter NhaA. *Receptors Channels* 7: 319-328.
- Rimon, A., Gerchman, Y., Kariv, Z., and Padan, E.** (1998) A point mutation (G338S) and its suppressor mutations affect both the pH response of the NhaA-Na⁺/H⁺ antiporter as well as the growth phenotype of *Escherichia coli*. *J. Biol. Chem.* 273: 26470-26476.

- Rimon, A., Tzuberly, T., Galili, L., and Padan, E.** (2002) Proximity of cytoplasmic and periplasmic loops in NhaA Na⁺/H⁺ antiporter of *Escherichia coli* as determined by site-directed thiol cross-linking. *Biochemistry* 41: 14897-14905.
- Rothman, A., Padan, E., and Schuldiner, S.** (1996) Topological analysis of NhaA, a Na⁺/H⁺ antiporter from *Escherichia coli*. *J. Biol. Chem.* 271: 32288-32292.
- Schuldiner, S., and Padan, E.** (1993) Molecular analysis of the role of Na⁺/H⁺ antiporters in bacterial cell physiology. *Int. Rev. Cytol.* 137C: 229-266.
- Steinhoff, H.-J., Radzwill, N., Thevis, W., Lenz, V., Brandenburg, D., Antson, A., Dodson, G., and Wollmer, A.** (1997) Determination of interspin distances between spin labels attached to insulin: comparison of electron paramagnetic resonance data with the X-ray structure. *Biophys. J.* 73: 3287-3298.
- Steinhoff, H.-J.** (2002) Methods for study of protein dynamics and protein-protein interaction in protein-ubiquitination by electron paramagnetic resonance spectroscopy. *Front. Biosci.* 7: c97-110.
- Taglicht, D., Padan, E., and Schuldiner, S.** (1991) Overproduction and purification of a functional Na⁺/H⁺ antiporter coded by *nhaA* (*ant*) from *Escherichia coli*. *J. Biol. Chem.* 266: 11289-11294.
- Taglicht, D., Padan, E., and Schuldiner, S.** (1993) Proton-sodium stoichiometry of NhaA, an electrogenic antiporter from *Escherichia coli*. *J. Biol. Chem.* 268: 5382-5387.
- Veenhoff, L. M., Heuberger, E. H. M. L., and Poolman, B.** (2002) Quaternary structure and function of transport proteins. *Trends Biochem. Sci.* 27: 242-249.
- Weese, J.** (1992) A reliable and fast method for the solution of Fredholm integral-equations of the 1st kind based on Tikhonov regularization. *Comput. Phys. Commun.* 69: 99-111.
- Williams, K. A., Geldmacher-Kaufner, U., Padan, E., Schuldiner, S., and Kühlbrandt, W.** (1999) Projection structure of NhaA, a secondary transporter from *Escherichia coli*, at 4.0 Å resolution. *EMBO J.* 18: 3558-3563.
- Williams, K. A.** (2000) Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature* 403:112-115.

CHAPTER 7***High-resolution structure of a Na⁺/H⁺ antiporter dimer obtained by pulsed EPR distance measurements*****7.1. Abstract**

Transient or partial formation of complexes between biomacromolecules is a general mechanism used to control cellular functions. Several of these complexes escape structure determination by crystallographic means. We developed a new approach for determining the structure of protein dimers in the native environment (e.g., in the membrane) with high resolution in cases where the structure of the two monomers is known. The approach is based on measurements of distance distributions between spin labels in the range between 2 and 6 nanometers by a pulsed electron paramagnetic resonance technique and explicit modeling of spin label conformations. By applying this method to the membrane protein homodimer of the Na⁺/H⁺ antiporter NhaA of *Escherichia coli*, the structure of the presumably physiological dimer was determined. It reveals two points of contact between the two monomers, with one of them confirming results of earlier cross-linking experiments.

7.2. Introduction

Many processes in living cells depend on the transient or partial formation of complexes between biomacromolecules. These processes include signal transduction via protein cascades, regulation of gene expression, and the control of individual activities of transporters and enzymes. The structure of such complexes may be difficult to elucidate by crystallographic techniques, as packing interactions in the crystal compete with the weak interactions holding together a complex that is in equilibrium with its components. The Na⁺/H⁺ antiporter NhaA of *Escherichia coli* (Padan *et al.*, 2004) is a case in point. This antiporter is involved in the regulation of intracellular pH, cellular Na⁺ content, and cell volume - functions shared with many other Na⁺/H⁺ antiporters of microorganisms, plants, and animals. Although regulation of the Na⁺ content is based on the level of expression of *nhaA*, the pH response occurs on the protein level. NhaA is downregulated below pH 6.5 and exhibits an activity increase by a factor of 1000 upon shift to alkaline pH with maximum activity at pH 8.5. Dimer formation of NhaA was observed in a low-resolution structure obtained by cryoelectron microscopy (cryo-EM) on two-dimensional (2D) crystals (Williams *et al.*, 1999; Williams, 2000), and functional complementation of mutants that are lethal at high pH and Na⁺ concentration suggests that the antiporter functions as a dimer (Gerchmann *et al.*, 2001).

This suggestion is also supported by cross-linking experiments. Reversible incomplete pH-dependent dimerization was observed in the pH range between 5.8 and 8.0 by pulsed electron paramagnetic resonance (EPR) distance measurements between spin-labeled mutants of NhaA reconstituted into liposomes (Hilger *et al.*, 2005). Yet, a recently obtained highly resolved X-ray structure displays nonphysiological association of the molecules in three-dimensional (3D) crystals (Hunte *et al.*, 2005), even though the monomer structure appears to be in good agreement with the structure observed in the dimers in 2D crystals (Screpanti *et al.*, 2006). The crystal structure provides substantial insight into the possible locations of the ion-translocation site and of the pH sensor. It suggests that long-range interactions in the protein are required to transmit the signal from this pH sensor to the translocation site. However, the detailed structural changes during ion transport, and in particular the possible role of dimerization, remain unknown. In this context, we report here on a novel approach of protein-protein interaction analysis and its application to the determination of the structure of the presumably physiological NhaA dimer in its downregulated state.

Assuming that the structure of the two NhaA monomers is nearly unchanged in the dimer with respect to the crystal structure (Protein Data Bank identifier 1ZCD), only the relative arrangement of the two molecules has to be determined to obtain a structure of the dimer. This corresponds to solving the problem of docking two biomacromolecules based on long-range distance constraints. To discuss interactions in the interface between the two moieties, this arrangement needs to be known with atomistic resolution. Our approach to this type of problem relies on site-directed spin labeling (SDSL) (Altenbach *et al.*, 1990) combined with measurements of label-to-label distances (Rabenstein and Shin, 1995) and restraint-driven rigid body transformations (Sompornpisut *et al.*, 2001). By applying a pulsed electron-electron double-resonance technique (Milov *et al.*, 1981; Pannier *et al.*, 2000), we obtain distance constraints in the range between 20 Å and 60 Å. As this range matches the dimensions of biomacromolecules, a relatively small number of distance constraints provides a unique structure. Compared to a recently introduced approach that also combined X-ray data and pulsed EPR distance measurements for reconstruction of the chemotaxis receptor-kinase assembly (Park *et al.*, 2006), our method for the first time, to our knowledge, utilizes information on the distribution of distances (Jeschke *et al.*, 2001) to match a modeled conformational distribution of the spin labels (Sale *et al.*, 2005; Borovykh *et al.*, 2006; Jeschke and Polyhach, 2007).

By this technique we overcome the intrinsic imprecision of SDSL EPR caused by its reliance on labels with a size of ~1 nm and obtain a resolution of the dimer model that is limited only by the resolution of the X-ray structure of the monomer. As EPR does not require any long-range order in the packing, this approach allows to determine the structure of a

membrane protein dimer in liposomes, i.e., in an environment that is closer to the physiological environment than a protein crystal or 2D crystal. Furthermore, external parameters such as the pH value can be controlled. Similar techniques should be applicable to heterodimers of both soluble and membrane proteins or to complexes of proteins with RNA or DNA. Therefore, we first discuss the general approach for structure elucidation of a complex of two macromolecules, while pointing out simplifications that could be made for the symmetric homodimer of NhaA. Then we describe the obtained structure of the NhaA dimer, and, finally, we discuss its functional implications.

7.3. Materials and Methods

7.3.1. Preparation and labeling of NhaA

The generation of *nhaA* alleles encoding single-Cys NhaA molecules used in this study has already been described (Gerchmann *et al.*, 2001; Rimon *et al.*, 2002). For overexpression, the resulting single-Cys NhaA encoding plasmids were transformed into *E. coli* TA16. TA16 is *nhaA*⁺ *nhaB*⁺ *lacIQ* and otherwise isogenic to TA15 (Taglicht *et al.*, 1991). Cells were grown aerobically in Luria-Bertani medium (Miller, 1992) containing 100 µg/ml ampicillin and 50 µg/ml kanamycin at 37 °C to an $A_{420\text{ nm}}$ of 1.0. After induction with 0.5 mM isopropyl-thio-β-D-galactoside and further growth for 3 h, membranes were prepared as described (Jung *et al.*, 1998) using a cell disruptor (IUL Instruments, Königswinter, Germany). NhaA was solubilized and purified by Ni-nitrilotriacetic acid affinity chromatography as described (Rimon *et al.*, 1998). Labeling the single-Cys NhaA with (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate (MTSSL, Toronto Research Chemicals, Toronto, Canada) was performed on the column.

For this purpose, 1 mM MTSSL in buffer W (20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5 mM imidazole, 10% glycerol (v/v), 0.1% β-D-dodecylmaltoside (w/v)) was applied to the column and incubated at 4 °C for 3 hours. Unbound label was removed by washing the column with buffer W, and labeled protein was eluted with 300 mM imidazole in buffer W. After elution the protein was reconstituted under nonreducing conditions into liposomes composed of *E. coli* lipids (Avanti Polar Lipids, Alabaster, AL, 67% phosphatidylethanolamine, 23.2% phosphatidylglycerol, and 9.8% cardiolipin) at a lipid/protein ratio of 20:1 (w/w) as described (Jung *et al.*, 1998). Finally, the proteoliposomes were washed twice with 50 mM KPi , pH 7.5, and resuspended in 10 mM KPi , pH 5.8, 100 mM KCl, and 5 mM MgCl_2 to yield a NhaA concentration of 50-100 µM. Proteoliposomes were frozen and stored in liquid nitrogen until use.

7.3.2. EPR measurements

Four-pulse double electron-electron resonance (DEER) measurements were performed with a Bruker Elexsys 580 spectrometer (Bruker Biospin, Karlsruhe, Germany) equipped with a 3 mm split-ring resonator under conditions of strong overcoupling ($Q \approx 100$) at a temperature of 50 K. Before insertion into the probe head, the samples were shock-frozen in liquid nitrogen to avoid crystallization of water. The four-pulse DEER sequence $(\pi/2)_{\nu_1} - \tau_1 - (\pi)_{\nu_1} - t' - (\pi)_{\nu_2} - \tau_1 + \tau_2 - t' - (\pi)_{\nu_1} - \tau_2$ -echo was used (Pannier *et al.*, 2000). The $\pi/2$ and π pulses at the observer frequency ν_1 had equal pulse lengths of 32 ns to ensure equal excitation bandwidths, whereas the π pulse at the pump frequency ν_2 had a length of 12 ns to maximize modulation depth while still keeping the two excitation bands separate. These conditions maximize sensitivity of the experiment (Jeschke *et al.*, 2004a). The long interpulse delay was $\tau_2 = 2000$ ns except for mutants E82R1, S87R1, and N177R1, where it was 2300 ns, and A118R1, where it was 2500 ns. An initial value of $t' = 80$ ns and an increment $\Delta t' = 8$ ns were used to acquire the time trace. To suppress proton modulation, data were added for eight equidistant values of τ_1 between 200 and 256 ns. A phase cycle $[+(+x), -(-x)]$ was applied to the first pulse and the two signals were subtracted. The pump frequency ν_2 (typically 9.33 GHz) was set to the center of the resonator mode and to coincide with the global maximum of the nitroxide spectrum. The observer frequency ν_1 was set to the local maximum at the low-field edge of the spectrum ($\nu_1 - \nu_2 = 65$ MHz). Accumulation times for the data sets varied between 8 and 14 h. Data were analyzed for dipolar evolution times $t = t' - \tau_1 \geq 0$.

7.3.3. Estimates of mean distances

The home-written program DeerAnalysis2006 (Jeschke *et al.*, 2006) was used for processing of the primary experimental data $V(t)$. To account for suppression of short distances due to low modulation depth (Jeschke *et al.*, 2006; Milov *et al.*, 2004), the background for a two-dimensional distribution of the spin-labeled objects was computed from a distance distribution that is defined as $P_B(r) = 0$ for $r < 1.5$ nm and otherwise as $P_B(r) = 2c_2r$, where c_2 is the area density of singly spin-labeled objects. The background factor $B(t)$ was determined with an optimized fit range (Jeschke *et al.*, 2006), and the baseline corrected form factor $F(t) = [V(t) - B(t)]/B(t)$ was computed. The distance distribution $P(r)$ for the spin labels within the same dimer was then estimated by Tikhonov regularization (Tikhonov, 1995), using the L curve as a criterion for selecting the optimum regularization parameter (Hansen, 1992; Chiang *et al.*, 2005). The mean distance of the distance

distribution was computed, disregarding any minor peaks that were well separated from the major peak.

7.3.4. Grid search for determination of the initial structure

A grid of angle values θ and ϕ and of translation parameters x and y is set up. For each parameter set $(\theta_k, \phi_k, x_k, y_k)$, mean distances $\langle r_i \rangle_{MO}$ for the model are computed, where index k runs over all grid points and index i runs over the selected labels (A202R1: $I = 1$, K221R1: $I = 2$, H225R1: $I = 3$, and E241R1: $I = 4$). The initial structure corresponds to the parameter set, for which the root mean-square deviation (RMSD)

$$RMSD = \sqrt{\sum_i (\langle r_i \rangle_T - \langle r_i \rangle_{MO})^2} \quad (7.1.)$$

is minimum. We find $\theta_0 = 45^\circ$, $\phi_0 = 320^\circ$, $x_0 = -40 \text{ \AA}$, $y_0 = 20 \text{ \AA}$.

7.3.5. Structure refinement by direct fit to the primary experimental data

Nelder-Mead simplex minimization of the RMSD between simulated and experimental four-pulse DEER data was used for structure refinement. Simulation of four-pulse DEER data involved the following steps. For a given parameter set (θ, ϕ, x, y) , the theoretical distance distribution $P_{T,i}(r)$ between N-O midpoints for all nine labels was computed as described above. This distance distribution was then converted to a theoretical form factor $F_{T,i}(t)$ by applying subroutine pcf2deer of the DeerAnalysis2006 package (Jeschke *et al.*, 2006). The theoretical four-pulse DEER data are computed as

$$V_{T,i}(t) = [1 - \lambda_i F_{T,i}(t)] B_{T,i}(t) \quad (7.2.)$$

where the theoretical background factor $B_{T,i}(t)$ is computed with subroutine pcf2deer from a distance distribution $P_{B,i}(r)$ which is identical zero for $r < 15 \text{ \AA}$ and given by $P_{B,i}(r) = c_{2,i} r$ for $r \geq 15 \text{ \AA}$. The modulation depth λ_i and the concentration parameter $c_{2,i}$ are determined by minimizing the RMSD between $V_{T,i}(t)$ and the primary experimental data $V_i(t)$ for this label, i.e., by a fit within a fit procedure. The figure of merit of the model is the sum of these minimum RMSD values for all nine labels, it depends on (θ, ϕ, x, y) . Minimization with respect to these four parameters provides the geometric parameters for the final structure.

7.3.6. Estimate of the resolution of the structure

The final structure was systematically tested with respect to the influence of the number of constraints and of the experimental noise on backbone atom positions. To this end the 36 possible combinations of seven out of the nine existing DEER data sets were used, and the noise level in each individual DEER data set was estimated from the difference

between experimental and simulated data for the final structure obtained with all nine data sets. For each of the 36 combinations the whole procedure of structure determination was repeated four times. In each trial the noise level of the DEER traces was doubled with respect to the original data by adding different pseudorandom artificial noise. This procedure gave a family of 144 approximate structures of the dimer. Each of these structures was superimposed onto the final structure from refinement with all nine constraints using the Magic Fit function of DeepView (Guex and Peitsch, 1997), and the backbone RMSD of each structure from the final structure was computed using DeepView scripting language. The total RMSD of the whole structure family from the final structure was 0.6 Å, which compares to a 3.45 Å resolution of the crystal structure used in deriving the dimer structure.

7.3.7. Superposition of the EPR and cryo-electron microscopy structures

A 2D electron density map of a single dimer was cut out from Fig. 4 of Williams (2000) and was converted to a smooth grayscale picture using Corel PhotoPaint to ensure better visibility. Utmost care was taken not to suppress or enhance any features of the original density map as published in Williams (2000). Both in the cryo-EM and in the EPR structure, the membrane normal is known. Superposition thus involves fitting of only one unknown parameter, an angle θ_{EM} corresponding to an Euler rotation about the C_2 symmetry axis of the dimer. This angle was fitted by visual inspection of the superposition picture.

7.4. Results and Discussion

7.4.1. Structure determination and refinement

In the first stages of structure determination, we treat the two component molecules as rigid bodies with structures known from previous crystallographic or NMR studies. The relative arrangement of two rigid bodies is fully described by six parameters: the three Euler angles accounting for the rotational degrees of freedom and the three components of a translation vector. A unique structure can thus be obtained by measuring six distances between the components. In practice, the actual number of distance constraints should exceed the number of free parameters at least by a factor of two. This allows us to test whether the components can indeed be treated as rigid bodies and to estimate the uncertainty of the obtained structure. Such a number of distance constraints can be obtained for proteins by current SDSL and pulsed EPR methodology (Borbat *et al.*, 2006). In the case of a symmetric homodimer, as assumed for NhaA, C_2 symmetry reduces the number of parameters to four (Fig. 7.1.). Two angles, θ and ϕ , determine the orientation of the symmetry axis in the frame of the monomer coordinates.

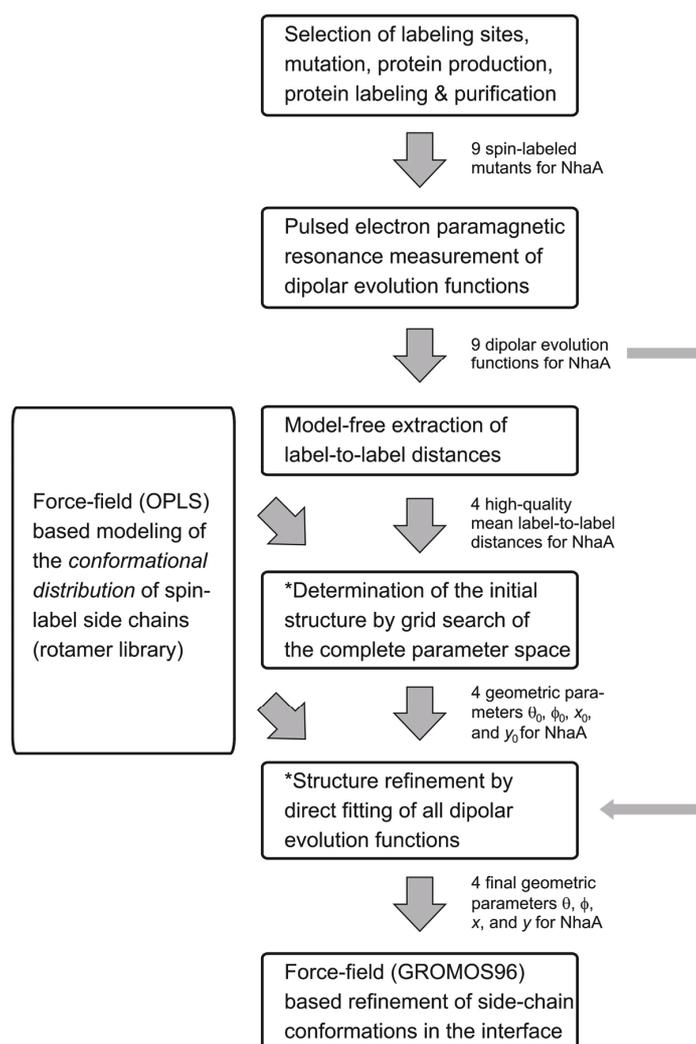


Figure 7.1.: Overview schematic for the general EPR structure determination approach for protein dimers. Steps labeled with an asterisk require an existing structure of the monomers, e.g., from previous X-ray or NMR work.

The coordinates of the second molecule are then generated from the coordinates of the first molecule by a rotation about the C_2 axis (z axis) by 180° , taking the origin of the frame as the geometric center of the protein, and by a subsequent translation by a vector $(x, y, 0)$. The translation is within the membrane plane. We decided to determine the four structural parameters θ , ϕ , x , and y based on nine distance measurements by the four-pulse DEER sequence (Pannier *et al.*, 2000), using single-cysteine mutants E82R1, S87R1, A118R1, N177R1, A202R1, K221R1, H225R1, E241R1, and V254R1 labeled with the methanethiosulfonate spin label (MTSSL) (R1 corresponds to MTSSL-labeled cysteine). As a basis for the structure determination, we first converted all DEER data sets to distance distributions by Tikhonov regularization (Tikhonov, 1995) with the optimum regularization parameter determined from the L curve (Hansen, 1992; Chiang *et al.*, 2005) and computed mean distances. Previously, mean distances have been found to be stable parameters of broad distance distributions, which are not much affected by the ill-posedness of the data conversion from time to distance domain (Jeschke *et al.*, 2004b). This first step of data

analysis is illustrated in Fig. 7.2. for the mutant K221R1, and the mean distances are given in Table 7.1..

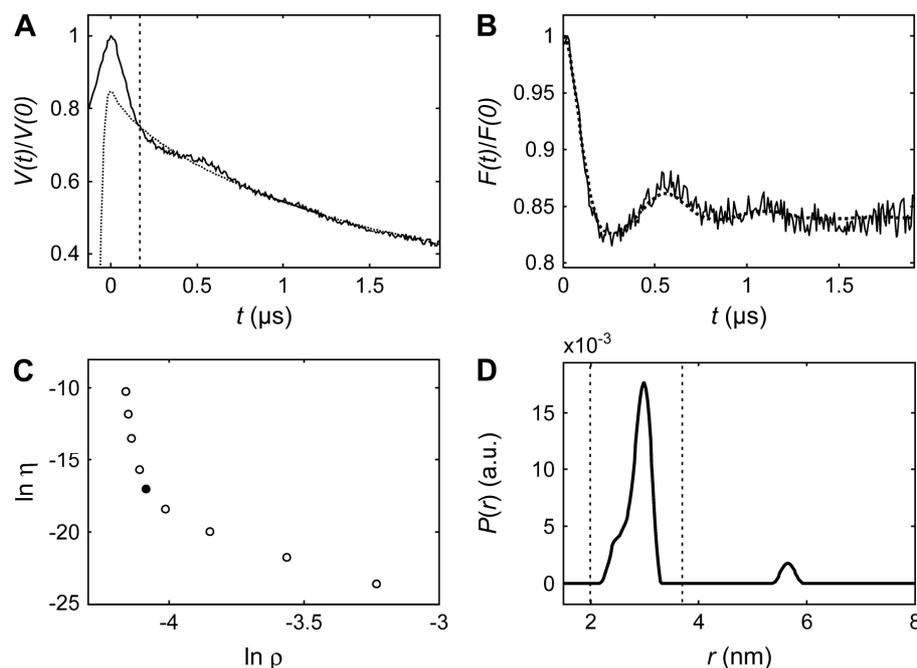


Figure 7.2.: Estimate of the mean distance on the example of spin label K221R1. (A) Normalized primary experimental data (solid line) and fit by a homogeneous background distribution in a plane for $r > 1.5 \text{ nm}$ (dotted line, fit range right from the vertical dashed line). Left from the dashed line the fit is extrapolated and is invalid for $t < 0$. (B) Normalized form factor (solid line) and best fit by Tikhonov regularization with optimized regularization parameter (dotted line). (C) L curve of Tikhonov regularization. The solid circle denotes the selected corner of the L curve, corresponding to a regularization parameter $\alpha = 10$. (D) Distance distribution obtained by Tikhonov regularization. The mean value of the distance for the peak between the two vertical dashed lines, $\langle r \rangle_T = 2.83 \text{ nm}$, was used in the first stage of structure determination.

Table 7.1.: Mean distances between spin labels in the NhaA dimer (in angstroms).

Mutant	E82R1	S87R1	A118R1	N177R1	A202R1	K221R1	H225R1	E241R1	V254R1
$\langle r \rangle_T (\text{\AA})$	44.8	47.3	45.8	48.4	34.1	28.3	44.5	25.7	18.9
$\langle r \rangle_M (\text{\AA})$	50.3	53.6	81.8	54.7	36.0	28.6	46.1	27.5	16.2

The value obtained by Tikhonov regularization is $\langle r \rangle_T$, and the value corresponding to the final structural model is $\langle r \rangle_M$. Values that were used in the first stage of structure determination are marked by bold typeface.

To obtain a first set of geometrical parameters (θ , ϕ , x , y) without making any further assumptions on the structure, the full parameter space should be searched. We performed a grid search as described in the Materials and Methods section for the range of θ from 0° to 90° and of ϕ from 0° to 355° , both in steps of 5° . For the translation parameters x and y we tested a range from -100 to 100 \AA in steps of 5 \AA . These ranges ensured that all relative arrangements were included for which the two molecules can actually be in contact with each other. During the grid search we fitted the mean distances between the spin labels in the structural model to mean experimental distances. The mean experimental distances were obtained by model-free analysis of the label-to-label distance distribution (Jeschke *et al.*, 2001; Sale *et al.*, 2005; Jeschke *et al.*, 2006). For the modeling of the conformational distribution, we used a library of 62 rotamers of MTSSL attached to a cystein residue and considered restrictions due to clashes with neighboring protein side groups or with the backbone by a Lennard-Jones potential parameterized as in the optimized potential for liquid

simulations (OPLS) force field (Jorgensen and Tirado-Rives, 1988). The label site was identified with the midpoint of the N-O bond of the spin label. In the grid search we used the minimum number of four distance constraints. That way we could select residues with a mean distance in the range where it can be determined most precisely by pulsed EPR (20-45 Å). For NhaA, these are residues A202R1, K221R1, H225R1, and V241R1. The initial parameter set obtained by the grid search is $\theta_0 = 45^\circ$, $\phi_0 = 320^\circ$, $x_0 = -40$ Å, $y_0 = 20$ Å.

In the first refinement step we fitted the parameter set by minimizing the RMSD between the primary experimental DEER time traces and simulated DEER time traces. This fitting procedure utilizes the full information on the distance distribution contained in the primary DEER data while avoiding the ill-posed problem of explicitly converting the time-domain data to a distance distribution. The simulated DEER time traces are based on the modeled distance distribution between the label sites, considering all 62 rotamers in each moiety with their appropriate weighting factors derived from the Lennard-Jones potential. A 2D background function was used in the DEER simulations to account for contributions from neighboring dimers (Hilger *et al.*, 2005). In this refinement step all nine distance constraints were used. The best-fit parameter set is $\theta_0 = 42.7^\circ$, $\phi_0 = 324.0^\circ$, $x_0 = -39.7$ Å, $y_0 = 20.1$ Å.

The change compared to the values previously found in the grid search with only four of the nine constraints is relatively small. This suggests a good internal consistency of the experimental data and thus validity of the hypothesis that the two monomers move as rigid bodies. The quality of the final fit is illustrated in Fig. 7.3.. The modulation depth λ_i ranges between 0.09 and 0.23 for most residues, probably reflecting differences in labeling efficiency. We may not exclude that the extent of dimer formation also varies somewhat for the different labeling sites. Exceptionally low modulation depths are observed for mutants E82R1 ($\lambda = 0.02$) and V254R1 ($\lambda = 0.04$). In the latter case, the low modulation depth is probably due to suppression effects for distances shorter than 1.8 nm (Jeschke *et al.*, 2006; Milov *et al.*, 2004). The background density ranges between 0.043 and 0.102, with the lowest value again corresponding to mutant E82R1. The relatively narrow range for the background density suggests that separation into background and form factor work well.

Although the fit quality is generally good, significant deviations of the fits (*red lines*) from the experimental data (*black lines*) are observed for the two cases with the most nicely resolved oscillations, A202R1 and K221R1. This indicates limitations in the precision of the modeling of the conformational distributions by our rotamer approach, which may be due to our neglect of any interactions of the label with the protein except for van der Waals interactions. The modeled distance distributions appear to be somewhat broader than the true distributions. More sophisticated modeling of the label conformations by molecular dynamics simulations (Sale *et al.*, 2005; Steinhoff and Hubbell, 1996) could be integrated into our approach. It was recently demonstrated that such modeling significantly improves

agreement with experimental spin-to-spin distances compared to simply interpreting them as C^β - C^β distances (Sale *et al.*, 2005).

This raises the question how imperfections of the conformational modeling and noise in the experimental data influence the precision of the four geometrical parameters and thus the resolution of the structure. We checked this by examining a family of structures that was obtained by considering only seven out of the nine constraints. Structures that result from using any combination of seven constraints in both the initial grid search and first refinement have the same or a larger final RMSD compared to our best-fit structure. To estimate the reliability of the final coordinates, we repeated refinement four times for any of the 36 combinations of seven out of nine constraints while doubling the noise amplitude of the experimental data by adding pseudo-random numbers. Using only seven of the nine constraints ensured that errors in modeling of the label conformations and in the assumption of rigid-body behavior of the components contribute to the scatter in the coordinates. The backbone RMSD of the resulting set of 144 structures with respect to the best-fit structure is only 0.6 Å, which suggests that the quality of the final structure is limited by the resolution of 3.45 Å of the crystal structure of the monomer rather than by the SDSL EPR docking fit.

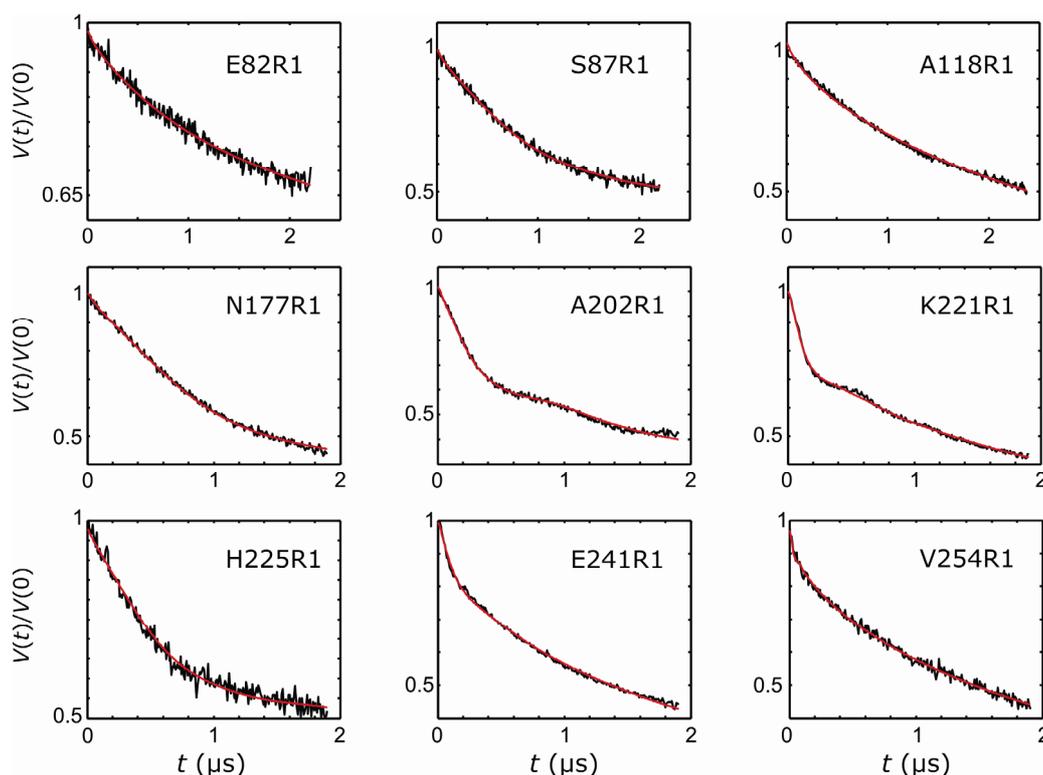


Figure 7.3.: Fits of primary experimental DEER data (*black lines*) by simulated data (*red lines*) corresponding to the final structure of the NhaA dimer and a distribution of spin label conformations modeled by a rotamer library.

This resolution is significantly better than the one achieved in an earlier EPR docking approach that was based on continuous-wave EPR-derived distance constraints (Sompornpisut *et al.*, 2001). The main improvements are the use of longer-range and higher

precision distance constraints from a pulsed EPR method and utilization of information on the whole distribution of distances rather than of only a single number for the site-to-site distances. These two features also distinguish our approach from protein-protein docking approaches based on NMR shift data or other 'ambiguous interaction restraints' (Dominguez *et al.*, 2003). Due to the requirement for spin labels our approach is more invasive compared to NMR titration. However, if labeling sites are carefully selected and are reexamined after solving the structure, any influence of the labels on protein-protein complex formation can be excluded.

The structure obtained at this point exhibits contacts between the two component molecules in three parts of the sequence, from residue 39 to 60 (β -sheets), from 202 to 211 (helix VII), and from 253 to 259 (helix IX). There exist no strong clashes between side group atoms and no clashes at all between backbone atoms of the two monomers. Nevertheless, it has to be realized that the side chains of the contact-forming residues are unlikely to be oriented in the same way as in the crystal structure of the monomer. As we do not have experimental information on side-chain orientation, we optimized the interface by repeated energy minimization using the GROMOS96 force field (van Gunsteren and Berendsen, 1990) as implemented in DeepView (Guex and Peitsch, 1997). The converged structure after 10 minimizations (Fig. 7.4.a) is the final structure discussed in the following.

7.4.2. Contacts in the NhaA dimer

Contacts between the two monomers in the dimer were analyzed with the CSU software (Sobolev *et al.*, 1999). The main intermolecular contact is mediated by the two-stranded antiparallel β -sheets, which are probably hold together by four hydrogen bonds formed between residues Q47 and R49 and backbone groups of the respective other β -sheet (Fig. 7.4.b) and by a number of hydrophobic contacts. Residue Q47 is also involved in stabilizing the nonphysiological dimer observed in 3D crystals (Hunte *et al.*, 2005). Contact between helices VII and IX (Fig. 7.4.c) is restricted to a few residues and most likely involves a hydrogen bond between R204 and V254 as well as hydrophobic contacts of R204 and L210 with W258. In general the interface is not densely packed, which indicates a relatively weak binding between the moieties in agreement with the pH-dependence of dimer formation (Hilger *et al.*, 2005) and the dissociation of the physiological dimer during 3D crystallization. Location of residue V254 in the interface was suggested before by a cross-linking study (Gerchman *et al.*, 2001). Indeed, we find a distance of 8.1 Å between the S atoms of the corresponding cysteins. This is in good agreement with the finding that the rigid cross-linking agent *p*-phenylenedimaleimide, which can cross-link only residues that are at least 10 Å apart, causes a change in the pH profile of NhaA activation (Williams, 2000). In contrast, the

flexible cross-linking agent 1,6-bismaleimido-hexane does not cause such a change. Location of residue V254 in the interface may also explain the low modulation depth, i.e. lower propensity for dimer formation observed in our earlier work when mutating this residue (Hilger *et al.*, 2005).

Contacts between the two monomers are restricted to one of the two bundles of six helices, the interfacial domain (*blue line* in Fig. 7.5.). According to previous work, this interfacial domain contains the pH sensor of NhaA around residues 241-252 (Hunte *et al.*, 2005). The other bundle of six helices contains the putative ion-translocation pathway and thus forms the translocation domain (*red line* in Fig. 7.5.). Although the two domains appear distinct, there are interdomain contacts. Helix I (residues 12-30) in the interfacial domain interacts with helix XII (357-382) near the periplasmic side and with helix IVc near the cytoplasmic side (Hunte *et al.*, 2005).

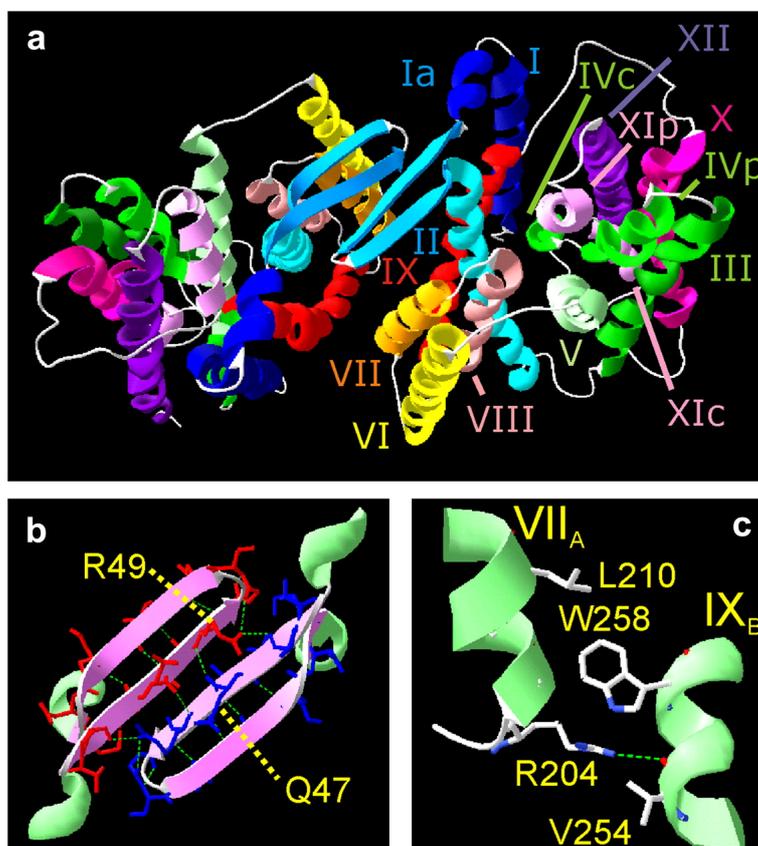


Figure 7.4. EPR structure of the dimer of the Na^+/H^+ antiporter NhaA of *E. coli*. (a) Overview of the whole structure. (b) Contacts between the two-stranded antiparallel β -sheets (residues 39-60). (c) Contacts between helices VII and IX. Highlighted residues are proposed to be directly involved in intermolecular interactions. Created with DeepView (Guex and Peitsch, 1997).

7.4.3. Comparison with the low-resolution structure

Although no information from the cryo-EM structure in 2D crystals (Williams *et al.*, 1999; Williams, 2000) was used, there is a rather good agreement between this structure and our EPR structure, in particular in the interfacial domain (Fig. 7.5.). This is not a trivial result, as the present EPR study is performed on NhaA dimers in liposomes rather than in 2D

crystals where packing effects might have influenced the structure of the weakly bound dimer. Indeed, our finding contrasts with a fit of the high-resolution monomer 3D crystal structure to the electron density obtained by cryo-EM, which predicted that the two-stranded antiparallel β -sheets would collide and should thus be displaced by ~ 6 Å out of the membrane plane and by ~ 6 Å towards the molecule center (Screpanti *et al.*, 2006). However, such a structure would imply shorter distances for residues A202R1 and K221R1 than we observe. As the DEER data for these two residues are of high quality, we can exclude this arrangement of the two moieties at least in liposomes. The disagreement may arise from the attempt to perfectly superimpose not only the interfacial domain but also the ion-translocation domain TrI, which appears to be slightly displaced in our structure with respect to the cryo-EM data.

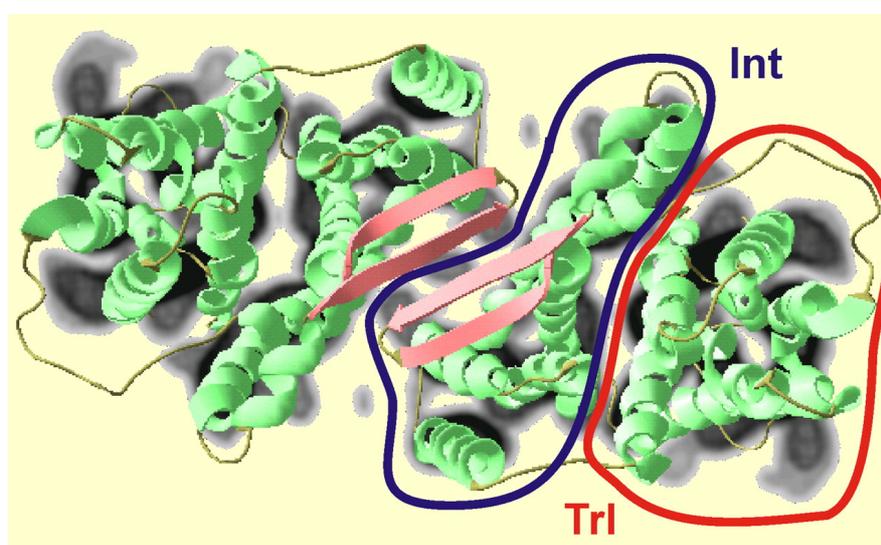


Figure 7.5.: Comparison between the EPR structure and the electron density projection to the membrane plane obtained by cryo-EM on 2D crystals (Williams *et al.*, 1999; Williams, 2000). The interfacial domain Int and the ion-translocation domain TrI involved in Na^+ and H^+ transport are indicated by the blue and red lines, respectively.

The displacement of the translocation domain relative to the interfacial domain could be a packing artifact of either the high-resolution monomer structure or of the cryo-EM structure. In the former case, the artifact would transmit to the EPR structure of the dimer, as we assumed rigid body motion of the two domains as a whole. Such a displacement in the 3D crystals may be induced by crystal contacts that stabilize the nonphysiological dimer, which are observed for residues D354, E356, and N359 in the translocation domain. This domain includes helices III (residues 95-116), IV (121-143), and V (150-175) and the helices X, XI, and XII with their intervening loops comprising residues 290 to the C terminus. Among the residues selected for mutation and distance measurements, only A118 is situated within this domain. Because of the long-distance between sites A118R1 in the two moieties of ~ 82 Å, our DEER data may not be sensitive to a displacement of this residue by up to 10 Å. A displacement of the translocation domain in liposomes compared to the crystal structure of

the monomer would thus be consistent with our data, but we cannot exclude that the difference is rather between liposomes and 2D crystals.

7.4.4. Functional implications of the EPR structure

For residues H225R1 and V254R1 (Hilger *et al.*, 2005) as well as A118R1 and N177R1, we tested for pH-dependence of the distance distribution by performing measurements at pH 5.8, 7.0 and 8.0. Quite unexpectedly, in none of these cases any significant changes were detected. Only the depth of the DEER modulation, which corresponds to the extent of dimer formation, varies with pH (Hilger *et al.*, 2005). This indicates that the structure of the interfacial domain, i.e., of the antiparallel β -sheets and helices I-III, and VI-IX, is conserved during pH changes. Also the interface between these domains appears to be conserved although propensity for dimer formation varies slightly.

Due to the scarcity of labeled residues in the translocation domain and the large separation between the two translocation domains in the dimer, we cannot exclude that this domain moves with respect to the interfacial domain on changing pH. In fact, with the pH sensor being located around residues 241-252 and Na^+/H^+ transport probably involving relative movement of helices IVc (residues 134-143) and Xlp (residues 340-350) (Hunte *et al.*, 2005), a change in the relative arrangement of the two domains might be the most likely explanation for transmittance of the signal from the pH sensor to the translocation pathway. The putative pH sensor is located near the contact region between helices VII and IX in the dimer. Mutation of V254, which forms a hydrogen bond with R204 across the interface, was found to change the pH-dependence of activity (Gerchmann *et al.*, 1999). Note also that the interface between the two monomers is lined by a number of polar residues (Fig. 7.6.). Among them R204 has a predicted pK_a value of 7.1 in a membrane environment (Olkhova *et al.*, 2006). Protonation/deprotonation of this residue may influence contacts between the monomers.

Signal transmittance between the translocation and interfacial domains is also required to rationalize functional complementation between the conditional lethal mutants. For instance, in the mutant H225R, antiporter activity increases between pH 6 and 7 but is shut off again at pH > 7, so that the mutant is lethal at pH 8.5 and high Na^+ concentration (Gerchman *et al.*, 2001). Likewise, mutant G338S is lethal, but for a different reason: in this mutant antiporter activity persists at pH below 7 so that at high Na^+ concentration pH in the cell becomes too low (Gerchmann *et al.*, 2001). If both mutants are coexpressed, the cells survive under the conditions that kill them if only one of the mutants is present. Such functional complementation between remote sites is only feasible if ion translocation is a process that involves a cooperative, but asymmetric motion of both moieties of the dimer.

This raises the questions how local structural changes in the two domains of a monomer are coupled and how structural changes are transmitted between the monomers in the dimer. Remarkable in this respect is the particularly high number of potentially positively charged residues (R203, R204, H243, R245, K249, R250, H253, and H256) which are at or close to the cytoplasmic contact site of the monomers (Fig. 7.6.). Taking into account that the predicted pK_a values for R204, H253, and H256 (Olkhova *et al.*, 2006) allow protonation/deprotonation events in the range between pH 5-7.5 it can be envisaged that pH directly influences contacts between the monomers. This idea is in agreement with the previous finding that pH influences the monomer-dimer equilibrium of NhaA (Hilger *et al.*, 2005).

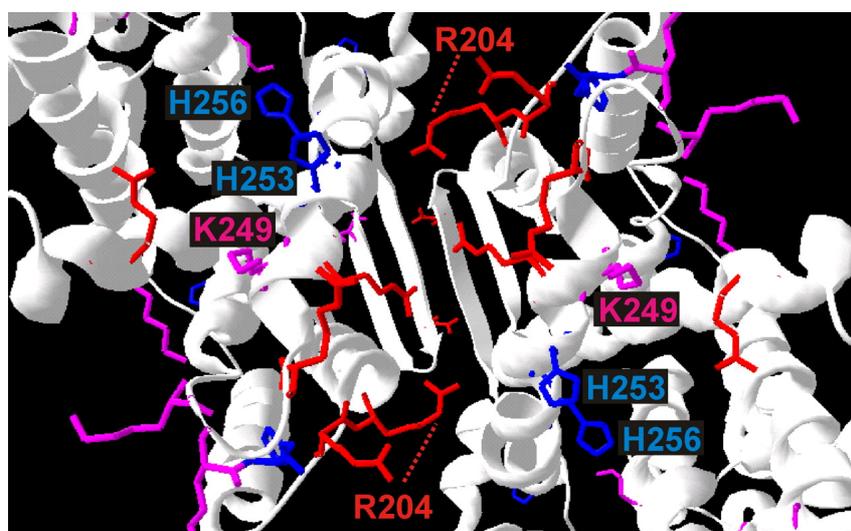


Figure 7.6.: Distribution of arginine (red), lysine (purple), and histidine (blue) residues near the interface. Figure created with DeepView (Guex and Peitsch, 1997).

Among the residues that can be protonated/deprotonated in the transition region, R204 may play a pivotal role in this process since it appears to form a hydrogen bond across the interface with the carbonyl group of V254 in the other monomer. V254 in turn is in hydrophobic contact with R250 in its own moiety. These interactions may also ensure contact between large clusters of residues strongly interacting within a monomer across the interface. These clusters were identified by a recent continuum electrostatics analysis of the monomer and involve E78, R81, E82, E252, H253, and H256 belonging to the pH sensor and D133, R203, H243, K249, and R250 (Rimon *et al.*, 1998). The clusters provide a coupling between the putative ion-translocation domain and the interfacial domain, as residue D133 is located in the narrow passage of the translocation channel (Hunte *et al.*, 2005).

Residue H225, which is spatially close to the pH sensor cluster, mediates contact between the interfacial domain, in which it is located, and residues A172 and Y175 in the translocation domain of the same moiety. Replacement of the histidine by arginine in the conditional lethal mutant H225R may modify or destroy this contact, which needs to exist in

one of the two moieties to ensure activity. A further contact between the translocation and interfacial domain in the same moiety is provided by residues 135-143 in helix IVc, which is in the cytoplasmic part of the putative ion-translocation channel, and residues 12-19 in helix I. Likewise some residues in the range between 64 and 72 in helix II have contact with residues in the range between 341 and 350 in helix XIp, which is in the periplasmic part of the ion-translocation channel. Interestingly, helix IVc and helix XIp were suggested to be the moving parts in pH-regulation of NhaA activity (Hunte *et al.*, 2005).

Based on the EPR structure of the dimer, one may now speculate that the proposed movement of helix IVc during activation is related to a movement of the cytoplasmic end of helix I towards the interface, whereas the proposed movement of helix XIp is related to a movement of the periplasmic end of helix II towards the interface. Activation would thus correspond to a clockwise rotation of the interface or, keeping the interfacial domains fixed, to an anticlockwise rotation of the two translocation domains with respect to the interface (seen from the periplasmic side). The displacement of the translocation domain seen in the EPR structure compared to the cryo-EM structure roughly corresponds to such an anticlockwise rotation, although it has to be said that the wildtype is still inactive at the pH of 5.8 where most of the EPR distance measurements were performed and that the reasons for the displacement are currently unknown. Clearly, structural information at pH above 6.5 is required to prove or refute the hypothesis and to obtain further insight into the cooperative motion of the two moieties of the dimer that appears to be the basis of ion translocation.

In conclusion, it is demonstrated that SDSL combined with pulsed EPR measurements of distances in the range between 20 and 60 Å and modeling of the conformational distribution of the spin label side groups provides a new approach for elucidating structures of dimers or complexes of biomacromolecules with essentially the same resolution that can be obtained for the structures of the constituents. The method does not provide experimental information on a possible repacking of side chains in the interface. In contrast to most established techniques, this approach is applicable even if interactions between the molecules are rather weak and complex or dimer formation is thus incomplete. It does not require crystallization; so that structures of membrane protein complexes can be determined in liposomes where external parameters such as pH can be varied.

7.5. Acknowledgments

We thank H. W. Spiess for access to the EPR spectrometer at Max Planck Institute for Polymer Research in Mainz and C. Bauer for technical assistance. This work was supported by Deutsche Forschungsgemeinschaft JE 246/3-2 and Ju 333/4-2.

7.6. References

- Altenbach, C., Marti, T., Khorana, H. G., and Hubbell, W. L.** (1990) Transmembrane protein-structure - Spin labeling of bacteriorhodopsin mutants. *Science* 248: 1088-1092.
- Borbat, P., Ramlall, T. F., Freed, J. H., and Eliezer, D.** (2006) Inter-helix distances in lysophospholipid micelle-bound alpha-synuclein from pulsed ESR measurements. *J. Am. Chem. Soc.* 128: 10004-10005.
- Borovykh, I. V., Ceola, S., Gajula, P., Gast, P., Steinhoff, H.-J., and Huber, M.** (2006) Distance between a native cofactor and a spin label in the reaction centre of *Rhodobacter sphaeroides* by a two-frequency pulsed electron paramagnetic resonance method and molecular dynamics simulations. *J. Magn. Reson.* 180: 178-185.
- Chiang, Y. W., Borbat, P. P., and Freed, J. H.** (2005) The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J. Magn. Reson.* 172: 279-295.
- Dominguez, C., Boelens, R., and Bonvin, A. M. J. J.** (2003) HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* 125: 1731-1737.
- Gerchmann, Y., Rimon, A., and Padan, E.** (1999) A pH-dependent conformational change of NhaA Na⁺/H⁺ antiporter of *Escherichia coli* involves loop VIII-IX, plays a role in the pH response of the protein, and is maintained by the pure protein in dodecyl maltoside. *J. Biol. Chem.* 274: 24617-24624.
- Gerchman, Y., Rimon, A., Venturi, M., and Padan, E.** (2001) Oligomerization of NhaA, the Na⁺/H⁺ antiporter of *Escherichia coli* in the membrane and its functional and structural consequences. *Biochemistry* 40: 3403-3412.
- Guex, N., and Peitsch, M. C.** (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 18: 2714-2723.
- Hansen, P. C.** (1992) Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.* 34: 561-580.
- Hilger, D., Jung, H., Padan, E., Wegener, C., Vogel, K.-P., Steinhoff, H.-J., and Jeschke, G.** (2005) Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na⁺/H⁺ antiporter of *E. coli*. *Biophys. J.* 89: 1328-1338.
- Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E., and Michel, H.** (2005) Structure of a Na⁺/H⁺ antiporter and insights into mechanism of action and regulation by pH. *Nature* 435: 1197-1202.
- Jeschke, G., Koch, A., Jonas, U., and Godt, A.** (2001) Direct conversion of EPR dipolar time evolution data to distance distributions. *J. Magn. Reson.* 155: 72-82.
- Jeschke, G., Bender, A., Paulsen, H., Zimmermann, H., and Godt, A.** (2004a) Sensitivity enhancement in pulse EPR distance measurements. *J. Magn. Reson.* 169: 1-12.

- Jeschke, G., Panek, G., Godt, A., Bender, A., and Paulsen, H.** (2004b) Data analysis procedures for pulse ELDOR measurements of broad distance distributions. *Appl. Magn. Reson.* 26: 223-244.
- Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H.** (2006) DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Reson.* 30: 473-498.
- Jeschke, G., and Polyhach, Y.** (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* 9: 1895-1910.
- Jorgensen, W. L., and Tirado-Rives J.** (1988) The OPLS potential functions for proteins - Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110: 1657-1666.
- Jung, H., Tebbe, S., Schmid, R., and Jung, K.** (1998) Unidirectional reconstitution and characterization of purified Na⁺/proline transporter of *Escherichia coli*. *Biochemistry* 37: 11083-11088.
- Milov, A. D., Salikhov, K. M., and Shirov, M. D.** (1981) Application of ELDOR in electron spin echo for paramagnetic center space distribution in solids. *Fiz. Tverd. Tela (Leningrad)* 23: 975-982.
- Miller, J. H.** (1992) *A short course in bacterial genetics. A laboratory manual and handbook for Escherichia coli and related bacteria*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Milov, A. D., Naumov, B. D., and Tsvetkov, Y. D.** (2004) The effect of microwave pulse duration on the distance distribution function between spin labels obtained by PELDOR data analysis. *Appl. Magn. Reson.* 26: 587-599.
- Olkhova, E., Hunte, C., Screpanti, E., Padan, E., and Michel, H.** (2006) Multiconformation continuum electrostatics analysis of the NhaA Na⁺/H⁺ antiporter of *Escherichia coli* with functional implications. *Proc. Natl. Acad. Sci. U. S. A.* 103: 2629-2634.
- Padan, E., Tzuberly, T., Herz, K., Kozachkov, L., Rimon, A., and Galili, L.** (2004) NhaA of *Escherichia coli*, as a model of a pH-regulated Na⁺/H⁺ antiporter. *Biochim. Biophys. Acta* 1658: 2-13.
- Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W.** (2000) Dead-time free measurement of dipole-dipole interactions between electron spins. *J. Magn. Reson.* 142: 331-340.
- Park, S. Y., Borbat, P. P., Gonzalez-Bonet, G., Bhatnagar, J., Pollard, A. M., Freed, J. H., Bilwes, A. M., and Crane, B. R.** (2006) Reconstruction of the chemotaxis receptor-kinase assembly. *Nat. Struct. Mol. Biol.* 13: 400-407.
- Rabenstein, M., and Shin, Y. K.** (1995) Determination of the distance between 2 spin labels attached to a macromolecule. *Proc. Natl. Acad. Sci. U. S. A.* 92: 8239-8243.
- Rimon, A., Gerchman, Y., Kariv, Z., and Padan, E.** (1998) A point mutation (G338S) and its suppressor mutations affect both the pH response of the NhaA-Na⁺/H⁺

- antiporter as well as the growth phenotype of *Escherichia coli*. *J. Biol. Chem.* 273: 26470-26476.
- Rimon, A., Tzuberly, T., Galili, L., and Padan, E.** (2002) Proximity of cytoplasmic and periplasmic loops in NhaA Na⁺/H⁺ antiporter of *Escherichia coli* as determined by site-directed thiol cross-linking. *Biochemistry* 41: 14897-14905.
- Sale, K., Song, L., Liu, Y. S., Perozo, E., and Fajer, P.** (2005) Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *J. Am. Chem. Soc.* 127: 9334-9335.
- Screpanti, E., Padan, E., Rimon, A., Michel, H., and Hunte, C.** (2006) Crucial steps in the structure determination of the Na⁺/H⁺ antiporter NhaA in its native conformation. *J. Mol. Biol.* 362: 192-202.
- Sobolev, V., Sorokine, A., Prilusky, J., Abola, E. E., and Edelman, M.** (1999) Automated analysis of interatomic contacts in proteins. *Bioinformatics* 15: 327-332.
- Sompornpisut, P., Liu, Y. S., and Perozo, E.** (2001) Calculation of rigid-body conformational changes using restraint-driven Cartesian transformations. *Biophys. J.* 81: 2530-2546.
- Steinhoff, H.-J., and Hubbell, W. L.** (1996) Calculation of electron paramagnetic resonance spectra from Brownian dynamics trajectories: Application to nitroxide side chains in proteins. *Biophys. J.* 71: 2201-2212.
- Taglicht, D., Padan, E., and Schuldiner S.** (1991) Overproduction and purification of a functional Na⁺/H⁺ antiporter coded by *nhaA* (*ant*) from *Escherichia coli*. *J. Biol. Chem.* 266: 11289-11294.
- Tikhonov, A. N.** (1995) *Numerical methods for the solution of ill-posed problems*, Kluwer Academic Publishers, Dordrecht, Boston.
- van Gunsteren, W. F., and Berendsen, H. J. C.** (1990) Computer simulation of molecular dynamics - Methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.* 29: 992-1023.
- Williams, K. A., Geldmacher-Kaufer, U., Padan, E., Schuldiner, S., and Kühlbrandt, W.** (1999) Projection structure of NhaA, a secondary transporter from *Escherichia coli*, at 4.0 Å resolution. *EMBO J.* 18: 3558-3563.
- Williams, K. A.** (2000) Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature* 403: 112-115.

CHAPTER 8***DeerAnalysis2006 - A comprehensive software package for analyzing pulsed ELDOR data*****8.1. Abstract**

Pulsed electron-electron double resonance techniques such as the four-pulse double electron-electron resonance experiment measure a dipolar evolution function of the sample. For a sample consisting of spin-carrying nanoobjects, this function is the product of a form factor, corresponding to the internal structure of the nanoobject, and a background factor, corresponding to the distribution of nanoobjects in space. The form factor contains information on the spin-to-spin distance distribution within the nanoobject and on the average number of spins per nanoobject, while the background factor depends on constraints, such as a confinement of the nanoobjects to a two-dimensional layer. Separation of the dipolar evolution function into these two contributions and extraction of the spin-to-spin distance distribution require numerically stable mathematical algorithms that can handle data for different classes of samples, e.g. spin-labelled biomacromolecules and synthetic materials. Furthermore, experimental imperfections such as the limited excitation bandwidth of microwave pulses need to be considered. The software package DeerAnalysis2006 provides access to a comprehensive set of tools for such data analysis within a common user interface. This interface allows for several tests of the reliability and precision of the extracted information. User-supplied models for the spin-to-spin distance distribution within a certain class of nanoobjects can be added to an existing library and be fitted with a universal algorithm.

8.2. Introduction

Pulsed electron-electron double resonance (PELDOR) experiments for the determination of distances between paramagnetic centers (Milov *et al.*, 1981; Milov *et al.*, 1984) have recently become a very active field of research with numerous applications on nanostructured materials (Pannier *et al.*, 2000a; Pannier *et al.*, 2001; Jeschke, 2002a; Jeschke and Godt, 2003; Hinderberger *et al.*, 2004a; Hinderberger *et al.*, 2004b), soluble proteins (Persson *et al.*, 2001; Elsässer *et al.*, 2002; van Amsterdam *et al.*, 2003; Zhou *et al.*, 2005; Nakamura *et al.*, 2005; Fajer, 2005; Bennati *et al.*, 2005; Kay *et al.*, 2006; Banham *et al.*, 2006), membrane proteins (Jeschke *et al.*, 2004a; Jeschke *et al.*, 2005; Hilger *et al.*, 2005; Borovykh *et al.*, 2006), nucleic acids (Schiemann *et al.*, 2004), and peptides (Milov *et al.*, 2000). In many of these studies, careful analysis of the dipolar evolution data was

recognized as a crucial precondition for their reliable interpretation. For the well-structured biomacromolecules such analysis also needs to consider the flexibility of the spin label (Borovykh *et al.*, 2006; Sale *et al.*, 2005). For broad or multimodal spin-to-spin distance distributions, as often encountered in materials (Pannier *et al.*, 2000a; Pannier *et al.*, 2001; Jeschke, 2002a; Jeschke and Godt, 2003; Hinderberger *et al.*, 2004a; Hinderberger *et al.*, 2004b), analysis in terms of a single value for the distance is not appropriate. These problems have inspired the development of several approaches for the conversion of dipolar evolution data to distance distributions (Jeschke *et al.*, 2002; Milov *et al.*, 2004a; Bowman *et al.*, 2004; Jeschke *et al.*, 2004b; Jeschke *et al.*, 2004c; Chiang *et al.*, 2005). The experience with these approaches suggested a number of modifications and extensions to our previous open-source program package DeerAnalysis2004 (Jeschke *et al.*, 2002; Jeschke *et al.*, 2004b; Jeschke *et al.*, 2004c) as well as a complete reprogramming of the graphical user interface. In addition to model-free computation of distance distributions, a flexible interface for model-based fitting (Pannier *et al.*, 2000a; Pannier *et al.*, 2001; Jeschke and Godt, 2003) is now included and recent work on effects of limited excitation bandwidth of the microwave (m.w.) pulses (Maryasov and Tsvetkov, 2000; Milov *et al.*, 2004b) is considered. Spin counting (Milov *et al.*, 1984; Hilger *et al.*, 2005) and concentration measurements (Milov *et al.*, 1984; Jeschke, 2002b) are implemented. Experimental background functions can be derived from singly labelled biomacromolecules and used for the analysis of data from doubly labelled molecules. This paper describes the underlying theory and algorithms and illustrates features of the program on data for three model systems: a shape-persistent biradical (Godt *et al.*, 2000), spin-labelled gold nanoparticles (Chechik *et al.*, 2004), and a spin-labelled membrane protein (Jeschke *et al.*, 2004a).

The paper is organized as follows. After the Materials and Methods section, the theoretical foundations of the program are described. In particular, we consider the sample as a collection of nanoobjects carrying one or more spins, with a spatial distribution of the nanoobjects that may be homogeneous or may be confined on a nanoscopic length scale to less than three dimensions. We discuss separation of the primary experimental data into a form factor, corresponding to spin-to-spin distances within the nanoobject, and a background factor, corresponding to the distribution of nanoobjects. Tikhonov regularization with L curve computation and model-based fitting are considered as techniques for extracting the distance distribution from the form factor. After explaining the determination of the average number of spins per nanoobject, the theoretical part concludes with a discussion of corrections that are necessary when the excitation bandwidth of the m.w. pulses is of the same order of magnitude or smaller than the largest dipole-dipole couplings between spins. The next section gives a brief description of the algorithms underlying the steps in data processing. This description is intended to provide the user of the program with sufficient background

knowledge for deciding which kind of data processing is most appropriate for a given sample. The final section illustrates the concepts of DeerAnalysis2006 on three samples that are representative of the main application scenarios. The determination of the shape of a narrow distance distribution is demonstrated for a shape-persistent biradical. This example also presents a critical test for the reliability of model-free extraction of the distance distribution, as the result can be compared with the wormlike chain model from polymer physics. Analysis of a broad distribution, again both with a model-free and model-based approach is illustrated for gold nanoparticles with a spin-labelled surface layer of alkyl chains. This example provides a test for excitation bandwidth correction. The use of experimentally derived background functions is tested on two singly labelled mutants of the membrane protein PutP and their mixture.

8.3. Materials and Methods

Syntheses of the shape-persistent biradical (Godt *et al.*, 2000) and of the spin-labelled gold nanoparticles (Chechik *et al.*, 2004) have been described elsewhere. The gold nanoparticle sample is based on butanethiol-coated particles (C4S) exchanged with the disulfide-modified nitroxide spin label N4. Spin-labelled mutants of the Na⁺/proline transporter PutP of *Escherichia coli* were prepared and reconstituted into proteoliposomes as described in Jeschke *et al.* (2004a), apart from the labelling procedure. This was done by adding washing buffer (50 mM KP_i, pH 8.0; 300 mM KCl; 10 mM imidazole; 10% glycerol [v/v]; 0.04% β -D-dodecylmaltoside [w/v]) containing 1 mM (1-oxyl-2,2,5,5,-tetramethylpyrroline-3-methyl)-methanethio-sulfonate (Toronto Research Chemicals, Canada) to the Ni²⁺-NTA column, followed by overnight incubation at 4 °C (hereafter we will refer to the spin-labeled residue as R1). Afterwards, unbound label was removed by washing the column with washing buffer.

Dipolar time evolution data were obtained at X-band frequencies (ca. 9.3-9.4 GHz) on a Bruker ElexSys 580 spectrometer equipped with a Bruker Flexline split-ring resonator ER 4118X_MS3 using the four-pulse DEER experiment $(\pi/2)_{v1} - \tau_1 - (\pi)_{v1} - t' - (\pi)_{v2} - \tau_1 + \tau_2 - t' - (\pi)_{v1} - \tau_2$ -echo (Pannier *et al.*, 2000b). The dipolar evolution time in this experiment is $t = t' - \tau_1$. A *phase cycle* [(+x)-(-x)] was applied to the $\pi/2$ pulse to cancel any receiver offset. Data were analysed only for $t > 0$. The resonator was overcoupled to $Q \approx 100$, the pump frequency ν_{pump} was set to the centre of the resonator dip and coincided with the maximum of the nitroxide EPR spectrum, while the observer frequency ν_{obs} was 65 MHz higher and coincided with the low-field local maximum of the spectrum. All experiments were performed with a pump pulse length of 12 ns and with observer pulse lengths of 32 ns at a temperature of 50 K. Proton modulation was averaged

by adding traces at 8 different τ_1 values, starting at $\tau_1(0) = 200$ ns and incrementing by $\Delta\tau_1 = 8$ ns. Deuterium modulation was averaged by adding 8 traces starting at $\tau_1(0) = 400$ ns and incrementing by $\Delta\tau_1 = 56$ ns.

Numerical computations of the modulation depth as a function of dipolar frequency were performed with a home-written MatLab program based on the EasySpin (Stoll and Schweiger, 2006) library. The program performs density operator computations of the whole four-pulse double electron-electron resonance (DEER) sequence, considering both the m.w. and static Hamiltonian during the pulses. The resonance offsets of both the observer and pumped spins are varied throughout the nitroxide line shape with proper weighting derived from an experimental echo-detected EPR spectrum. Details will be published elsewhere (J. Banham *et al.*, *unpublished*).

Algorithms for data processing are described below. The MatLab source codes are available at <http://www.mpip-mainz.mpg.de/~jeschke/distance.html>.

8.4. Theory

8.4.1. Separation into form factor and background factor

PELDOR experiments measure dipole-dipole couplings between an observer spin A and pumped spins B_k . For a given observer spin A, the signal $V(t)$ can be written as a product of the signals of all spin pairs involving A and any spin B_k in the sample,

$$V(t) = \prod_k V_{A,B_k}(t).$$

The signal can be observed until times t_{\max} (Fig. 8.1.) that are a small multiple of the phase relaxation time T_m of the A spin. Hence, we can neglect all B spins in the sample for which $V_{A,B_k} \equiv 1$ at times $t \leq t_{\max}$. Usually, this condition is fulfilled for B spins at distances larger than $r_{\max} = 40$ nm, so that only B spins within a sphere with radius r_{\max} need to be considered (Jeschke *et al.*, 2002). Many samples of interest are dilute solutions of spin-labelled nanoobjects, such as labelled biomacromolecules, synthetic macromolecules, or nanoparticles. It is thus convenient to further classify B spins into spins within the same nanoobject as spin A and spins in other nanoobjects. The signal can then be written as a product of a form factor $F(t)$ from B spins within the same nanoobject and a background factor $B(t)$ from spins in neighboring nanoobjects. Note that B spins at distances $r > r_{\max}$ do not contribute at all. This separation is analogous to the separation into a form factor and structure factor used in the analysis of scattering data (Lindner and Zemb, 2002). If the paramagnetic centers are homogeneously distributed rather than grouped into nanoobjects,

one has $F(t) \equiv 1$. In the limit of infinite dilution of the nanoobjects, one has $B(t) \equiv 1$. In all other cases, the experimentally detected dipolar evolution function $V(t)$ has to be separated into $F(t)$ and $B(t)$ before further data analysis.

Separation can be based on the fact that $F(t)$ decays to a constant value $1-\Delta$ within a time t_{decay} (Fig. 8.1.). The decay time t_{decay} is determined by the maximum distance between spins and the width of the spin-to-spin distance distribution within the nanoobject (Jeschke *et al.*, 2002). Here Δ is the modulation depth that depends on the number of spins within the nanoobject (Milov *et al.*, 1984; Jeschke *et al.*, 2004b). At times $t > t_{\text{decay}}$, the shape of the dipolar evolution function $V(t)$ is governed by the background factor $B(t)$.

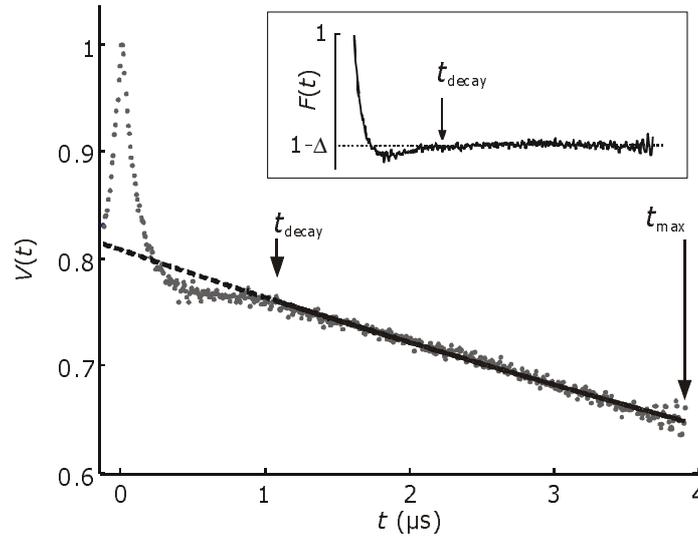


Figure 8.1: Separation into background factor $B(t)$ and form factor $F(t)$. Dots are experimental data points for surface-labelled gold nanoparticles. The solid line is the fitted background factor $B(t)$ assuming a homogeneous distribution of the nanoparticles in three dimensions, the dashed line is the back extrapolation of $B(t)$ to times shorter than t_{decay} . The inset shows the form factor $F(t)$ obtained by dividing the primary data $V(t)$ by $B(t)$ and renormalization according to Eq. (8.6.).

In the most simple case, where the nanoobjects or paramagnetic centers are distributed homogeneously throughout the observation sphere with radius r_{max} around spin A, and when longitudinal relaxation of A and B spins during the measurement can be neglected, $B(t)$ is an exponentially decaying function (Milov *et al.*, 1981; Milov *et al.*, 1984). When a dead-time-free experiment, such as four-pulse DEER (Pannier *et al.*, 2000b), is used, $V(t)$ can be normalized to $V(0) = 1$. Furthermore, by using a phase cycle for cancelling receiver offsets, it can be assured that $V(t \rightarrow \infty) = 0$. At sufficiently long times we then have

$$V(t > t_{\text{decay}}) = (1 - \Delta) \exp(-kt),$$

so that only two parameters Δ and k need to be fitted. Such fits tend to be stable even in the case $t_{\text{decay}} \approx t_{\text{max}}$ where the form factor does not fully decay within the observation time.

The background function differs from an exponential decay whenever the nanoobjects are confined to a domain which has an extension $d < r_{\text{max}}$ in at least one dimension. A typical

example are membrane proteins in liposomes. The lipid bilayer of the liposome is planar on the length scale of r_{\max} . Hence, to a first approximation the spatial distribution of membrane proteins in liposomes is a two-dimensional (2-D) rather than 3-D homogeneous distribution (Hilger *et al.*, 2005). For the general case of a D -D homogeneous distribution, the background function is given by Milov and Tsvetkov (1997)

$$B(t) = \exp(-kt^{D/3}). \quad (8.1.)$$

Back extrapolation of $B(t)$ from the fit range at long times t to shorter times becomes less reliable with decreasing dimensionality D , as the background functions then become steeper at short times compared to long times. Fits with dimensionality $D = 2$ are usually reliable if $t_{\text{decay}} < t_{\max}/4$. The dimensionality D itself can also be treated as a fit parameter and can be noninteger, corresponding to a fractal dimension of the distribution of nanoobjects. Such fits, however, tend to be unstable unless $t_{\text{decay}} \ll t_{\max}$.

Another complication results from the fact that the spins within the nanoobject may have a minimum distance, d_{\min} , to the surface of the object. As at the closest approach the nanoobjects can only have surface-to-surface contact but cannot overlap, contributions to $B(t)$ from distances shorter than $2d_{\min}$ are suppressed. Generally, any heterogeneous distribution of spins within the nanoobject causes slight deviations from the expression given in Eq. (8.1.). The exact form of $B(t)$ can then only be obtained experimentally. For measurements of doubly labelled biomacromolecules with site-directed spin labelling techniques (Hubbell and Altenbach, 1990), this problem can be solved by preparing the two corresponding singly labelled macromolecules for which $F(t) \equiv 1$, so that DEER measurements directly provide the background functions $B_1(t)$ and $B_2(t)$. The background function for the doubly labelled biomacromolecule can then be approximated by

$$B(t) = B_1(t)B_2(t). \quad (8.2.)$$

Such an approach is tested in Sect. 8.6.3..

8.4.2. Tikhonov regularization

Converting the form factor $F(t)$ into a distribution of distances $P(r)$ is a moderately ill-posed problem. This means that small changes in the input data $F(t)$, for instance due to noise or small errors in the separation of the form and background factor, can result in large changes in the output data $P(r)$. A number of different approaches for overcoming this problem have been tested (Jeschke *et al.*, 2002; Milov *et al.*, 2004a; Bowman *et al.*, 2004; Jeschke *et al.*, 2004b; Jeschke *et al.*, 2004c; Chiang *et al.*, 2005). Among them, approximate Pake transformation (APT) (Jeschke *et al.*, 2002) is computationally most efficient, but suffers from the drawback that it cannot incorporate the constraint $P(r) > 0$, which strongly

stabilizes the solution. Nevertheless, APT is useful for optimizing the separation into $F(t)$ and $B(t)$ (Sect. 8.5.2.) and for obtaining starting values for model fits (Sect. 8.5.3.). However, for the final computation of $P(r)$ an approach should be used that incorporates the constraint $P(r) > 0$. Our own experience with Hermite polynomial interpolation and Tikhonov regularization (Jeschke *et al.*, 2004b) suggests that different approaches with this constraint provide very similar results if an appropriate smoothing in the distance domain is selected. As a guideline for estimating the optimum degree of smoothing the L curve criterion appears to be better (Chiang *et al.*, 2005) than the self-consistency criterion previously used by us. In the following, we thus explain exclusively Tikhonov regularization with L curve computation, which is the main approach for model-free conversion of $F(t)$ to $P(r)$ in DeerAnalysis2006.

For a given distance distribution $P(r)$, the corresponding form factor can be computed in a straightforward manner by multiplication with a kernel function $K(t,r)$

$$F(t) = K(t,r)P(r).$$

This is a linear transformation which poses no particular problems. The inverse problem of computing $P(r)$ from $F(t)$ is ill-posed. Straightforward minimization of the root mean square deviation (RMSD) of KP from F typically leads to distance distributions that consist of a large number of very narrow peaks, even if the constraint $P(r) > 0$ is applied. Position and amplitudes of the narrow peaks depend strongly on noise. Characteristics of the whole distribution, such as the mean distance $\langle r \rangle$ and the standard deviation σ_r are stable (Jeschke *et al.*, 2004b). A physically reasonable solution for $P(r)$ can be found by requiring a certain smoothness of the distribution. This can be achieved by minimizing the target function

$$G_\alpha(P) = \|KP - F\|^2 + \alpha \left\| \frac{d^2}{dr^2} P \right\|^2, \quad (8.3.)$$

rather than only the RMSD. This approach is a Tikhonov regularization (Tikhonov, 1995). In Eq. (8.3.), $\rho = \|KP - F\|^2$ is the mean square deviation and the second term on the right hand side is the product of the regularization parameter α with the square norm η of the second derivative of the distance distribution. The regularization parameter α weighs the smoothness criterion (second term) with respect to the criterion of minimum RMSD (first term). For very small α , the simulated form factor replicates parts of the noise in the experimental form factor, which leads to a $P(r)$ with many unrealistically narrow peaks (Fig. 8.2.b, c). At the optimum α , the experimental $F(t)$ is fitted well and $P(r)$ is reasonably smooth (Fig. 8.2.d, e). For large α , the distribution is very smooth but the simulated form factor may deviate from the measured one (Fig. 8.2.f, g).

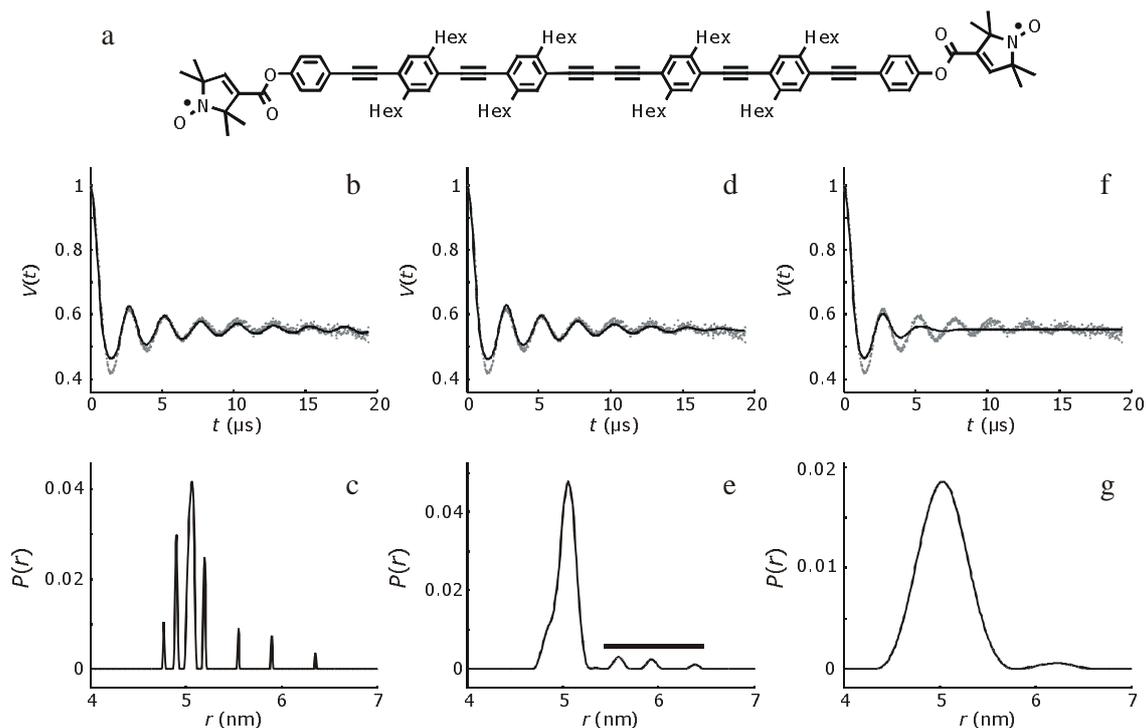


Figure 8.2: Tikhonov regularization with the additional constraint $P(r) > 0$ for a shape-persistent biradical. The dependence of the fit of the experimental form factor $F(t)$ and of the smoothness of the distance distribution $P(r)$ on the regularization parameter α is shown. Data were acquired for 12 h at $T = 50$ K on a sample of 0.4 mmol/liter biradical in perdeuterated *o*-terphenyl with a 12 ns pump pulse. a) Structure of the biradical. b) Experimental $F(t)$ (grey dots) and fit (black solid line) for $\alpha = 10^{-6}$. c) The RMSD of the fit is 0.000471. $P(r)$ for $\alpha = 10^{-6}$. d) Experimental $F(t)$ and fit for optimum $\alpha = \sqrt{10}$. e) RMSD: 0.000490. $P(r)$ for optimum $\alpha = \sqrt{10}$. The bar denotes artefact peaks as explained in Sect. 8.5.7.. f) Experimental $F(t)$ and fit for $\alpha = 10^3$. g) RMSD: 0.000762. The bar denotes artefact peaks as explained in the text. g) $P(r)$ for $\alpha = 10^3$.

The solution $P(r)$ thus depends strongly on the choice of the regularization parameter α . This choice can be guided by the L curve (Fig. 8.3.), which is a plot of $\ln \eta$ vs. $\ln \rho$ for the different values of α . For small α (left side of the L curve) the square norm η of the second derivative of $P(r)$ is large as the distance distribution is not smooth. Hence, $\log \eta$ is large and depends strongly on α . This causes a steeply decaying branch of the curve at small α , as the mean square deviation, ρ , is small and does not depend very strongly on α for weak smoothing. In contrast, for large α data are oversmoothed, which leads to a larger ρ term that increases strongly with further smoothing (right side of the L curve). In this situation, η depends only weakly on α , as $P(r)$ is already rather smooth. This behavior causes an only weakly decaying branch of the curve. The optimum compromise between smoothness of $P(r)$ and quality of the fit of $F(t)$ corresponds to the corner of the L curve (circles in Fig. 8.3.).

Comparison of Fig. 8.3.a and b demonstrates that an experimental L curve conforms best to these considerations when the distance distribution is narrow and the signal-to-noise ratio is high. For a narrow distance distribution, oversmoothing leads to an overdamping of the dipolar oscillation (Fig. 8.2.c) that in turn causes a strong increase in ρ if the dipolar oscillation at long times has a higher amplitude than the noise. In contrast, for broad distributions minor oversmoothing has only a small effect on ρ . Furthermore, assignment of

the optimum regularization parameter to the most pronounced corner of the L curve is a purely mathematical criterion that does not take into account any previous knowledge about the expected resolution of the distance distribution. It is therefore advisable to inspect the smoothness of the distance distribution and the quality of the fit of $F(t)$ for different values of α rather than to rely only on the L curve. For this purpose, DeerAnalysis2006 stores fits and distance distributions for all α values used in an L curve computation.

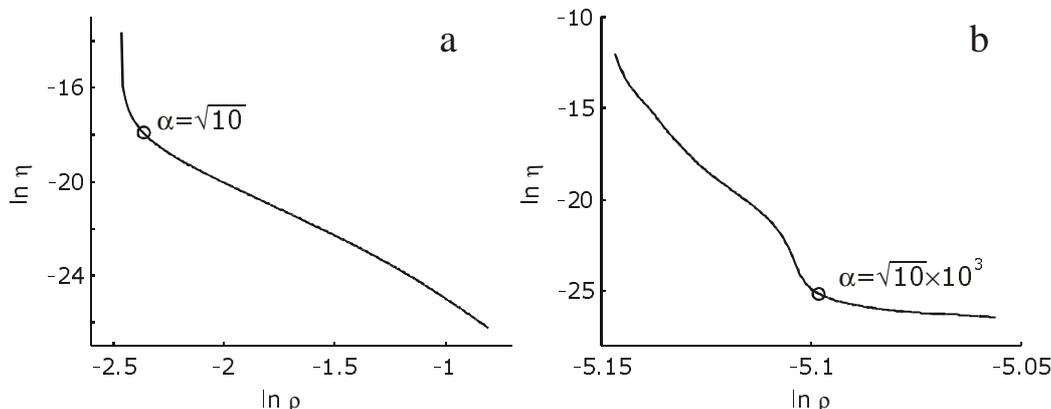


Figure 8.3.: Tikhonov L curves for two experimental data sets. a) Narrow distance distribution and high signal-to-noise ratio for a shape-persistent biradical in perdeuterated α -terphenyl. The circle denotes the optimum regularization parameter $\alpha = \sqrt{10}$. b) Broad distance distribution and moderate signal-to-noise ratio for surface-labelled gold nanoparticles. The circle denotes the optimum regularization parameter $\alpha = \sqrt{10} \cdot 10^3$.

8.4.3. Fitting of the form factor by a model

Tikhonov regularization involves compromise between distortions of $P(r)$ by noise artefacts for small α and artificial broadening for large α . Indeed, such a compromise needs to be made in any model-free computation of $P(r)$ from $F(t)$, so that all these approaches have difficulties with distance distributions that contain both narrow and broad contributions (Jeschke *et al.*, 2004b). Artificial broadening can be avoided and noise artefacts still be suppressed if there exists a good model for the shape of $P(r)$. For instance, a homogeneous distribution of spin labels on the surface of a sphere with diameter d_{sph} corresponds to a distance distribution $P(r)$ with $P(0) = 0$ that linearly increases for $0 \leq r \leq d_{\text{sph}}$ and is zero for $r > d_{\text{sph}}$. Furthermore, if one allows for a Gaussian distribution of sphere diameters $d_{\text{sph}} > 0$ with standard deviation σ_d , the resulting $P_{\text{model}}(r)$ is expected to be a good model for the spin-to-spin distance distribution in surface-labelled spherical nanoparticles. This model has only two free parameters, while the number of free parameters in a model-free computation is the number of r values at which $P(r)$ is defined. Fits of an appropriate model are thus much less susceptible to noise artefacts and can be performed by minimization of the RMSD between $K(r,t)P_{\text{model}}(r)$ and $F(t)$. Furthermore, there is no obvious problem with fitting a distribution $P_{\text{model}}(r)$ that contains both narrow and broad features.

The quality of a model for $P(r)$ can be judged by comparing the RMSD of the model-based fit to the RMSD for Tikhonov regularization with an optimum regularization parameter. If, and only if, the RMSD with Tikhonov regularization is significantly lower, differences between the model-free distance distribution and $P_{\text{model}}(r)$ can be discussed in terms of a deviation of the real nanoobject from the model.

8.4.4. Spin counting

The modulation depth Δ in a PELDOR measurement depends on the average number $\langle n \rangle$ of spins per nanoobject and on experimental parameters such as pulse lengths, observer and pump position in the spectrum, and the shape and width of the resonator mode. The experimental parameters can be conveniently expressed by a modulation depth factor λ , which is the modulation depth for a nanoobject that contains exactly two spins ($\lambda = \Delta$ for $\langle n \rangle = 2$). For modulation depth factors $\lambda \ll 1$, the total modulation depth Δ of the form factor $F(t)$ is then given by $\Delta = \lambda (\langle n \rangle - 1)$ (Milov *et al.*, 1984). Using an equation of Jeschke *et al.* (2004b), we can relax the condition $\lambda \ll 1$. We find

$$\lim_{t \rightarrow \infty} \ln F(t) = -\lambda (\langle n \rangle - 1)$$

and thus

$$1 - \Delta = \lim_{t \rightarrow \infty} F(t) = \exp[-\lambda (\langle n \rangle - 1)] .$$

Hence, the average number of spins per nanoobject can be determined from the modulation depth Δ of $F(t)$ at times $t > t_{\text{decay}}$ (inset in Fig. 8.1.) using the expression (Hilger *et al.*, 2005)

$$\langle n \rangle = 1 + \frac{\ln(1 - \Delta)}{\lambda} . \quad (8.4.)$$

In contrast to Hilger *et al.* (2005), we here define the modulation depth factor λ as a positive number. In addition to the parameters mentioned earlier, λ is influenced by the shape of the pump pulse and the position of the pump frequency ν_{pump} with respect to the center of the mode (Jeschke *et al.*, 2006). Because of the dependence on so many parameters that are hard to estimate precisely, reliable spin counting based on Eq. (8.4.) requires determination of λ by an experimental calibration using, for example, an authentic biradical such as the one shown in Fig. 8.2.a. In a calibration using five biradicals and one triradical, λ could be determined with a correlation coefficient of 0.9965 (Hilger *et al.*, 2005). The default value of λ in DeerAnalysis2006 is based on these data sets, which are distributed together with the program. It is strongly advisable to recalibrate spin counting with at least one biradical sample for each different combination of a spectrometer and probe head.

8.4.5. Excitation bandwidth correction

PELDOR measurements of distances between spins require that the excitation bandwidth of the pulses is at least of the order of magnitude of the dipole-dipole coupling. Contributions from distances shorter than 1.5 nm are largely suppressed with the usual m.w. pulse powers (Pannier *et al.*, 2000b; Martin *et al.*, 1998). Furthermore, it has been pointed out that, unless the excitation bandwidth is much larger than the dipole-dipole coupling, the static Hamiltonian needs to be considered during the m.w. pulses (Maryasov and Tsvetkov, 2000; Milov *et al.*, 2004b). The usual theoretical expressions for the dipolar evolution function (Milov *et al.*, 1981; Milov *et al.*, 1984; Jeschke *et al.*, 2004b; Jeschke, 2002b) are derived for ideal pulses and thus do not apply. Analytical expressions for the three-pulse PELDOR experiment could be obtained by keeping the assumptions that the observer pulses exclusively excite A spins, the pump pulse exclusively excites B spins, and only the secular part of the dipole-dipole Hamiltonian

$$\hat{H}_{\text{sec}} = \omega_{dd} \hat{S}_{A,z} \hat{S}_{B,z},$$

with

$$\omega_{dd}(r, \theta) = (3 \cos^2 \theta - 1) \frac{\mu_B^2 \mu_0 g_A g_B}{\hbar} \frac{1}{r^3},$$

needs to be considered (Maryasov and Tsvetkov, 2000). Here, θ is the angle between the spin-spin vector r and the direction of the static magnetic field, μ_B is the Bohr magneton, μ_0 the permeability of the vacuum, g_A and g_B are the g values of A and B spins, respectively, and r is the spin-to-spin distance. This approach neglects the pseudosecular dipole-dipole Hamiltonian

$$\hat{H}_{p\text{sec}} = -\frac{\omega_{dd}}{2} (\hat{S}_{A,x} \hat{S}_{B,x} + \hat{S}_{A,y} \hat{S}_{B,y}),$$

and thus requires that the difference between resonance frequencies of the A and B spins is much larger than the dipole-dipole coupling ω_{dd} (Jeschke *et al.*, 2000). This is equivalent to the assumption that the difference of the two m.w. frequencies is much larger than ω_{dd} . Closer analysis of the results reveals that with these assumptions the frequency of the dipolar oscillation remains unaffected by nonideality of the pulses but that the modulation depth factor λ becomes dependent on ω_{dd} (Maryasov and Tsvetkov, 2000; Milov *et al.*, 2004b). Furthermore, independent of the magnitude of ω_{dd} , the dipolar evolution function is time shifted by half the duration of the pump pulse.

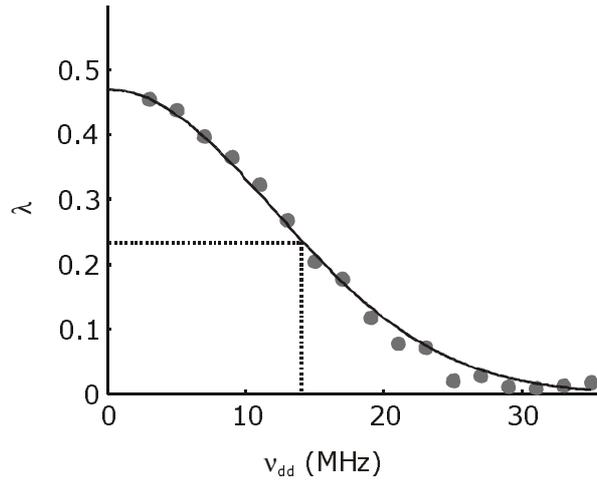


Figure 8.4.: Dependence of the modulation depth factor λ on the dipole-dipole coupling $\nu_{dd} = \omega_{dd}/(2\pi)$ for a four-pulse DEER sequence with observer pulses of 32 ns length and a pump pulse of 12 ns length (simulation). Full grey circles are results from numerical computations using the density operator formalism and the solid black line is a fit by a Gaussian with standard deviation of 17 MHz. The dotted lines denote the half width at half height of the excitation profile, corresponding to $\nu_{dd} = 14$ MHz and a distance of approximately 1.55 nm.

Despite the remaining assumptions, the full analytical expressions are rather extensive and slow down computation of the kernel function $K(t,r)$ considerably. Therefore, we studied the dependence of the dipolar oscillation frequency, modulation depth factor λ , and time shift of the oscillation on ω_{dd} for the four-pulse DEER experiment by numerical density operator computations taking into account both \hat{H}_{sec} and \hat{H}_{psec} and excitation of both A and B spins by both observer and pump pulses. Details of these computations will be published elsewhere (J. Banham *et al.*, *unpublished*). We find that within the whole range of dipolar couplings where the modulation depth factor λ differs significantly from zero, the dipolar oscillation frequency and the time shift are independent of ω_{dd} . Hence, the kernel function for ideal pulses needs to be corrected only by the dependence of λ on ω_{dd} . This dependence is plotted in Fig. 8.4. for observer pulse lengths of 32 ns and a pump pulse length of 12 ns, as used in this work. The numerical results can be fitted quite well by a Gaussian excitation profile of the whole pulse sequence

$$\lambda(\omega_{dd}) = \lambda_0 \exp\left(-\frac{\omega_{dd}^2}{\sigma_{exc}^2}\right),$$

where σ_{exc} is the excitation bandwidth parameter, which assumes a value of $2\pi \cdot 17$ MHz for the given pulse lengths. In terms of a half width at half height, the excitation bandwidth of four-pulse DEER with respect to dipolar couplings is thus 14 MHz for these pulse lengths. This corresponds to a distance of 1.55 nm. The general shape of the excitation profile remains the same for different pulse lengths, just σ_{exc} and the modulation depth factor in the limit of very small dipole-dipole couplings, λ_0 , vary. Knowledge of λ_0 is not required for the extraction of distance distributions. Rather, the constant contribution $1-\Delta$ can be subtracted

from $F(t)$ and the remaining pure dipolar modulation $F_{\text{mod}}(t)$ can be renormalized to $F_{\text{mod}}(0) = 1$. The excitation bandwidth corrected kernel function corresponding to F_{mod} is given by

$$K(t, r) = \int_0^1 \exp\left(-\frac{\omega_{dd}^2(r, \theta)}{\sigma_{exc}^2}\right) \cos[\omega_{dd}(r, \theta)t] d \cos \theta. \quad (8.5.)$$

8.5. Algorithms

8.5.1. Preprocessing

As noted above, computation of the distance distribution from the form factor is an ill-posed problem. Hence, experimental imperfections of the form factor have to be avoided as far as possible. This requires careful correction of the signal phase and of the time shift (Sect. 8.4.5.) of the dipolar evolution function and limitation of the form factor $F(t)$ to a time range $t < t_{\text{max}}$ where it can be measured free from artefacts and with sufficient signal-to-noise ratio. In the following these three preprocessing steps are discussed in turn.

Exact adjustment of the phase of the echo signal in PELDOR experiments can be difficult if the signal is weak. Furthermore, during the long accumulation times used for biomacromolecules, the signal phase may drift slightly. It is therefore advisable to record complex data by quadrature detection and perform phase correction after the measurement. This phase correction can be based on the fact that the imaginary part of the signal is zero after proper adjustment. Deviations may occur at times where the pump pulse is close to an observer pulse. Automatic phase correction in DeerAnalysis2006 is thus performed in the time range selected for background fitting. Directly after loading, this range is selected as the last three quarters of the data set. The optimum phase is determined by minimizing the RMSD of the imaginary part of the signal from zero. This presupposes that receiver offsets have been eliminated by $[(+x)-(-x)]$ phase cycling of the observer $\pi/2$ pulse. Such phase cycling is also required to ascertain that $V(t)$ approaches zero at infinite times, which is an assumption implicit in separation of form factor and background factor.

Determination of the time origin of the dipolar evolution, e.g., its zero time, can be based on the fact that, except for noise, the time origin corresponds to a maximum of the dipolar evolution function and that four-pulse DEER data should be symmetric about the time origin. After loading a data set and correcting phase, DeerAnalysis2006 determines the time t_0 where the real part is maximum. If $t_0 > 0$, the program considers a window with width t_0 that moves through the whole data set. The first moment of the real part of the data with respect to the center of this window is determined for all positions of the window. For a symmetric function the first moment is zero. Thus, the best guess for the time origin is the center of the

window with the minimum absolute value of the first moment. This algorithm is affected by any distortion of the data at early times that break symmetry about the time origin and by strong noise, although it is more stable with respect to noise than just assuming that the maximum of the real part is the time origin. It is good practice to determine zero time on a given spectrometer with given interpulse delays and pulse lengths for a standard sample with good signal-to-noise ratio and to use this experimental value in later measurements on weaker samples. Zero time is not affected by the choice of the interpulse delay τ_2 .

Dipolar evolution data $V(t)$ may be distorted at times near t_{\max} by approach of the pump pulse to the last observer pulse. Moreover, extraction of $F(t)$ involves dividing $V(t)$ by the background factor $B(t)$. If $B(t)$ has decayed to a small value at t_{\max} , this causes a poor signal-to-noise ratio of $F(t)$ near t_{\max} . In such situations, the reliability of the extracted distance distribution may be better when t_{\max} is redefined to a smaller value, i.e., if some data points at the end of $V(t)$ are cut off. The cutoff should be set at a time t_{cutoff} where the RMSD from the theoretical form factor exceeds a certain multiple of the RMSD in the undistorted and less noisy part of the data at early times. Determination of a suggested cutoff time is thus not strictly preprocessing - it already requires an estimate of $P(r)$ and of the corresponding theoretical form factor $F_{\text{theor}}(t)$. After loading a data set, an initial estimate for F_{theor} is provided by APT.

DeerAnalysis2006 determines the RMSDs for a moving 11-point window of data points and estimates the noise level at early times t as the minimum RMSD of any of these windows. The program suggests to cut off a data point at the end if the RMSD in a window centered at this point exceeds the noise level by at least a factor of 6. Then the next data point towards shorter times is tested in the same way until a data point is encountered for which the RMSD in the window is less than 6 times the noise level. After performing the suggested cutoff, $P(r)$ and $F_{\text{theor}}(t)$ change. This may result in a new suggested cutoff time t_{cutoff} that is even shorter. The cutoff correction thus needs to be performed iteratively until the suggestion does not change anymore. In some cases the suggestion may change once more after F_{theor} is improved by Tikhonov regularization or model-based fitting. It is then advisable to correct t_{cutoff} and repeat the fit.

8.5.2. Separation into form factor and background factor

The form factor $F(t)$ is computed in the interval from $t = 0$ (time origin) to t_{cutoff} as

$$F(t) = B(0) \frac{V(t)}{B(t)}, \quad (8.6.)$$

where multiplication with $B(0)$ assures that a normalized form factor, $F(0) = 1$, is obtained from a normalized dipolar evolution function. Separation of $V(t)$ into $F(t)$ and $B(t)$ requires an appropriate choice of the background function (Sect. 8.4.1.) and of the range in which this function is fitted to $V(t)$. The upper limit for the fit range is t_{cutoff} (Sect. 8.5.1.). The initial value for the lower limit is $t_{\text{max}}/4$, which assumes that $t_{\text{decay}} < t_{\text{max}}/4$. If t_{decay} can be estimated from the appearance of the primary data (Fig. 8.1.), and if it is smaller than $t_{\text{max}}/2$, the lower limit should be readjusted to t_{decay} . If an estimate of t_{decay} from the appearance of $V(t)$ is difficult or if t_{decay} approaches or exceeds t_{max} , an appropriate lower limit can be found by minimizing $|P(r)|$ at distances that are larger than the size of the nanoobjects.

This algorithm relies on the fact that for a proper separation, $P(r)$ should be identical zero at distances larger than the maximum extension of the nanoobject. For a $P(r)$ computed by APT (Jeschke *et al.*, 2002) undercorrection of the background contribution results in $P(r) > 0$ at such distances, while overcorrection results in $P(r) < 0$. DeerAnalysis2006 uses the sum of the last three data points of $|P(r)|$ as a figure of merit. The lower limit of the background fit range is set to the time where this figure of merit is minimum. This algorithm presupposes that the maximum distance r_{max} for which $P(r)$ can be computed by APT (Jeschke *et al.*, 2002) is larger than the maximum extension of the nanoobject. A good approximation is

$$r_{\text{max}} \approx \sqrt[3]{104t_{\text{max}}/\mu\text{snm}}.$$

For example, if data are measured to $t_{\text{max}} = 2 \mu\text{s}$, one cannot expect to obtain any information on distances larger than 5.9 nm.

8.5.3. Determining and using experimental background functions

If the spin labels inside the nanoobject have a minimum distance to its surface that exceeds 0.75 nm, the minimum label-to-label distance is larger than 1.5 nm. The background function $B(t)$ may then not fully conform to Eq. (8.1.). Likewise, if the nanoobjects are confined to a domain with at least one dimension longer than 1.5 nm but shorter than 40 nm, such a description by a homogeneous distribution with fractal dimension D is not necessarily appropriate. Such a case may be encountered for membrane proteins. Computation of $F(t)$ for the doubly labelled protein by Eq. (8.6.) is then best performed with an experimental background function $B_{\text{exp}}(t)$ which can be obtained by measurements on the corresponding two singly labelled proteins. However, direct use of the sum of the experimental dipolar evolution functions $V_1(t)$ and $V_2(t)$ of the singly labelled proteins is not appropriate for two reasons. First, noise from the measurements on the single mutants would be introduced into $F(t)$ of the double mutant. Second, any differences in protein concentration could not be

accounted for. It is thus necessary to parametrize the experimental background functions. For this purpose, a fit of the logarithm of $V(t)$ is most appropriate, as $\ln V(t)$ scales linearly with concentration. Furthermore, by taking the logarithm, Eq. (8.2.) simplifies to an addition of the coefficients in series expansions of $B_1(t)$ and $B_2(t)$. As a flexible background model function, we thus use

$$\ln B_i(t) = \sum_{k=1}^n a_{i,k} t^k, \quad (8.7.)$$

where the $a_{i,k}$ are real coefficients of the background polynomial and n is the order of this polynomial. The $a_{i,k}$ are determined by fitting the $V_i(t)$ for the singly labelled nanoobjects in the range from $t = 0$ to $t = t_{\max}$. Note that t_{\max} should be at least as long as in the intended measurement of the doubly or multiply labelled nanoobjects. The order n of the polynomial is adjusted so that the residual of the fit is a flat line except for noise. Background functions of two or more singly labelled nanoobjects are combined by adding the coefficients,

$$\ln B(t) = c \sum_i w_i \sum_{k=1}^n a_{i,k} t^k, \quad (8.8.)$$

with weighting factors w_i that account for any differences in concentration. The weighting factor of the first component is set to $w_1 = 1$. In the most simple case where all singly labelled objects are measured at the same concentration, one thus has $w_i = 1$ for all i . The parameter c in Eq. (8.8.) serves for fitting $B(t)$ to $V(t)$ of the doubly labelled or multiply labelled nanoobjects.

8.5.4. Concentration measurements

For a homogeneous 3-D distribution of nanoobjects ($D = 3$), parameter k in Eq. (8.1.) can be expressed as $k = xc$, where c is the local concentration of spins within a sphere with a radius of 20-40 nm around observer spin A. The factor x depends on the modulation depth factor λ and on the product of the g values of the observer and pumped spins. Because of the dependence of λ on many experimental variables (see Sect. 8.4.5.), x is best determined by calibration. The default value of DeerAnalysis2006 corresponds to a 12 ns pump pulse applied at the global maximum of a nitroxide spectrum. It was determined from a measurement on a solution of TEMPOL in toluene with a concentration of 2 mmol/liter at room temperature, corresponding to a concentration of 2.5 mmol/liter at 80 K. The calibration data set is distributed with the program. For measurements under the same conditions, the displayed density in the subpanel 'Background model' of the user interface of

DeerAnalysis2006 corresponds to a concentration in units of millimoles per liter. For any new combination of spectrometer and resonator a recalibration is strongly suggested.

8.5.5. Tikhonov regularization with L curve computation

The L curve for Tikhonov regularization can be computed very efficiently if Eq. (8.3.) is applied without further constraints (Chiang *et al.*, 2005; Hansen, 1992). DeerAnalysis2006 does use the constraint $P(r) > 0$, which strongly stabilizes the solution, hence a computationally somewhat less efficient approach was implemented. As an advantage, this approach provides not only the L curve, but also the distance distributions $P(r)$ and corresponding form factors $F_{\text{theor}}(t) = K(t,r)P(r)$ for all predefined values of the regularization parameter α . By default these values are the integer powers of 10 between 10^{-3} and 10^5 . This range and spacing of regularization parameters was found to be sufficient for all applications that we encountered so far. If necessary, the set of regularization parameters can be changed by editing the file *Lcurve_abscissa.dat* that is read on starting DeerAnalysis2006.

For these values of α , Tikhonov regularization according to Eq. (8.3.) and incorporating the constraint $P(r)$ is performed using the external program FTIKREG (Honerkamp and Weese, 1990; Weese, 1992) and the kernel function from Jeschke *et al.* (2004b). If excitation bandwidth correction is selected, the kernel function is modified according to Eq. (8.5.). The L curve is displayed in the form of discrete points corresponding to the regularization parameters α at which $P(r)$, $F_{\text{theor}}(t)$, ρ , and η were computed. The corner is located by forward linear extrapolation of the first three points, backward linear extrapolation of the last three points, and selection of the point that is closest to the intersection of the two lines. It is checked whether the curve makes an angle that deviates less from 180° than at the next point. If so, the next point is selected as the corner. For well-behaved, i.e., nicely L-shaped, curves this simple procedure usually finds the corner. In cases such as the one in Fig. 8.3.b, interactive correction is required. The user interface allows for scanning through the whole set of regularization parameters and observing how $P(r)$, the fit of $F_{\text{theor}}(t)$ to $F(t)$ and the RMSD of this fit vary. Even if the corner is not very pronounced, this usually allows one to make a reasonable estimate of the optimum regularization parameter.

8.5.6. Fitting of models for the distance distribution

DeerAnalysis2006 provides an interface for user-defined models of $P(r)$ that does not suppose any knowledge of the inner working of the program. A user-defined model can have

up to six parameters. Any parameter needs to be defined with a symbol, a default starting value and upper and lower limits. These items are given in comment lines of the function head (see Appendix). The model has to be defined as a Matlab function in the *models* subdirectory of DeerAnalysis2006. The function receives a vector of distance values $r0$ and a vector of parameters par . It has to compute a distance distribution $distr$ defined at the values of $r0$. The distribution must depend on all parameters provided in the vector par .

All functions defined in that way can be selected from the user interface. The program then simultaneously displays the $P(r)$ computed by APT and the $P_{\text{model}}(r)$ provided by the model function in output variable $distr$, first using the default values for the parameters (Fig. 8.5.). It also displays $F_{\text{theor}}(t)$ computed from $P_{\text{model}}(r)$. The starting parameters can then be adjusted interactively so that $F_{\text{theor}}(t)$ roughly fits $F(t)$. During subsequent least-squares fitting, individual parameters can be fixed if necessary.

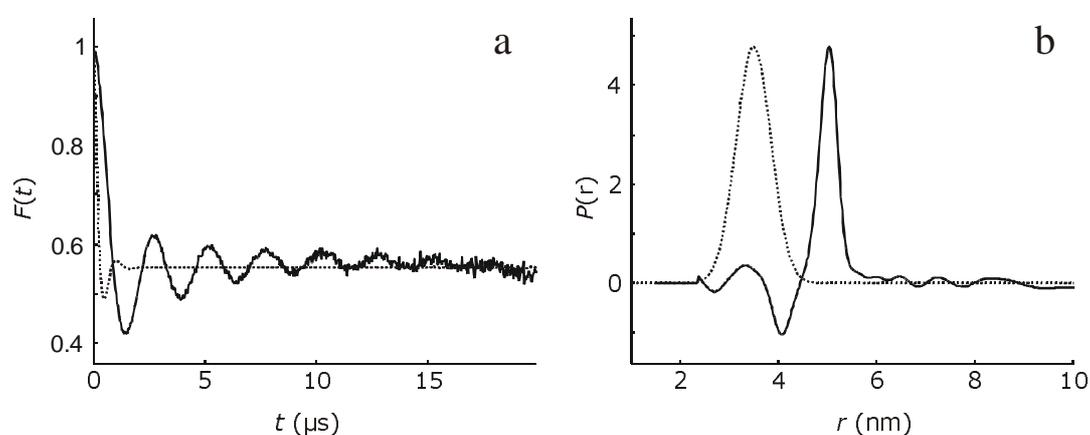


Figure 8.5.: Selection of starting values in model-based fitting. A single Gaussian peak with default mean distance $\langle r \rangle = 3.5$ nm and default standard deviation $s(r) = 0.5$ nm is assumed (see Appendix). Experimental data are for the biradical shown in Fig. 8.2.. a) Experimental form factor $F(t)$ (solid line) and theoretical form factor $F_{\text{theor}}(t)$ (dotted line). The faster oscillation of $F_{\text{theor}}(t)$ suggests to increase the starting value $\langle r \rangle$, while the faster damping of the oscillation suggests to decrease $s(r)$. b) Distance distribution $P(r)$ provided by APT (solid line) and distance distribution $P_{\text{model}}(r)$ provided by the model function (dotted line). Starting parameters can be optimized by improving the superposition of the two lines. Negative contributions in the $P(r)$ provided by APT are due to orientation selection and noise in the input data.

8.5.7. Suppression test for artefact peaks in the distance distribution

For ill-posed problems, such as the conversion of $F(t)$ into a distance distribution $P(r)$, noise and other distortions of the input data transform in a nonlinear and nonintuitive way to features in the output data. Thus, the question arises whether minor features, such as the small peaks between 5.4 and 6.6 nm in the distance distribution of the shape-persistent biradical (bar in Fig. 8.2.e), correspond to spin-to-spin distances actually present in the sample or are artefacts. This question can be answered by suppressing these peaks in $P(r)$, i.e., by setting $P(r) = 0$ in the interval from 5.4 to 6.6 nm and recomputing the theoretical form factor with this changed distribution. DeerAnalysis2006 allows for such suppression of peaks in a selected distance interval.

Whether or not the peaks are artefacts can best be estimated by changes in the RMSD of $F_{\text{theor}}(t)$ from $F(t)$ due to their suppression and by visual changes in the fit residual. In the case at hand, the root mean square value changes only slightly from 0.000479 for the original distribution to 0.000498 with suppression of the peaks. The reason for the appearance of these peaks becomes apparent by considering the change in the fit residual (arrows in Fig. 8.6.a). This change is confined to a range where the residual anyway deviates systematically from zero. The systematic deviation, in turn, is caused by orientation selection (Larsen and Singel, 1993; Schweiger and Jeschke, 2001) that suppresses the outer shoulders of the Pake pattern (arrows in Fig. 8.6.b), since orientations where the spin-spin vector is parallel to the external magnetic field are underrepresented. This example illustrates that it may be necessary to inspect different representations of the data, for instance, the form factor in time domain (Fig. 8.2.d), the corresponding dipolar spectrum (Fig. 8.6.b), and the fit residual with and without peak suppression (Fig. 8.6.a) to understand the reason for unexpected features in the distance distribution.

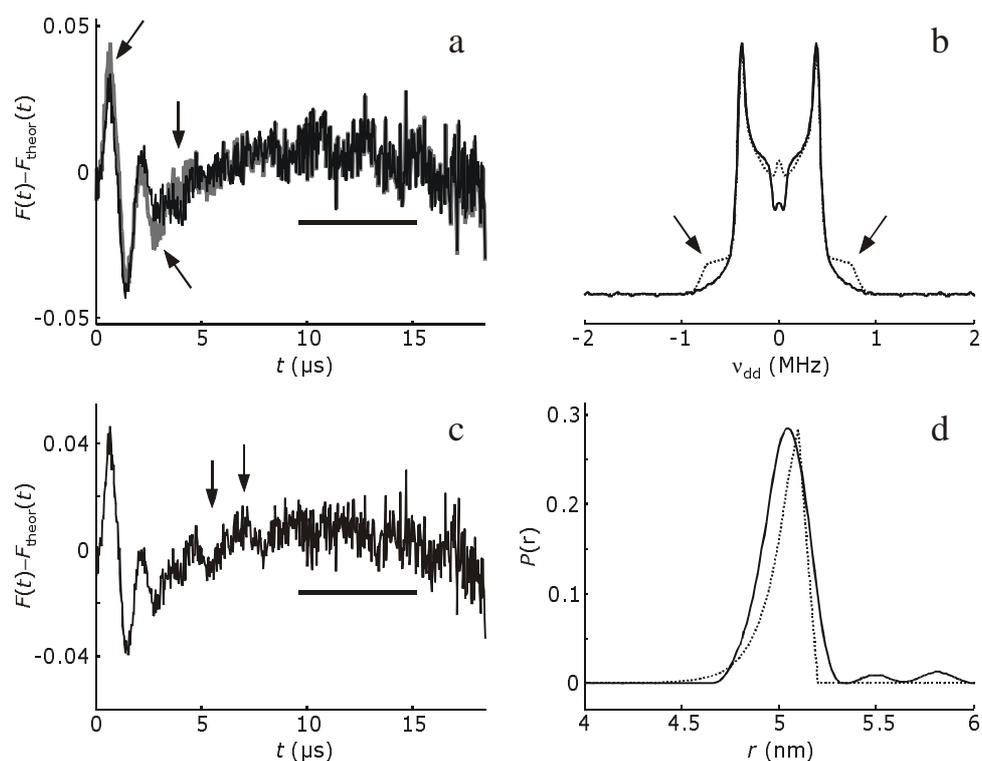


Figure 8.6.: Suppression test for artefacts and model-based fitting for a shape-persistent biradical (see also Fig. 8.2.). a) Fit residuals of the form factor from Tikhonov regularization (regularization parameter $\alpha = 1$) with (grey line) and without (black line) suppression of the small peaks marked by a bar in Fig. 8.2.e. The arrows mark significant changes in the residual. The bar denotes an additional range where some oscillatory deviations are apparent. b) Dipolar spectra obtained by Fourier transformation of the experimental form factor (solid line) and of the theoretical form factor without artefact suppression (dotted line). The arrows denote attenuation of the outer shoulders of the Pake pattern by orientation selection. c) Residual for a fit by the theoretical end-to-end distance distribution of a wormlike chain. Fit parameters are a contour length of 5.2 nm and a persistence length of 23.8 nm. The arrows denote stronger oscillatory deviations than for Tikhonov regularization at early times, the bar denotes weaker deviations at late times. d) Distance distributions obtained with Tikhonov regularization ($\alpha = \sqrt{10}$, solid line) and with the worm-like chain model (dotted line).

8.5.8. Dual display-scaling procedures for direct comparison of two data sets

In many applications it is of interest whether the distance distribution changes significantly on changing external conditions or the composition of the investigated system. Due to the nonintuitive influence of noise and other nonidealities in the dipolar evolution function $V(t)$ on the distance distribution $P(r)$, it is not advisable to decide this by merely comparing distance distributions. Rather, the primary data $V_1(t)$ and $V_2(t)$ of the two samples should be compared (Jeschke *et al.*, 2004b). In particular for biomacromolecules, where labelling efficiency may not be exactly reproducible, this introduces the problem that modulation depths, Δ , may differ between the two experiments. Small differences in the Δ may also arise from imperfections in experimental setup such as differences in the effective flip angle of the pump pulse or in the overcoupling of the resonator, which both influence the modulation depth factor λ . As pointed out earlier (Jeschke *et al.*, 2004b), the logarithm of $V(t)$ is proportional to λ , so that such differences can be corrected for by a scaling factor

$$f_\lambda = \frac{\sum_{k=1}^N [\ln V_1(t_k)]^2}{\sum_{k=1}^N \ln V_1(t_k) \ln V_2(t_k)} ,$$

where the summation runs over the data points. Superposition of two data sets with or without such modulation depth scaling is implemented in the dual display feature of the graphical user interface of DeerAnalysis2006. If modulation depth scaling is applied, both the primary data and the form factor of the second data set are scaled to the modulation depth of the first data set. Without modulation depth scaling, the distance distribution of the second data set is displayed with the same maximum amplitude as the one of the first data set. This guarantees that all features in both distributions are visible. With modulation depth scaling, the distance distributions are normalized so that they have the same integral, which in turn is proportional to the modulation depth of the first data set. This scaling is best suited for detecting changes in the distance distribution. Such changes are likely to be significant, if a difference between the modulation depth scaled primary data sets V_1 and $\exp(f_\lambda \ln V_2)$ is apparent.

8.6. Examples

8.6.1. Asymmetric distance distribution in shape-persistent biradicals

Shape-persistent biradicals were previously used for testing the precision of distance measurements by PELDOR (Jeschke *et al.*, 2004c; Godt *et al.*, 2000; Pannier *et al.*, 2000b; Martin *et al.*, 1998). The most precise of these measurements on the compounds with the

longest *p*-phenylenethynylene backbones (Jeschke *et al.*, 2004c) revealed that molecular force fields do not quantitatively account for the length of the backbone. This strongly suggests that current force field parameters cannot predict the flexibility of poly(*p*-phenyleneethynylene). A precise measurement of the end-to-end distance distribution of such compounds allows for quantification of the flexibility, as this distribution exhibits a distinct asymmetry and a shape and width that depends on the persistence length of the polymer backbone (Wilhelm and Frey, 1996). Such an asymmetry is apparent in the model-free analysis of the data for the biradical with a length of approximately 5 nm (Fig. 8.2.e).

An analytical expression for the end-to-end distance distribution of a semiflexible polymer has been derived on the basis of a continuous form of the worm-like chain model (Wilhelm and Frey, 1996). The distribution depends on only two parameters, the length L of the fully stretched chain, which is the contour length of the polymer, and the persistence length l_p that quantifies flexibility. Here we use these expressions to illustrate the combined use of model-free and model-based data analysis.

A fit of the form factor of the biradical by the continuous worm-like chain model provides a distance distribution with similar asymmetry but less general broadening than model-free data analysis by Tikhonov regularization with an optimized regularization parameter of $\sqrt{10}$ (Fig. 8.6.d). The RMSD increases from 0.000490 for Tikhonov regularization to 0.000515 for the worm-like chain model. Closer inspection of the fit residuals (Fig. 8.6.a, c) reveals that at early times the model-based fit displays more oscillatory deviations (arrows in Fig. 8.6.c), while at late times stronger deviations are found in the model-free fit (bars in Fig. 8.6.a, c). The deviations at late times suggest that Tikhonov regularization slightly oversmooths the distance distribution. On the other hand, the deviation of the model-based approach at early times suggests that the distribution does not exactly have the shape implied by the model. This is not surprising, as we have neglected the conformational distribution of the nitroxide endlabellers, which introduces additional broadening of the distance distribution. The significant deviations for the simple model suggest that this conformational distribution should be taken into account in an improved model.

8.6.2. Broad distance distribution in spin-labelled gold nanoparticles

Broad distance distributions present a more critical test for data analysis procedures (Jeschke *et al.*, 2004b). Surface-labelled gold nanoparticles are particularly well suited for such a test, as a good model for the distance distribution exists (Sect. 8.4.3.). The surface labels (Fig. 8.7.a) are supposed to be distributed randomly on the surface. Spin counting indicates an average number of $\langle n \rangle = 1.4$ labels per particle. Much lower degree of labelling

leads to a low modulation depth and, hence, to a poor signal-to-noise ratio. For a much higher degree of labelling, relaxation time is expected to decrease. The number of labels per particle is expected to follow a Poisson distribution. Particles with no label at all do not contribute to $V(t)$, particles with one label contribute only to $B(t)$. Particles with two (Fig. 8.7.b) or more labels contribute to both $B(t)$ and $F(t)$. As the form factor $F(t)$ corresponds to the pair distribution function, it is the same for two and more than two labels, except for differences in modulation depth. The model for the distribution of label-to-label distances r (Fig. 8.7.b) is based on the assumption of spherical particles and a Gaussian distribution of the effective sphere diameter due to a combined distribution of particle diameters and conformations of the label-carrying surface-attached alkyl chains. Such a model has two parameters, the mean diameter $\langle d_s \rangle$ of the sphere and the standard deviation σ_d of this diameter. The distance distribution extends from $r = 0$ to the effective sphere diameter, so that the data also allow for testing whether excitation bandwidth correction can recover suppressed contributions to the distance distribution at short distances.

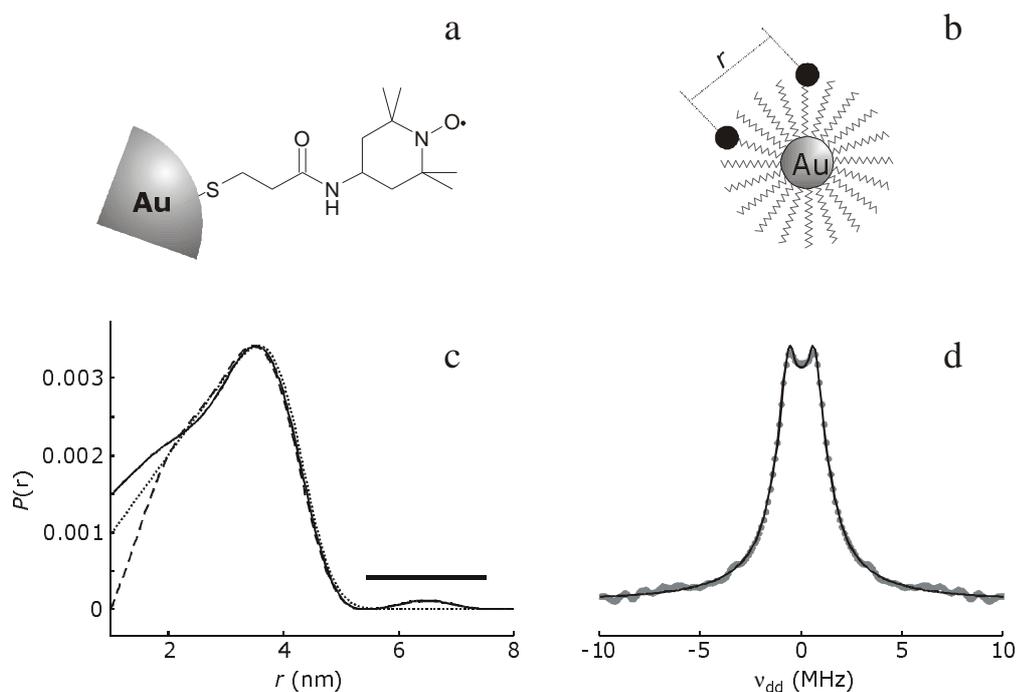


Figure 8.7.: Model-free and model-based fitting of the distance distribution for surface-labelled gold nanoparticles with and without excitation bandwidth correction. a) Structure of the surface label. b) Schematic structure of the particle, assuming a case with two surface labels. c) Distance distribution $P(r)$ obtained with Tikhonov regularization with $\alpha = \sqrt{10} \cdot 10^3$ without (dashed line) and with (solid line) excitation bandwidth correction according to Eq. (8.5.), assuming $\sigma_{\text{exc}} = 2\pi \cdot 17$ MHz. The dotted line is the distance distribution obtained with excitation bandwidth correction for a model of labels homogeneously distributed on the surface of a sphere with a normally distributed diameter d_s . Parameters of the normal distribution are mean diameter $\langle d_s \rangle = 4.18$ nm and standard deviation $\sigma_d = 0.66$ nm. d) Experimental dipolar spectrum (grey dots) and fit by the sphere model (black line).

As can be seen in Fig. 8.7.c, the distance distributions obtained by Tikhonov regularization with (solid line) and without (dashed line) excitation bandwidth correction indeed differ for distances shorter than about 2 nm. The RMSD of the fit of the form factor is slightly better without (0.000234) than with the correction (0.000236); however, the difference

is hardly significant. The form factor can be fitted well without the correction, but this is expected to cause a systematic distortion of the distance distribution. Comparison with the model-based fit with excitation bandwidth correction (dotted line in Fig. 8.7.c) reveals good agreement of all approaches at distances larger than 2 nm. The model-based fit has an RMSD of 0.000240 from the experimental form factor. As revealed by the superposition of the experimental and simulated dipolar spectrum (Fig. 8.7.d), this corresponds to a very good fit quality. The fit residual is white noise without any oscillatory deviations (data not shown). The model-based fit can thus be used to discuss the performance of excitation bandwidth correction.

As expected, neglect of the excitation bandwidth correction in Tikhonov regularization (dashed line in Fig. 8.7.c) leads to an underestimate of the contribution of short distances with respect to the model-based fit (dotted line). Somewhat unexpectedly, use of the kernel function given in Eq. (8.5.) with the numerically determined excitation bandwidth $\sigma_{\text{exc}} = 2\pi \cdot 17$ MHz (solid line) appears to overcompensate this suppression. We checked that the exact shape of the distance distribution cannot be recovered by adjusting σ_{exc} . We tentatively assign this overcompensation to contributions from distances below 1.3 nm, where the assumptions of a constant oscillation frequency and a constant time shift fail. A more detailed account will be given elsewhere (J. Banham *et al.*, *unpublished*). Despite this shortcoming, excitation bandwidth correction does improve agreement with the model-based distribution. In any case, a significant difference between results with and without correction points to a distortion of the uncorrected result.

Not unexpectedly, model-based fitting is much more robust with respect to limited excitation bandwidth than Tikhonov regularization. In the case at hand, $F(t)$ is dominated by contributions from distances larger than 2 nm, so that a two-parameter-fit is hardly affected by these effects. With excitation bandwidth correction, we find $\langle d_s \rangle = 4.18$ nm and $\sigma_d = 0.66$ nm, without such correction $\langle d_s \rangle = 4.20$ nm and $\sigma_d = 0.63$ nm. The RMSD of the fit of $F(t)$ does not change. The values for the sphere diameter and its standard deviation can be compared with previous work by transmission electron microscopy (TEM) and continuous-wave (CW) EPR (Ionita *et al.*, 2005). For the diameter of the gold core, analysis of TEM pictures gives an estimate of 1.99 ± 0.42 nm, while the average distance between the label and its point of attachment to the surface is at least 0.8 nm according to CW EPR. For a fully stretched label, this length would be 1.2 nm. From these values, a mean sphere diameter $\langle d_s \rangle$ between 3.6 and at most 4.4 nm and a standard deviation between 0.5 and 0.7 nm would be expected. The values from DEER fall within this range. In fact, the TEM estimate of the gold core diameter is probably the most significant error source in this comparison.

8.6.3. Derivation and test of an experimental background function for a membrane protein

As explained in Sect. 8.5.3., for a spin label inside a membrane protein, the intermolecular background function may not conform to Eq. (8.1.). In Fig. 8.8. this is illustrated for a methane thiosulfonate (MTSSL) spin label attached to residue 41 in PutP after mutation of this residue from serine to cysteine. For this singly labelled membrane protein, a fit by a spatially unconstrained homogeneous distribution ($D = 3$) is particularly poor (Fig. 8.8.a), corresponding to an RMSD $\|V(t) - B(t)\| = 0.001880$. Considering that the protein is reconstituted into proteoliposomes with a radius much larger than the length scale of pulsed ELDOR distance measurements, confinement to two dimensions ($D = 2$) appears likely. Indeed, this does improve the fit so that the RMSD decreases to 0.000823, but significant deviations are still apparent (Fig. 8.8.b). Relaxing the strictness of the 2-D confinement by fitting both k and D in Eq. (8.1.) improves the RMSD to 0.000743 at the best-fit fractional dimension of 2.17, but the deviations do not vanish completely (data not shown). This becomes apparent when $\ln V(t)$ is fitted by a fifth-order polynomial according to Eq. (8.7.). The RMSD again decreases drastically to 0.000480 and the fit is now almost perfect (Fig. 8.8.c). In fact, even a fit by a second-order polynomial with only one more parameter than the fit for a fractal dimension improves the root mean square value already to 0.000490 (data not shown).

Similar observations are made for a label attached to residue 423 in PutP. Here the RMSD decreases from 0.001715 for $D = 3$ via 0.000664 for $D = 2$ and 0.000648 for the best-fit fractal dimension $D = 2.08$ to 0.000565 for the fifth-order polynomial. The improvement by the polynomial compared to $D = 2$ is not drastic in this case, which probably means that this site is closer to the surface of the protein.

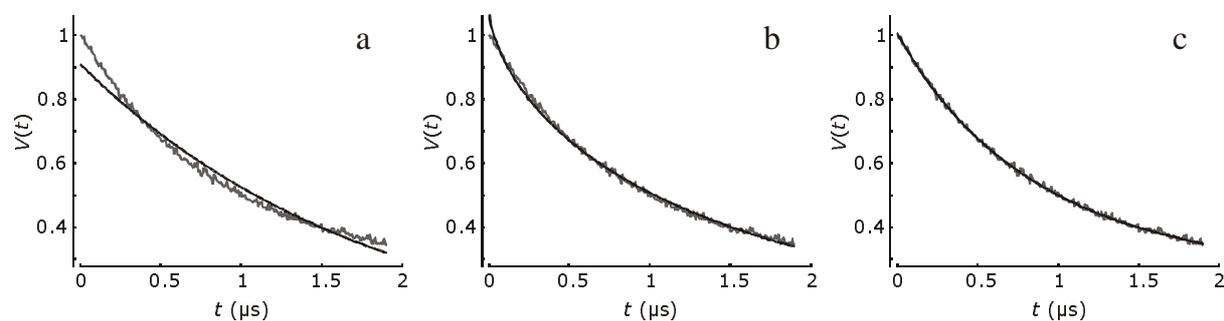


Figure 8.8.: Different background fits for the singly labelled mutant S41R1 of PutP. a) Experimental four-pulse DEER data (grey line) and fit by a background function (black line) assuming a homogeneous distribution of spin labels in three dimensions. b) Experimental data (grey line) and fit (black line) assuming a homogeneous distribution confined to two dimensions. c) Experimental data (grey line) and polynomial background fit (black line) based on Eq. (8.7.) with polynomial order $n = 5$.

The quality of the experimentally derived background model can be tested on a mixture of both singly labelled proteins. In this case, the RMSDs with analytical background models are 0.000645 for $D = 2$ and 0.000579 for the best-fit fractal dimension $D = 2.15$.

Using the background polynomial obtained with the spin label at residue 41 improves the root mean square value to 0.000409, while using the polynomial obtained with the spin label at residue 423 improves the root mean square value to 0.000430. Note that these fits with the experimental background parameter have only a single free fit parameter, namely parameter c in Eq. (8.8.), as the $a_{i,k}$ were determined before on the single mutants. Using the sum of both experimental background polynomials with $w_1 = w_2 = 1$, the fit again improves slightly to an RMSD of 0.000403. Again, only c is a fit parameter. As $B(t)$ is well reproduced for the mixture of the single mutants, one can expect that the approach is well suited for the corresponding double mutant. Indeed, we find that fit parameter c depends only slightly on the selected range for the background fit, i.e., the procedure of fitting $B(t)$ at later times and extrapolating it back to early times works well for the mixture of single mutants. The same procedure can be applied for nanoobjects that can be prepared both with a single label and with multiple labels. It is not applicable to samples with statistical labelling, as for instance the surface-labelled gold nanoparticles in Sect. 8.6.2..

8.7. Appendix

Example for a user-defined model function for the distance distribution

```
function distr=Gaussian(r0,par),
```

```
%
```

```
% Model library of DeerAnalysis2006: Gaussian
```

```
%
```

```
% single Gaussian peak with mean distance <r> and width
```

```
% (standard deviation) s(r)
```

```
% (c) G. Jeschke, 2006
```

```
%
```

```
% PARAMETERS
```

```
% name symbol default lower bound upper bound
```

```
% par(1) <r> 3.5 1.5 10
```

```
% par(2) s(r) 0.5 0.05 5
```

```
gauss0=(r0-par(1)*ones(size(r0)))/par(2);
```

```
distr=exp(-gauss0.^2);
```

8.8. Acknowledgments

We thank T. Roths and J. Honerkamp from Freiburger Materialforschungszentrum for kindly providing the source code of the program FTIKREG for Tikhonov regularization.

8.9. References

- Banham, J. E., Timmel, C. R., Abbott, R. J. M., Lea, S. M., and Jeschke, G.** (2006) The characterization of weak protein-protein interactions: evidence from DEER for the trimerization of a von Willebrand Factor A domain in solution. *Angew. Chem. Int. Ed.* 118: 1058-1061.
- Bennati, M., Robblee, J. H., Mugnaini, V., Stubbe, J., Freed, J. H., and Borbat, P.** (2005) EPR distance measurements support a model for long-range radical initiation in *E. coli* ribonucleotide reductase. *J. Am. Chem. Soc.* 127: 15014-15015.
- Borovykh, I. V., Ceola, S., Gajula, P., Gast, P., Steinhoff, H.-J., and Huber, M.** (2006) Distance between a native cofactor and a spin label in the reaction centre of *Rhodobacter sphaeroides* by a two-frequency pulsed electron paramagnetic resonance method and molecular dynamics simulations. *J. Magn. Reson.* 180: 178-185.
- Bowman, M. K., Maryasov, A. G., Kim, N., and DeRose, V. J.** (2004) Visualization of distance distribution from pulsed double electron-electron resonance data. *Appl. Magn. Reson.* 26: 23-39.
- Chechik, V., Wellsted, H. J., Korte, A., Gilbert, B. C., Caldararu, H., Ionita, P., and Caragheorgheopol A.** (2004) Spin-labelled Au nanoparticles. *Faraday Discuss.* 125: 279-291.
- Chiang, W.-Y., Borbat, P. P., and Freed, J. H.** (2005) The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J. Magn. Reson.* 172: 279-295.
- Elsässer, C., Brecht, M., and Bittl, R.** (2002) Pulsed electron-electron double resonance on multinuclear metal clusters: assignment of spin projection factors based on the dipolar interaction. *J. Am. Chem. Soc.* 124: 12606-12611.
- Fajer, P. G.** (2005) Site directed spin labelling and pulsed dipolar electron paramagnetic resonance (double electron-electron resonance) of force activation in muscle. *J. Phys.: Condens. Matter* 17: 1459-1469.
- Godt, A., Franzen, C., Veit, S., Enkelmann, V., Pannier, M., and Jeschke, G.** (2000) EPR probes with well defined, long distances between two or three unpaired electrons. *J. Org. Chem.* 65: 7575-7582.
- Hansen, P.C.** (1992) Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Rev.* 34: 561-580.
- Hilger, D., Jung, H., Padan, E., Wegener, C., Vogel, K.-P., Steinhoff, H.-J., and Jeschke, G.** (2005) Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na⁺/H⁺ antiporter of *E. coli*. *Biophys. J.* 89: 1328-1338.
- Hinderberger, D., Spiess, H. W., and Jeschke, G.** (2004a) Dynamics, site binding, and distribution of counterions in polyelectrolyte solutions studied by EPR spectroscopy. *J. Phys. Chem. B* 108: 3698-3704

- Hinderberger, D., Schmelz, O., Rehahn, M., and Jeschke, G.** (2004b) Electrostatic site attachment of divalent counterions to rodlike ruthenium(II) coordination polymers characterized by EPR spectroscopy. *Angew. Chem. Int. Ed.* 43: 4616-4621.
- Honerkamp, J., and Weese, J.** (1990) Tikhonovs regularization method for ill-posed problems. *Cont. Mech. Therm.* 2: 17-30.
- Hubbell, W. L., and Altenbach, C.** (1994) Investigation of structure and dynamics in membrane proteins using site-directed spin labeling. *Curr. Opin. Struct. Biol.* 4: 566-573.
- Ionita, P., Caragheorgheopol, A., Gilbert, B. C., and Chechik, V.** (2005) Dipole-dipole interactions in spin-labeled au nanoparticles as a measure of interspin distances. *J. Phys. Chem. B* 109: 3734-3742.
- Jeschke, G., Pannier, M., Godt, A., and Spiess, H. W.** (2000) Dipolar spectroscopy and spin alignment in electron paramagnetic resonance. *Chem. Phys. Lett.* 331: 243-252.
- Jeschke, G.** (2002a) Determination of the nanostructure of polymer materials by electron paramagnetic resonance spectroscopy. *Macromol. Rapid Commun.* 23: 227-246.
- Jeschke, G.** (2002b) Distance measurements in the nanometer range by pulse EPR. *ChemPhysChem* 3: 927-932.
- Jeschke, G., Koch, A., Jonas, U., and Godt, A.** (2002) Direct conversion of EPR dipolar time evolution data to distance distributions. *J. Magn. Reson.* 155: 72-82.
- Jeschke, G., and Godt, A.** (2003) Co-conformational distribution of nanosized [2]catenanes determined by pulse EPR measurements. *ChemPhysChem* 4: 1328-1334.
- Jeschke, G., Wegener, C., Nietschke, M., Jung, H., and Steinhoff, H.-J.** (2004a) Interresidual distance determination by four-pulse double electron-electron resonance in an integral membrane protein: the Na⁺/proline transporter PutP of *Escherichia coli*. *Biophys. J.* 86: 2551-2557.
- Jeschke, G., Panek, G., Godt, A., Bender, A., and Paulsen, H.** (2004b) Data analysis procedures for pulse ELDOR measurements of broad distance distributions. *Appl. Magn. Reson.* 26: 223-244.
- Jeschke, G., Bender, A., Paulsen, H., Zimmermann, H., and Godt, A.** (2004c) Sensitivity enhancement in pulse EPR distance measurements. *J. Magn. Reson.* 169: 1-12.
- Jeschke, G., Bender, A., Schweikardt, T., Panek, G., Decker, H., and Paulsen, H.** (2005) Localization of the N-terminal domain in light-harvesting chlorophyll a/b protein by EPR measurements. *J. Biol. Chem.* 280: 18623-18630.
- Jeschke, G., Zimmermann, H., and Godt, A.** (2006) Isotope selection in distance measurements between nitroxides. *J. Magn. Reson.* 180: 137-146.
- Kay, C. W. M., Elsässer, C., Bittl, R., Farrell, S. R., and Thorpe, C.** (2006) Determination of the distance between the two neutral flavin radicals in augmenter of liver regeneration by pulsed ELDOR. *J. Am. Chem. Soc.* 128: 76-77.

- Larsen, R. G., and Singel, D. J.** (1993) Double electron-electron resonance spin-echo modulation spectroscopic measurement of electron-spin pair separations in orientationally disordered solids. *J. Chem. Phys.* 98: 5134-5146.
- Lindner, P., and Zemb T.** (2002) *Neutrons, X-rays and Light: Scattering methods applied to soft condensed matter*. North-Holland, Amsterdam.
- Martin, R. E., Pannier, M., Diederich, F., Gramlich, V., Hubrich, M., and Spiess, H. W.** (1998) Determination of end-to-end distances in a series of TEMPO diradicals of up to 2.8 nm length with a new four-pulse double electron electron resonance experiment. *Angew. Chem. Int. Ed.* 37: 2833-2837.
- Maryasov, A. G., and Tsvetkov, Y. D.** (2000) Formation of the pulsed electron-electron double resonance signal in the case of a finite amplitude of microwave fields. *Appl. Magn. Reson.* 18: 583-605.
- Milov, A. D., Salikhov, K. M., and Shirov, M. D.** (1981) Application of ELDOR in electron spin echo for paramagnetic center space distribution in solids. *Fiz. Tverd. Tela (Leningrad)* 23: 975-982.
- Milov, A. D., Ponomarev, A. B., and Tsvetkov, Y. D.** (1984) Electron-electron double resonance in electron spin echo: model biradical systems and the sensitized photolysis of decalin. *Chem. Phys. Lett.* 110: 67-72.
- Milov, A. D., and Tsvetkov, Y. D.** (1997) Double electron-electron resonance in electron spin echo: conformations of spin labeled poly-4-vinylpyridine in glassy solutions. *Appl. Magn. Reson.* 12: 495-504.
- Milov, A. D., Tsvetkov, Y. D., Formaggio, F., Crisma, M., Toniolo, C., and Raap, J.** (2000) Self-assembling properties of membrane-modifying peptides studied by PELDOR and CW-ESR spectroscopies. *J. Am. Chem. Soc.* 122: 3843-3848.
- Milov, A. D., Tsvetkov, Y. D., Formaggio, F., Oancea, S., Toniolo, C., and Raap, J.** (2004a) Solvent effect on the distance distribution between spin labels in aggregated spin labeled trichogin GA IV dimer peptides as studied by pulsed electron-electron double resonance. *Phys. Chem. Chem. Phys.* 6: 3596-3603.
- Milov, A. D., Naumov, B. D., and Tsvetkov, Y. D.** (2004b) The effect of microwave pulse duration on the distance distribution function between spin labels obtained by PELDOR data analysis. *Appl. Magn. Reson.* 26: 587-599.
- Nakamura, M., Ueki, S., Hara, H., and Arata, T.** (2005) Calcium structural transition of human cardiac troponin C in reconstituted muscle fibres as studied by site-directed spin labelling. *J. Mol. Biol.* 348: 127-137.
- Pannier, M., Schädler, V., Schöps, M., Wiesner, U., Jeschke, G., and Spiess, H. W.** (2000a) Determination of ion cluster sizes and cluster-to-cluster distances in ionomers by four-pulse double electron-electron resonance spectroscopy. *Macromolecules* 33: 7812-7818.
- Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W.** (2000b) Dead-time free measurement of dipole-dipole interactions between electron spins. *J. Magn. Reson.* 142: 331-340.

- Pannier, M., Schöps, M., Schädler, V., Wiesner, U., Jeschke, G., and Spiess, H. W.** (2001) Characterization of ionic clusters in different ionically functionalized diblock copolymers by CW EPR and four-pulse double electron-electron resonance. *Macromolecules* 34: 5555-5560.
- Persson, M., Harbridge, J. R., Hammarstrom, P., Mitri, R., Martensson, L.-G., Carlsson, U., Eaton, G. R., and Eaton, S. S.** (2001) Comparison of electron paramagnetic resonance methods to determine distances between spin labels on human carbonic anhydrase II. *Biophys. J.* 80: 2886-2897.
- Sale, K., Song, L., Liu, Y.-S., Perozo, E., and Fajer, P.** (2005) Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *J. Am. Chem. Soc.* 127: 9334-9335.
- Schiemann, O., Piton, N., Mu, Y. G., Stock, G., Engels, J. W., and Prisner, T. F.** (2004) A PELDOR-based nanometer distance ruler for oligonucleotides. *J. Am. Chem. Soc.* 126: 5722-5729.
- Schweiger, A., and Jeschke G.** (2001) *Principles of pulse electron paramagnetic resonance*, Oxford University Press, Oxford.
- Stoll, S., and Schweiger, A.** (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* 178: 42-55.
- Tikhonov, A. N.** (1995) *Numerical methods for the solution of ill-posed problems*, Kluwer Academic Publishers, Dordrecht, Boston.
- van Amsterdam, I. M. C., Ubbink, M., Canters, G. W., and Huber, M.** (2003) Measurement of a Cu-Cu distance of 26 Å by a pulsed EPR method. *Angew. Chem. Int. Ed.* 42: 62-64.
- Weese, J.** (1992) A reliable and fast method for the solution of Fredholm integral equations of the first kind based on Tikhonov regularization. *Comput. Phys. Commun.* 69: 99-111.
- Wilhelm, J., and Frey, E.** (1996) Radial distribution function of semiflexible polymers. *Phys. Rev. Lett.* 77: 2581-2584.
- Zhou, Z., DeSensi, S. C., Stein, R. A., Brandon, S., Dixit, M., McArdle, E. J., Warren, E. M., Kroh, H. K., Song, L., Cobb, C. E., Hustedt, E. J., and Beth, A. H.** (2005) Solution structure of the cytoplasmic domain of erythrocyte membrane band 3 determined by site-directed spin labeling. *Biochemistry* 44: 15115-15128.

CHAPTER 9**Discussion**

Secondary transporters are found in all species throughout the kingdoms of life (Ren and Paulsen, 2005). They catalyze the translocation of a variety of compounds and play vital roles in cellular physiology. Their dysfunction is implicated in several human diseases and many secondary transporters are important pharmaceutical targets (Dahl *et al.*, 2002). In the present study, the Na⁺/proline symporter PutP and the Na⁺/H⁺ antiporter NhaA were used as model systems to gain insights into the structure, function, and dynamics of ion-coupled secondary transporters. For structural characterization of these proteins, new approaches were developed based on DEER distance measurements. These allow for the determination of secondary, tertiary, and quaternary structures of membrane proteins in the native environment of a lipid bilayer. Furthermore, a new software package is presented that provides access to a comprehensive set of tools for pulsed ELDOR data analysis.

9.1. Functional role and structural dynamics of transmembrane domain IX of PutP

The Na⁺/proline transporter PutP of *E. coli* is composed of 13 putative transmembrane domains of which TMs II and IX harbor the highest density of conserved residues (suppl. Fig. 5.1.). Previous analyses of structure-function relationships of TM II suggest that this region of PutP is particularly important for Na⁺ and proline binding (Quick *et al.*, 1996; Quick and Jung, 1997; Pirch *et al.*, 2002; Pirch *et al.*, 2003). Furthermore, this domain appears to line a hydrophilic cytoplasmic cavity and is supposed to be involved in Na⁺- and proline-induced conformational alterations (Pirch *et al.*, 2003). In this study, the functional role and structural dynamics of TM IX of PutP were investigated using genetic and biochemical approaches. Individual site-directed replacement of all residues in TM IX and analysis of transport function of the respective PutP derivatives revealed that this transmembrane segment is of particular functional importance. Six residues of TM IX (Gly328, Ser332, Ser340, Thr341, Gln345, and Leu346) were identified as being relevant for Na⁺-coupled proline uptake. Thus, individual substitution of these amino acids results in significant alterations of the kinetic parameters of transport (Tables 2.1. and 3.1.). Most severe defects were observed upon replacement of the conserved residues Ser340 and Thr341. Removal of the polar hydroxyl group at position 340 or 341 drastically reduces the apparent affinity of PutP for Na⁺ and Li⁺ (Fig. 2.3. and Table 2.1.). However, the size of the side chain at position 340 has a stronger impact on transport kinetics than the presence of the hydroxyl group, whereas the hydroxyl oxygen at position 341 is found to be essential for high affinity ion binding (Fig. 2.3. and Table 2.1.). These findings suggest that Ser340 is

located in the immediate vicinity of the Na⁺-binding site or may participate in Na⁺ binding via its main-chain carbonyl group. In contrast, Thr341 is probably directly involved in Na⁺ binding via its polar side chain. Besides their putative primary role in coordination of the coupling ion, both residues were also shown to be important for substrate binding. This is documented by the fact that substitutions at position 340 or 341 affect the apparent proline affinity, albeit to a lesser extent than the Na⁺ affinity (Fig. 2.3. and Table 2.1.). Similar observations have already been made upon substitution of Asp55, Ser57, or Gly58 of TM II, residues supposed to be involved in ligand binding (Quick *et al.*, 1996; Quick and Jung, 1997; Pirch *et al.*, 2002). Obviously, the binding of the coupling ion and the substrate occurs in close proximity. This idea is further supported by the kinetic properties of PutP-Q345C (decreased apparent affinities for Na⁺ and proline, comparatively small changes of V_{max}) that are, to some extent, reminiscent of those found for replacements of Ser340 or Thr341 (Tables 2.1. and 3.1.). Furthermore, individual Cys substitution of Gly328, Ser332, and Leu346 leads to reduced V_{max} values that can be explained by defects in the transport cycle subsequent to ligand binding. The defects are probably caused by inhibition of conformational alterations or a hampered release of the ligands. The latter explanation agrees, in principle, with the increased apparent ligand affinities observed for PutP-S332C and -L346C (Table 3.1.).

To obtain further information on the role of TM IX in ligand binding and ligand-induced conformational alterations, the accessibility of Cys residues placed throughout TM IX to sulfhydryl reagents of various sizes and polarities was determined. Labeling of Cys side chains at distinct positions in the periplasmic half (L331C and A337C) and the middle of TM IX (T341C) by the membrane-impermeable MTSET efficiently inhibits Na⁺-coupled proline uptake into intact cells (Fig. 3.4.b). Furthermore, Cys placed at positions 340 and 341 reacts with MTS compounds of different polarities as shown in right-side-out membrane vesicles (Fig. 2.4.b). These results indicate an accessibility of the four positions from the water phase and thus provide first evidence for the existence of an outwardly oriented hydrophilic funnel in PutP. This water-filled crevice may connect the proposed ion-binding residue Thr341 in the middle of the lipid bilayer with the periplasmic surface of PutP (Fig. 3.9.). Differing from the MTSET inhibition pattern, labeling with the membrane-permeable NEM inhibits the Na⁺-coupled proline uptake of five individual Cys replacements at positions in the cytoplasmic half (C344, L347C, V348C, and S351C) and in the middle (T341C) of TM IX (Fig. 3.4.a). The same residues were also found to be accessible to the sulfhydryl reagent FM, suggesting that the corresponding amino acid positions are exposed to an inwardly oriented hydrophilic cavity (Figs. 3.6. and 3.9.). The fact that the residues in the periplasmic half of TM IX are not accessible to the bulky FM (molecular volume of ~352 Å³) but to the relatively small MTSET (~167 Å³) implicates that the cytoplasmic cavity is significant larger than the periplasmic one. The only position in the periplasmic half of the domain where the introduced Cys side chain

reacts with FM is P324C (Fig. 3.6.). This is probably due to the presumed location of P324C at the periplasmic boarder of TM IX, which may protrude into the periplasm and is therefore highly accessible even for relatively large sulfhydryl reagents. Remarkably, PutP(Δ Cys)-T341C is the only derivate whose Cys residue is highly reactive to all three sulfhydryl reagents (MTSET, NEM, and FM) (Figs. 3.4. and 3.6.). This observation underlines the particular functional importance of this residue in ligand binding and is consistent with the idea of Thr341 being located at the interface between the periplasmic and cytoplasmic cavities. Such a position is in agreement with the alternating access mode of solute transport in which the ligand-binding site in the middle of the membrane is alternately exposed to both faces of the lipid bilayer (Jardetzky, 1966). In comparison to T341C, the Cys replacement of Ser340 only shows reactivity with NEM and MTS compounds in right-side-out membrane vesicles (Fig. 2.4.b), but it is neither inhibited by MTSET or NEM in intact cells nor reacts with FM in randomly oriented membrane vesicles (Figs. 3.4. and 3.6.). These results are consistent with a buried position of the side chain of S340C in the protein interior in the presence of a membrane potential ($\Delta\Psi$), whereas in the absence of $\Delta\Psi$, it is somewhat exposed to the periplasmic hydrophilic crevice. Similar observations have been made, in principle, for the H⁺/lactose transporter LacY of *E. coli* (Nie *et al.*, 2007). Here, the generation of an H⁺ electrochemical gradient ($\Delta\tilde{\mu}_{H^+}$) increases the reactivity of Cys residues on the periplasmic side of the sugar-binding site and in the putative hydrophilic pathway of LacY.

The effect of Na⁺ and/or proline on the accessibility of a given Cys side chain in TM IX was shown to be dependent on the location of the residue in either the periplasmic or cytoplasmic half of this segment. Na⁺ but not proline inhibits labeling of positions in the periplasmic half (L331C and A337C) and middle (T341C) of TM IX, although it cannot completely protect the respective Cys residue from modification possibly due to the dynamics of the system (Fig. 3.8.). The reduced accessibility of these positions may be attributed to Na⁺-induced conformational alterations of the local environment around the Cys residue or of the pathway from the bulk solution to the position. This idea correlates well with the proposed ordered binding model of Na⁺-coupled proline symport of PutP (Yamato and Anraku, 1993; Zhou *et al.*, 2004). In this scheme, Na⁺ binds to the empty transporter first, thereby inducing a conformational alteration, which increases the affinity of the transporter for the substrate. Another possible explanation is that the Na⁺-dependent Cys protection is due to direct steric hindrance by the coupling ion. However, while the side chain of Thr341 is supposed to be directly involved in ion binding, the side chains of Leu331 and Ala337 are nonpolar and cannot interact with monovalent cations, e.g. via electrostatic interactions or hydrogen bond formation. Hence, it is highly unlikely that protection of the latter positions is achieved by direct steric hindrance by Na⁺. Differing from the periplasmic half of TM IX, Cys placed at individual positions in the cytoplasmic half (C344, L347C, V348C, and S351C) of the

segment is effectively blocked against modification by proline but not by Na^+ (Fig. 3.7.). Since replacement of Cys344 by Ser (Yamato and Anraku, 1988) and Leu347, Val348, and Ser351 by Cys (Fig. 3.3.) does not significantly impair transport activity of PutP, the proline protective effect is probably not attributed to direct steric hindering by the substrate. Rather, the findings may be explained by proline-induced conformational alterations of TM IX or other helices that bury the residues within the protein. Interestingly, no significant influence of proline was observed on Cys accessibility to FM at position 341 (Fig. 3.7.). This result can be interpreted in two different ways: proline binding takes place at a site different from this part of TM IX or binding is prevented by the alteration of the substrate-binding site caused by the Cys substitution itself. The kinetic data of the replacement of Thr341 strongly support the latter idea (Table 2.1. and Fig. 2.3.).

Previously, a similar Cys accessibility pattern to the cytoplasmic half of TM IX was found for the functionally important TM II of PutP (Pirch *et al.*, 2003). Proline blocked modification of Cys at respective positions in TM II as detected here for TM IX. Therefore, intramolecular cross-linking between Cys introduced in TMs II and IX was performed in this study with cross-linkers of different lengths to test for a possible proximity between residues of these segments. Cys pairs positioned at the cytoplasmic and periplasmic end or middle of TM II and TM IX are efficiently cross-linked, even with the zero-length cross-linker CuPh (Fig. 2.5.b). Although the capacity of cross-linking varied with the position analyzed, the data indicate an at least temporal proximity of TMs II and IX. Furthermore, the observations that cross-linkers of varying lengths cross-link positions in both TMs and ligands do not influence the cross-linking efficiency suggest a high conformational flexibility of these parts of the transporter (Fig. 2.5.b and suppl. Fig. 2.4.). The strongest intermolecular cross-linking (>90%) with each of the cross-linking reagents was observed between Cys replacements of the proposed ligand-binding residues Ser57 (TM II) and Thr341 (TM IX). Together with the kinetic data and accessibility analyses of these and other residues in both domains, this supports the idea of a ligand-binding site located at the intersection of TMs II and IX (Fig. 2.6.). An equivalent site was observed and assigned as a Na^+ -binding site in the recently obtained X-ray structure of the Na^+ /galactose symporter vSGLT of the same transporter (SSS) family, as well as in the structure of the Na^+ /leucine symporter LeuT_{Aa} of the NSS family, the Na^+ /benzyl-hydantoin symporter Mhp1 of the nucleobase:cation symporter-1 (NCS1) family (TC 2.A.39), and the Na^+ /glycine betaine symporter BetP of the betaine/carnitine/choline transporter (BCCT) family (TC 2.A.15) (Faham *et al.*, 2008; Yamashita *et al.*, 2005; Weyand *et al.*, 2008; Ressler *et al.*, 2009). Surprisingly, these transporters share a common ten-helix core structure even though they do not have significant similarity at the sequence level. In all of these structures TM I and TM VIII of the core domain are tilted against each other forming close contacts in the middle of the

membrane. At the intersection of both domains, two hydroxyl groups from adjacent Ser and/or Thr residues of TM VIII are located, of which one (vSGLT) or two (LeuT_{Aa}, Mhp1, BetP) are involved in Na⁺ binding. Although this arrangement is resembled by Ser340 and Thr341 in TM IX of PutP (corresponds to TM VIII of the core domain), the kinetic data suggest a direct participation in Na⁺ binding only for the hydroxyl group of Thr341, while Ser340 may only interacts via its main-chain carbonyl group (Table 2.1.). In LeuT_{Aa}, the Na⁺-binding residue Ser355 is also crucial for binding of the substrate, which is consistent with the results obtained for the corresponding residue Thr341 in TM IX of PutP (Table 2.1.).

In summary, the results indicate that TM IX of PutP is of particular functional importance. Six residues (Gly328, Ser332, Ser340, Thr341, Gln345, and Leu346) were found to be relevant for ligand binding and/or translocation. The periplasmic half of TM IX is proposed to line a periplasmic funnel that may be used by Na⁺ as an entrance to the ion-binding site in about the middle of the membrane. In this region, Ser340 and Thr341 are located of which the side chain of Thr341 is presumably directly involved in Na⁺ binding. As shown by cross-linking experiments, these residues are at least temporally in close proximity to putative ligand-binding residues in TM II of PutP. This suggests that TMs II and IX may participate in the formation of a ligand-binding site located at the intersection of both domains. Furthermore, TM II and TM IX are found to line a large cytoplasmic cavity that allows the release of the ligands from the binding site into the intracellular bulk. Finally, ligand-induced conformational changes of TM IX and/or other helices lead to structural alteration of the periplasmic and cytoplasmic cavities presumably associated with an alternating access mode of ligand transport.

9.2. Backbone structure of transmembrane domain IX of PutP

To learn more about the structure and function of TM IX of PutP, an approach was developed for *ab initio* determination of backbone folds of TMs based on SDSL and DEER distance distribution measurements. Since many TMs of transport proteins are not ideal straight helices but often bent or kinked (Screpanti and Hunte, 2007), the backbone structure of TM IX was described in terms of a helix-loop-helix model. Starting from an ideal helix backbone, the dihedral angles φ_i and ψ_i in the loop region were fitted to the raw experimental DEER time traces for various loop lengths and positions, thereby also taking into account spin label rotamer distributions. For validation of this approach, the shape of TM VIII in the structure of LeuT_{Aa} was reproduced from synthetic DEER data. By using 20 synthetic DEER time traces, an ensemble of 11 structure models was obtained with a backbone RMSD of 0.68 Å, whereas the backbone RMSD with respect to the experimental structure was determined with approximately 2 Å (Fig. 4.3.). The larger RMSD with regard to the experimental structure indicates a systematic deviation of the structural model. This

systematic deviation could be traced back mainly to imperfections in SCWRL3 side group packing and corresponding deviations in the spin label rotamer distribution. Nevertheless, the results demonstrate that the established approach is suitable for the determination of backbone structures of TMs with a resolution that allows discussion of structure-function relations.

For transmembrane helix IX of PutP the use of 16 experimental DEER data sets results in an ensemble of 26 structures with a backbone RMSD of 1.6 Å (Fig. 4.6.c). The representative structure from this ensemble exhibits a pronounced kink near the functional important residue Thr341 (Figs. 4.6.d and 9.1.a). In this region the α -helical structure is interrupted by a loop segment with residues Val338 to Ser340 adopting extended, non-helical conformations. Similar interruptions in transmembrane helices have been encountered in the X-ray structures of other secondary transporters, like LeuT_{Aa}, vSGLT, Mhp1, BetP, NhaA, and CIC, where they are thought to be involved in ligand binding (Yamashita *et al.*, 2005; Faham *et al.*, 2008; Weyand *et al.*, 2008; Ressler *et al.*, 2009; Hunte *et al.*, 2005; Dutzler *et al.*, 2002). The fact that in most of the 26 structural models both putative ligand-binding residues Ser340 and Thr341 are found in the loop segment supports the idea that this also applies to the interruption in TM IX of PutP. Possibly, this structural feature is favorable for the coordination of ligands, because it exposes main-chain hydrogen-bonding partners and partially charged helix dipoles that are well suited for ligand attraction and/or ligand binding within the lipid bilayer (Screpanti and Hunte, 2007). Together with polar side chains in close vicinity, these electrostatic elements may confer selectivity for the ligand based on its size, polarity, and charge. In this context it is of interest that Cys substitution of Ala337 in TM IX slightly decreases the apparent Na⁺ affinity of PutP (threefold) (Table 3.1.). Since Ala337 is located at the C-terminal end of the first helical segment of TM IX, the partially negatively charged helix dipole at this position and/or the main-chain carbonyl oxygen of this residue may participate in Na⁺ binding. This conclusion is strongly supported by the X-ray structures of LeuT_{Aa}, Mhp1, and vSGLT in which the corresponding residues in the putatively homologous TMs are proposed to fulfill similar functions (Yamashita *et al.*, 2005; Weyand *et al.*, 2008; Faham *et al.*, 2008). However, it must be stated that the observed binding defect of PutP-A337C can also be explained by indirect effects, e.g. distortion of the loop structure followed by changes in the structural arrangement of the Na⁺-binding residues Ser340 and Thr341. The four remaining functional important residues identified in TM IX (Gly328, Ser332, Gln345, and Leu346) are found in the helical sections of the representative structure of this domain (Table 3.1; Fig. 9.1.a). Interestingly, these positions cluster on the phase of TM IX that is located opposite to Ala337, Ser340, and Thr341. Considering that none of the residues on this side of the helical segment is accessible from the aqueous bulk (Fig. 9.1.b), it is tempting to speculate that they may

interact with other helices in the protein and do not line the ligand-translocation pathway. Supported by the kinetic parameters ascertained for individual replacements at positions 328, 332, 345, and 346, these interactions might be of functional importance. Probably, they stabilize the structure of TM IX or convey conformational changes induced by ligand binding at the unwound region to other parts of the protein. The latter idea correlates well with the observed ligand-dependent changes in the accessibility pattern of TM IX, discussed below. Due to the kinetic analysis, it is also possible that residue Gln345 is directly involved in substrate binding (Table 3.1.). However, because of its inaccessibility from the aqueous bulk, it might be only important for transient binding of proline in the occluded state of the transporter.

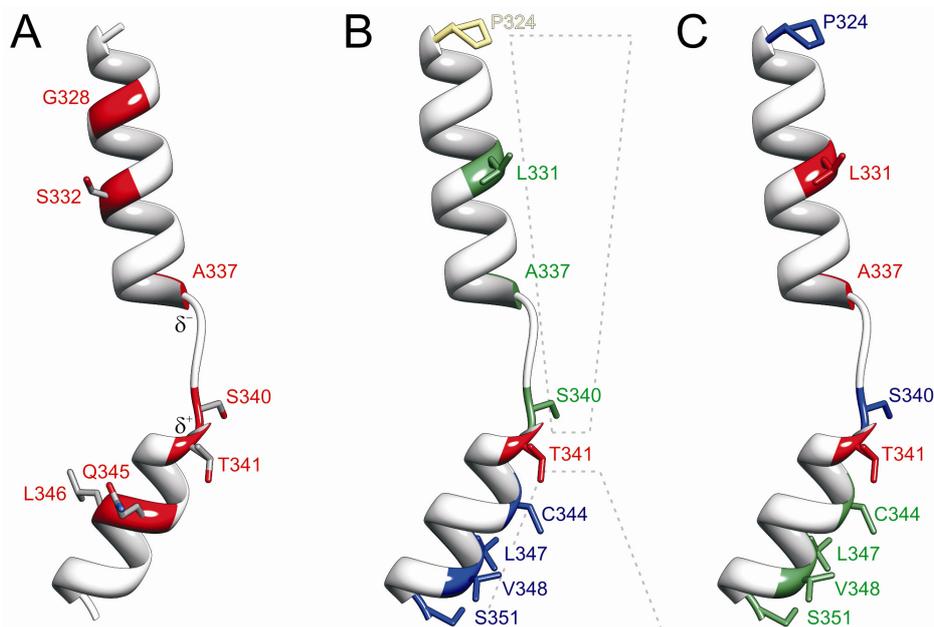


Fig. 9.1.: Localization of functionally and structurally important residues in the representative structure of TM IX of PutP. The TM is represented as a ribbon model and selected amino acids are shown as stick models. The single-letter amino acid code is used and the location of the residues in the amino acid sequence of the whole protein is indicated by a number. (A) Functional important amino acids are highlighted in red. The side chains are colored by atom type (grey: carbon, blue: nitrogen, red: oxygen). The labels δ^+ and δ^- indicate the partial charge at the helix ends. (B) Amino acid residues of which Cys replacements were found to be accessible to membrane-impermeable MTS-compounds are shown in green. Amino acid residues of which Cys replacements were found to be accessible to MTSET, NEM, and FM are colored in red. Amino acid residues of which Cys replacements were found to be accessible to NEM and FM are highlighted in blue. Amino acid residues of which Cys replacements were found to be accessible to FM are shown in yellow. Putative cytoplasmic and periplasmic aqueous cavities are marked by broken lines. (C) Amino acid residues of which Cys replacements are partially protected by Na^+ are colored in red. Amino acid residues of which Cys replacements are protected by proline are shown in green. Amino acid residues of which Cys replacements are not protected by Na^+ or proline are highlighted in blue. The figure was prepared with the program UCSF chimera (Pettersen *et al.*, 2004).

In contrast to the inaccessible residues, the accessible positions cluster on the phase of TM IX where the ligand-binding site appears to be located, suggesting that this side of the domain lines the translocation pathway (Fig. 9.1.b). Remarkably, the accessibility pattern of the different Cys-specific labels correlates well with the kinked structure of TM IX. Positions that are labeled by membrane-impermeable MTS-compounds from the outside of the lipid bilayer (Leu331, Ala337, and Ser340) are located above the kink, whereas positions that are accessible to the membrane-permeable NEM and FM from the inside of the membrane

(Cys344, Leu347, Val348, and Ser351) are found below the kink (Fig. 9.1.b). Furthermore, Thr341 of which the Cys replacement is the only one modified by all thiol-specific label reagents is situated in the immediate vicinity of the kink. Based on these observations it is likely that the periplasmic funnel and the cytoplasmic cavity may be due to the bend in TM IX (Fig. 9.1.b). Moreover, its central position at the interface between the crevices near the presumed ligand-binding site indicates that the kink might be associated with a hinge that facilitates movements required for ligand transport. This idea is supported by the observed ligand-dependent changes in the accessibility pattern of TM IX (Fig. 3.7. and Fig. 3.8.). As shown by the structure of TM IX, the sensitivity of the accessibility for either Na⁺ or proline depends on the position of the residue with respect to the kink (Fig. 9.1.c). Residues that are partially protected from labeling by Na⁺ binding to the transporter are located above (Leu331 and Ala337) or at the kink (Thr341). On the contrary, residues that show a proline-dependent protection of accessibility are positioned below the kink (Cys344, Leu347, Val348, and Ser351). Thus, it can be envisaged that the Na⁺-induced accessibility changes may reflect conformational alterations in the kink region of TM IX that lead to changes in the shape of the periplasmic funnel. According to the ordered binding mechanism proposed for PutP (Yamato and Anraku, 1993; Zhou *et al.*, 2004), these structural changes at the ligand-binding site might increase the affinity of the transporter to proline. In the case of the proline-protective effect, substrate binding probably induces hinge motions in TM IX followed by a closure of the cytoplasmic cavity that buries the residues inside the protein. Since no stimulatory proline effect is found for residues on the periplasmic side of the protein, this conformation may represent an occluded state of the transporter.

The X-ray structures of the structural related LeuT_{Aa}, vSGLT, and BetP provide evidence for different conformational states as they might proceed through the alternating access mechanism of Na⁺-coupled substrate transport (Yamashita *et al.*, 2005; Singh *et al.*, 2007; Zhou *et al.*, 2007; Faham *et al.*, 2008; Ressler *et al.*, 2009). Comparison of the structure of TM IX of PutP with the corresponding TMs in the crystal structures can thus provide new insights into conformational alterations of this domain. Interestingly, the structural model of TM IX of PutP and the putatively equivalent TM VIII in LeuT_{Aa} show similar kinked backbone structures, although the X-ray structure corresponds to a substrate-bound state (Singh *et al.*, 2007) (Fig. 4.6.e). If the two domains are aligned at Thr354/Ser355 of LeuT_{Aa} and Ser340/Thr341 of PutP, and are restricted to the same length, the model of TM IX superimposes on the structure of TM VIII with a backbone RMSD of 1.98 Å. On changing from the outward-facing conformation (LeuT_{Aa}) over the occluded state (BetP) to the inward-facing conformation of vSGLT, the transmembrane helix straightens and can be well described by a model of an ideal α -helix. Assuming that TM IX of PutP changes its conformation by a hinge motion that closes the cytoplasmic cavity and changes the shape of

the periplasmic funnel, it would also be expected to be approximately straight in the substrate-bound state. However, this notion needs to be proven by modeling the backbone structure of TM IX using DEER distance distribution measurements obtained in the presence of ligands.

Taken together, the newly developed approach based on DEER distance distribution measurements on doubly spin-labeled protein derivatives and modeling of backbone folds in terms of helix-loop-helix models is shown to provide an efficient mean for *ab initio* structure determination of transmembrane helices with a resolution that allows discussion of structure-function relations. For TM IX of PutP, a structural ensemble is obtained with an internal RMSD of less than 2 Å. The structures show a pronounced kink near the presumed ligand-binding site. This kink appears to be a hinge that closes the cytoplasmic hydrophilic cavity and changes the shape of the periplasmic funnel upon ligand binding.

9.3. Helix packing of PutP

In addition to information about the structure and function of single transmembrane helices, the architecture of the whole helix bundle of a polytopic membrane protein is of great interest for understanding its mechanism of function. However, although the X-ray structures of the presumably structurally related Na⁺-coupled solute transporters LeuT_{Aa}, Mhp1, BetP, and vSGLT could be solved during the past few years, attempts to crystallize PutP have failed so far (Yamashita *et al.*, 2005; Weyand *et al.*, 2008; Ressler *et al.*, 2009; Faham *et al.*, 2008). In this study, DEER distance measurements between spin labels attached near the helix ends were used to model the helix bundle of PutP. For this purpose, the TMs of the helix bundle were considered as cylinders whose orientation in space is defined by the Cartesian coordinates of their end points. The uncertainty associated with the size and the conformational distribution of the spin label was accounted for by describing each helix by a four-point model including the midpoints of the N-O bonds of the spin labels and the helix ends (Fig. 5.1.). For PutP that comprises 13 TMs a complete determination of such a model would require a minimum of 325 interspin distances. To reduce this considerable effort, internal restraints of the coarse-grained helix bundle and template restraints derived from the putative structurally homologous ten-helix core of vSGLT were utilized. Finally, the distance matrix geometry approach of Crippen and Havel (1988) was used for generation of helix bundle models that conform to all internal, experimental, and template restraints. The performance of this newly developed procedure was tested by modeling the LeuT_{Aa} helix bundle based on 46 synthetic EPR distance restraints and a template derived from the vSGLT core structure. The central structures of the two resulting ensembles of 200 structural models have RMSD Q_c values of 7.17 Å and 8.65 Å with respect to the X-ray structure, while the original template derived from the vSGLT structure has a Q_o of 11.65 Å. This

demonstrates that the use of EPR distance restraints significantly improves the model. In particular, the approach correctly establishes the relative arrangement of the two non-core TMs XI and XII of LeuT_{Aa}, which is poorly predicted by the vSGLT-based template (Fig. 5.2.). The refinement of the ten-helix core, however, is only marginal as shown by the RMSDs of 6.08 Å and 6.38 Å obtained for only the core helices of the two central structures in comparison to 6.49 Å of the original vSGLT template (Fig. 5.3.). Clearly, the number of the utilized synthetical EPR distance restraints is not high enough to achieve a resolution that allows the discussion of differences in the core between vSGLT and LeuT_{Aa}. Nevertheless, testing of the approach with core decoys generated by random permutations of the helix assignment of the vSGLT-derived template demonstrates that the approach can distinguish with high certainty between structurally homologous and structurally unrelated core templates. Hence, the fold of the core can be reliably recognized.

For PutP, 45 DEER distance restraints were used to generate two ensembles of 300 structures. The standard deviation of the 20% of structures closest to the central structures, which is shown by LeuT_{Aa} to be a good approximation of the RMSD with respect to the X-ray structure, is determined with 11.52 Å and 11.69 Å for the two PutP ensembles. Similar to the LeuT_{Aa} models, the RMSD of the ten-helix core of the central structures of PutP from the vSGLT core template is about 6.5 Å. This suggests that the arrangement of the core helices in PutP does not deviate more strongly from vSGLT than LeuT_{Aa}. Furthermore, as tested by decoys, it is very unlikely that the set of restraints is consistent with a core architecture that is significantly different from vSGLT. Together, these findings provide evidence that the core structure of PutP is structurally homologous to vSGLT. Like in vSGLT, the core of the PutP helix bundles contains two inverted repeat domains (TMs II-VI and TMs VII-XI), not detectable in the amino acid sequence, that are related by a pseudo-two-fold axis in the plane of the membrane (Faham *et al.*, 2008) (Fig. 5.6.). These structural repeats are intertwined, giving a central four-helix bundle consisting of TMs II, III, VII, and VIII that is surrounded by a scaffold defined by TMs IV, V, VI, IX, X, and XI. The most relevant differences with respect to the template based on the vSGLT structure concern the three non-core helices, whose positions are determined exclusively by internal and experimental restraints (Fig. 5.5.). TM I, which is situated in the back of the bundle of vSGLT, is localized in the front in both experimental PutP structures. For TM XIII the position of the periplasmic end is found to be similar in both proteins. However, the tilt direction of this domain differs and thus the cytoplasmic end near the C terminus of PutP is localized elsewhere. TM XII is similarly situated relative to the core, but it is more separated from the bundle in the PutP structures than in the vSGLT template. As the PutP models and the X-ray structure of vSGLT have the same common fold of the ten-helix core but show different orientations of the auxiliary transmembrane helices, the core domain may define the fundamental machinery of

these transporters. Furthermore, the homologous architecture of the core domains implies commonalities in mechanism, including principles in ligand binding. Indeed, the ligand-binding sites in the X-ray structures of the structural homologous vSGLT, LeuT_{Aa}, Mhp1, and BetP are similarly located at the center of the proteins, even though they belong to different transporter gene families and transport different substrates (Faham *et al.*, 2008; Yamashita *et al.*, 2005; Weyand *et al.*, 2008; Ressler *et al.*, 2009). Considering that the complete set of helices that participate in the formation of the ion- and/or substrate-binding site in PutP is not known, comparison of the helix arrangement in the PutP models with the X-ray structures of vSGLT, LeuT_{Aa}, Mhp1, and BetP can thus draw conclusion about the TMs that are involved in ligand binding.

In the two helix bundle models of PutP, the functional important TMs II and IX, which are proposed to participate in the formation of the ligand-binding site, are located in the core domain (Fig. 5.6.e, f). TM II forms part of the central four-helix bundle, whereas TM IX is situated in the surrounding scaffold. The position of both segments relative to each other is found to be somewhat different in the two models. In the first one TM II faces TM IX (Fig. 5.6.e) while in the second structure the two TMs are separated by TM VI (Fig. 5.6.f). Although both structures are consistent with the available restraints, effective cross-linking between residues in TM II and TM IX, even by the zero length cross-linker CuPh, favors an arrangement similar to that observed in the first model (Figs. 2.5.b and 5.6.e). As mentioned above, such an arrangement is also found in the X-ray structures of vSGLT, LeuT_{Aa}, Mhp1, and BetP (Faham *et al.*, 2008; Yamashita *et al.*, 2005; Weyand *et al.*, 2008; Ressler *et al.*, 2009). In the latter structures the corresponding TM I and TM VIII of the ten-helix core are in close vicinity to each other, forming a Na⁺-binding site at the intersection of these domains in approximately the middle of the membrane. There appears to be a conserved Na⁺-binding pattern made by main-chain carbonyl oxygens of residues in TM I (vSGLT: Ala62, Ile65; LeuT_{Aa}: Gly20, Val23; Mhp1: Ala38, Ile41; BetP: Ala147) and by hydroxyl groups from Ser and/or Thr residues in TM VIII of the core (vSGLT: Ser365; LeuT_{Aa}: Thr354, Ser355; Mhp1: Ser312, Thr313; BetP: Thr467, Ser468). In case of vSGLT, LeuT_{Aa}, and Mhp1, also the main-chain carbonyl oxygen of an Ala residue in TM VIII is proposed to be involved in Na⁺ binding (vSGLT: Ala361; LeuT_{Aa}: Ala351; Mhp1: Ala309). This coordination pattern is in part resembled by the situation suggested for PutP. In particular, the putative Na⁺-binding residues Ser340 and Thr341 in TM IX of PutP are at positions equivalent to the ion-binding polar residues in vSGLT, LeuT_{Aa}, Mhp1, and BetP. Although Ser340 is proposed to bind Na⁺ by the main-chain carbonyl oxygen, these residues may fulfill similar functions in PutP (Chapter 2). Additionally, the carbonyl oxygen of Ala337 might also participate in the coordination of Na⁺ as it is suggested by the results of the kinetic analyses, accessibility studies, and the backbone structure of TM IX of PutP (Table 3.1.; Figs. 3.4., 3.8., 9.1.). Thus,

it is likely that the Na⁺-binding motif AXXS/TS/T, found in the corresponding TMs of vSGLT, LeuT_{Aa}, and Mhp1, also exists in TM IX of PutP. In TM II of PutP the situation is less clear. Here, the carboxyl group of Asp55 is identified to be essential for Na⁺ dependence of the transporter. However, this residue is not conserved within the SSS family as Ser340 and Thr341 and there is no analog residue in the available X-ray structures of vSGLT, LeuT_{Aa}, Mhp1, or BetP. Probably, Asp55 stabilizes the empty and/or occupied Na⁺-binding site in PutP as it is most likely deprotonated at physiological pH by additional hydrogen bonds. Due to this deprotonation it is also possible that Asp55 permits charge compensation of the partially positively charged helix dipole near residue Ser340 and Thr341 of the adjacent discontinuous helix IX of PutP (Fig. 9.1.a). This may shape the ion-binding site and/or reduces the electrostatic repulsion between the dipole end and the bound Na⁺. Furthermore, because TM II is supposed to be involved in both ion and substrate binding, the properties of the substrate may also have influenced the precise mechanism of Na⁺ binding. Besides the carboxyl group of Asp55, individual replacements of Ala48 and Gly63 in TM II have an impact on Na⁺-dependent transport (Pirch *et al.*, 2003). Since TM I of the core region of vSGLT, LeuT_{Aa}, Mhp1, and BetP contributes main-chain carbonyl groups to the cation-binding site, it is tempting to speculate that these residues might also play a role in coordination of the coupling ion. Taken together, the helix bundle structures in conjunction with structure-function analyses and comparison with the available crystal structures of this structural family strongly support the idea of a Na⁺-binding site that is located at the interface between TMs II and IX of PutP. However, additional distance measurements are required for more reliable determination of the relative position of both helices in the bundle structure of PutP.

As mentioned above, TM II and TM IX of PutP are not only proposed to be important for Na⁺ binding, but also for coordination of proline. In fact, Ser57 and Gly58 of TM II and Ser340, Thr341, and Gln345 of TM IX are suggested to participate in the formation of a substrate-binding site that is situated in close vicinity or even overlaps with the ion-binding site (Quick *et al.*, 1996; Pirch *et al.*, 2002; Chapters 2 and 3). This putative proline-binding pocket at the interface between TMs II and IX is additionally enclosed by five TMs (IV, VI, and XI of the scaffold and III and VII of the four-helix bundle) that represent further appropriate candidates for substrate coordination (Fig. 5.6.b, e and c, f). Two of these helices, TMs IV and VII, have been recently identified to be crucial for transport function (Raba, M., Pirch, T., and Jung, H., unpublished data). Thereby, the conserved residues Trp244 and Tyr248 in TM VII and Tyr140 in TM IV were found to be important for Na⁺-coupled substrate transport (Fig. 5.4.). The available data support the notion of a proline-binding site formed by TMs II, IV, VII, and IX (TMs I, III, VI, and VIII of the core). Interestingly, this arrangement resembles only in part the situation found in vSGLT, BetP, and Mhp1, but is

fully consistent with the binding region in LeuT_{Aa}. In the sugar transporter vSGLT, helices I, II, VI, VII, and X of the core domain are involved in substrate binding (Faham *et al.*, 2008). These TMs coordinate the OH-moieties of the galactose by hydrogen bonds from polar side chains. For BetP, the substrate betaine is placed between TMs II and VI of the core domain and is bound by one Tyr and three Trp residues via cation- π and van der Waals interactions (Ressl *et al.*, 2009). In Mhp1, TMs I, III, VI, and VIII are involved in the coordination of the benzyl-hydantoin, similar to the proposed set of substrate-binding transmembrane helices in PutP (Weyand *et al.*, 2008). In Mhp1, however, the benzyl-hydantoin is coordinated exclusively by side chains (π -stacking interactions with Trp residues and hydrogen bonds with conserved Glu and Asn residues), whereas in PutP also main-chain atoms contribute to proline binding. Within the four crystallized Na⁺-coupled transporters of this structural family, the involvement of main-chain atoms in substrate binding is exclusively found in LeuT_{Aa} (Yamashita *et al.*, 2005). Here, the α -substitutions of leucine are predominantly bound by exposed carbonyl oxygens and amid nitrogens and of the partially charged helix dipoles of the unwound regions in TMs I and VI. Only two amino acid side chains participate in the coordination of the α -substitutions of leucine, Tyr108 in TM III and Ser256 in TM VI. As residues Tyr140 in TM IV and Tyr248 in TM VII of PutP are at equivalent position to the corresponding residues Tyr108 and Ser256 of LeuT_{Aa}, they may fulfill similar function in coordinating the carboxyl group of the substrate. Furthermore, the functional important residue Gly58 in TM II of PutP probably corresponds to Gly26 in TM I of LeuT_{Aa} that forms a main-chain interaction between the amide nitrogen and the α -carboxy group of leucine (Pirch *et al.*, 2002). In addition to these residues, the side chain of Ser57 in TM II of PutP is proposed to be crucial for high-affinity proline uptake (Quick *et al.*, 1996). However, in the corresponding TM I of LeuT_{Aa} no equivalent residue exists that is involved in substrate binding. This is most likely due to the different substrate specificity of PutP and LeuT_{Aa} that is ensured by variation of specific interaction sites. In contrast to this polar part of the substrate-binding pocket in LeuT_{Aa} that accommodates the α -substitutions of leucine, the aliphatic side chain of the substrate is located within a hydrophobic binding region formed by residues of TMs III, VI, and VIII (Yamashita *et al.*, 2005). As aforementioned, one of these residues, Ser355 of TM VIII, is engaged in substrate and Na⁺ binding as it is also proposed for the equivalent residue Thr341 in TM IX of PutP (Chapter 2). Thus, it is possible that Thr341 is also involved in the formation of a hydrophobic binding pocket, which might maintain the cyclic aliphatic side chain of proline. In addition to Thr341, Trp244 in TM VII and Gln345 in TM IX of PutP may also constitute part of this hydrophobic pocket. Both residues are conserved throughout the proline transporters of the SSS family and are crucial for Na⁺-dependent proline uptake (suppl. Fig. 5.1., Tables 2.1. and 3.1., Raba, M., Pirch, T., and Jung, H., unpublished data).

Altogether, the helix bundle models in combination with biochemical and mutagenesis studies on PutP indicate a proline-binding site, which is created by TMs II, IV, VII, and IX. Comparison of this arrangement with those in the X-ray structures of vSGLT, LeuT_{Aa}, Mhp1, and BetP shows that it is most similar to the situation found in LeuT_{Aa}. Since LeuT_{Aa} catalyzes the uptake of amino acids, like PutP, this suggests that the specific interactions within the common core structure of this structural family vary in order to ensure substrate specificity for a diverse range of substrates. However, it must be stated that the complete set of residues engaged in substrate binding is not known for PutP. Furthermore, no functional data is available for TMs III, VI, and XI that also represent potential helices for substrate binding. Clearly, more functional data for PutP is required to verify the presumed proline-binding site.

In conclusion, it is shown that the established approach for modeling of the helix bundle of membrane proteins based on DEER distance measurements, structural homology, and distance geometry is suitable to determine the fold of proteins in the native environment of a lipid bilayer. For PutP, two ensembles of structural models are obtained with a RMSD of 11.52 Å and 11.69 Å. The low resolution of these ensembles indicates that the number of utilized EPR distance restraints is not sufficient to discuss differences in the core structure between vSGLT and PutP. Nevertheless, the fold of the core domain of PutP is reliably recognized and the arrangement of the three non-core helices with respect to the core is established. Finally, comparison of the core structure of PutP with those of vSGLT, LeuT_{Aa}, Mhp1, and BetP leads to a model of putative ligand-binding sites. This model can guide functional studies to identify the complete set of residues essential for transporter function.

9.4. Structural dynamics of the NhaA dimer

Previously, genetic complementation studies, biochemical pull-down experiments, cross-linking analyses (Gerchman *et al.*, 2001), and cryoelectron microscopy of 2D crystals (Williams *et al.*, 1999; Williams, 2000) revealed that the Na⁺/H⁺ antiporter NhaA of *E. coli* functions as a homodimer in the native membrane. The dimeric state is probably important with respect to the pH regulation of NhaA, as suggested by intermolecular cross-linking between two NhaA monomers at the dimer interface that changed the pH profile of the antiporter (Gerchman *et al.*, 2001). However, the explicit role of dimerization for ion transport and/or pH-regulation of NhaA remains elusive. In this study, the pH dependence of the structure of NhaA and its dimer was investigated by means of CW and pulsed EPR techniques using singly spin-labeled derivatives H225R1 and V254R1, as they appear to be involved in the pH response of the antiporter (Rimon *et al.*, 1995; Gerchman *et al.*, 1999). The CW EPR spectra of the singly spin-labeled NhaA variant H225R1 reveal that the mobility of the nitroxide at position 225 slightly decreases as the pH is shifted from 5.8 to 7.0 and 8.0

(Fig. 6.2.a). Since the mobility of the spin label side chain depends on its topographical location, this indicates that a pH-shift towards the alkaline range induces small but significant conformational alterations in the region of position 225. The existence of these pH-induced structural changes is confirmed by low temperature EPR spectroscopy. Increasing the pH from 5.8 to 8.0 causes a movement of the nitroxide in a slightly more polar environment (Fig. 6.3.). Considering the drastic changes in activity in the studied pH range between 5.8 and 8.0 for the wild-type protein (Taglicht *et al.*, 1991) as well as for the derivative H225C (Rimon *et al.*, 1995), the pH dependence of the spectral changes is surprisingly weak. Thus, it is tempting to speculate that this residue may not be part of the pH sensor but is involved in the transduction of the signal from the site of the sensor to the site of the ion-translocation pathway. This is supported by the X-ray structure of pH-downregulated NhaA (Hunte *et al.*, 2005). In the structure, His225 is located at the periplasmic end of TM VIII in close vicinity to TM V that is proposed to participate in Na⁺ binding (Fig. 1.6.a, b). Possibly, the pH-induced conformational changes observed and predicted for the preceding loop 9 (Venturi *et al.*, 2000; Gerchman *et al.*, 1999; Olkhova *et al.*, 2007), which constitutes part of the pH sensor, have small impact on the conformation of TM VIII. Hence, the orientation of the side chain of His225 might alter slightly, allowing the formation of hydrogen bonds with the polar residue Tyr175 and/or with main-chain atoms of Ala172 at the periplasmic end of the ion-binding TM V. This, in turn, may stabilize the conformation of TM V in the active state of NhaA, promoting Na⁺ binding to the transporter. In line with this assumption, replacement of His225 by either Cys or Ser but not Ala yielded an antiporter with a wild-type pH-sensitive phenotype, suggesting that polarity and/or the ability to form hydrogen bonds are essential at this position for pH regulation of NhaA (Rimon *et al.*, 1995). As indicated by the CROX and oxygen accessibility measurements, the localization of the spin label at position 225 relative to the membrane does not significantly change during a pH-shift from 5.8 to 8.0 (Table 6.1.). The low CROX and intermediate oxygen accessibility of H225R1 implies that residue His225 faces the apolar environment of the lipid bilayer. Although this is in agreement with the crystal structure of NhaA, the low accessibility to water-soluble CROX is in apparent contradiction to the at least moderate accessibility of this residue to NEM, as this alkylation reagent requires ionization of the thiol group (Olami *et al.*, 1997). This contradiction might be explained by the different timescales of the two experiments. In a dynamic environment, labeling studies sense exposure of residues, even if it is realized only during small fractions of the conformational trajectory, as reaction times are in the range of tens of minutes. EPR experiments, on the other hand, measure the time average of accessibility. Alternatively, the residue might be transiently accessible to nearby water molecules that penetrate the membrane and help to ionize the Cys residue.

Similar to position 225, the nitroxide accessibility detected for V254R1 implies an exposure of the spin label side chain to the apolar environment of the membrane (Table 6.1.). Furthermore, the mobility analysis suggests a more buried location for V254R1 than for H225R1 (Fig. 6.1.). The latter result agrees well with the proposed position of V254 in the dimer interface, as indicated by cross-linking analysis (Gerchman *et al.*, 2001). Moreover, neither the mobility and polarity analyses nor the accessibility studies revealed significant pH-dependent structural alterations close to this residue (Fig. 6.2.b and Table 6.1.). This seems to be inconsistent with the proposed pH-induced conformational change of the periplasmic end of TM IX at alkaline pH (Tzuberly *et al.*, 2004; Gerchman *et al.*, 1999). However, residues for which the pH-dependent structural alterations were identified by changes in accessibility to trypsin (Lys249) and MANS (Glu252) are situated on the phase of TM IX that is located opposite to Val254 (Hunte *et al.*, 2005) (Fig. 1.6.b). Thus, conformational alterations of TM IX that affect the accessibility of Lys249 and Glu252 do not necessarily have an impact on residue Val254. Another explanation could be that the changes in accessibility of Lys249 and Glu252 are not due to structural alterations in TM IX but in TM II, which is adjacent to this side of TM IX. Finally, it has to be taken into account that the mobility and accessibility of the nitroxide and the polarity in its microenvironment are not necessarily affected by a conformational alteration. Therefore, DEER distance measurements were performed using the singly spin-labeled NhaA derivatives V254R1 and H225R1. For V254R1, a single peak at 2.00 nm with a width of 0.2 nm is found (Fig. 6.5.c), indicating that NhaA exists as a homodimer in the membrane with well-defined structure. In principle, the data could also be explained by a symmetric trimer. However, as the cross-linking analyses and the cryo-EM structure of NhaA imply dimerization, this seems to be unlikely (Gerchman *et al.*, 2001; Williams *et al.*, 1999; Williams, 2000). In line with the CW EPR measurements, no pH-dependent alteration of the interspin distance V254R1/V254R1' could be observed, suggesting that the dimer interface of NhaA does not undergo any large-scale changes in the range between pH 5.8 and 8.0. Most interestingly, the modulation depth Δ of the DEER time traces, which corresponds to the average number $\langle n \rangle$ of labels per oligomer, slightly increases with increasing pH. Similar results are obtained for H225R1. Here, the DEER measurements reveal an interspin distance of 4.36 nm with a width of ~ 0.3 nm (Fig. 6.5.a). The distance does not significantly change, but Δ varies during a pH shift from 5.8 to 7.0 and 8.0. Due to the significant better DEER data of NhaA-H225R1, Δ was quantitatively analyzed and the reversibility of the pH effect was tested for this derivative. The quantification of Δ provides evidence that the average number $\langle n \rangle$ of labels per protein oligomer slightly but significantly increases upon shifting pH from 5.8 to 8.0 (Fig. 6.6.d). On the other hand, lowering pH from 8.0 to 5.8 decreases $\langle n \rangle$, indicating that the pH effect is completely reversible (Fig. 6.6.d). This observation is most easily explained by a dimerization

equilibrium that depends moderately on pH. Assuming that labeling of the protein molecules is complete ($f = 1$), the degree of dimerization α of derivative H225R1 increases from approximately $\alpha = 0.35$ at pH 5.8 to $\alpha = 0.6$ at pH 8.0. With $f = 0.6$, which is estimated to be the lower limit of the degree of labeling, α is about 0.6 at pH 5.8 and 1.0 at pH 8.0. Independently of the value of f , $\alpha > 0$ implies that dimerization is substantial at pH 5.8 where the NhaA derivative H225C is almost inactive (Olami *et al.*, 1997). It may thus be safe to conclude that dimerization is not a sufficient condition for antiporter activity of NhaA. Dimerization might be important for pH regulation of NhaA. This is supported by intermolecular cross-linking of the V254C derivative at the dimer interface that alters the pH profile of the antiporter and by functional complementation of pairs of conditional lethal mutants, impaired in pH response (Gerchman *et al.*, 2001). Based on the DEER distance measurements between the V254R1 derivatives and cross-linking analyses, the putative pH sensor appears to be located at or near the dimer interface. It can thus be envisaged that dimerization serves to optimize the pH sensor. The pH dependence of the dimerization would then be an indirect effect of the presumed pH-induced conformational change of the periplasmic end of TM IX and/or the loop 9 that influences the intermolecular contact site but not be directly related to control of the ion flow by pH. Furthermore, two residues near position 254, His253 and His256, have predicted pK_a values of 5.2 and 7.5 in a membrane environment (Olkhova *et al.*, 2006). Thus, protonation/deprotonation events at these positions may also have an impact on the dimer equilibrium.

In summary, the DEER distance measurements between singly spin-labeled NhaA derivatives H225R1 and V254R1 indicate that NhaA exists as a dimer in the membrane, as previously suggested by a low-resolution electron density map and cross-linking experiments. The modulation depth is shown to decrease reversibly when pH is shifted from pH 8.0 to 5.8. A quantitative analysis suggests a dimerization equilibrium, which depends moderately on pH. Furthermore, only slight pH-dependent changes are observed for the mobility of the nitroxide and polarity of the microenvironment of the spin label attached to position 225, while no other changes are detected by CW EPR. This indicates that His225 and Val254 may not contribute to the pH sensor or the ion-binding site but are probably involved in the transmittance of the pH signal from the sensor to the translocation site.

9.5. High-resolution structure of the NhaA dimer

To better understand the role of the dimeric state of NhaA for its function and/or pH regulation, knowledge about the high-resolution structure of the dimer is indispensable. Since the X-ray structure of NhaA displays a nonphysiological association of the NhaA monomers in 3D crystals, a novel approach was developed to determine the structure of the presumably

physiological dimer in the native membrane based on pulsed EPR spectroscopy. As the X-ray structure of the NhaA monomer was shown to be in good agreement with the low-resolution structure observed in the dimers in 2D crystals (Screpanti *et al.*, 2006), the structure of the dimer was obtained only by determining the relative arrangement of the two molecules. For this purpose, DEER distance measurements between nine singly spin-labeled NhaA derivatives (E82R1, S87R1, A118R1, N177R1, A202R1, K221R1, H225R1, E241R1, and V254R1) at pH 5.8 and restraint-driven rigid body transformations were used to determine the four geometric parameters that fully describe the relative arrangement of the monomers in the dimer. Thereby, information on the distance distribution was utilized to match a modeled conformational distribution of the spin labels. The resulting set of 144 structures shows a backbone RMSD of 0.6 Å with respect to the best-fit structure. This implies that the quality of the final structure is limited by the resolution of the crystal structure of the monomer rather than by the SDSL EPR docking fit. The best-fit structure shows two points of contacts between the two monomers in the dimer (Fig. 7.4.). The main contact is built by the two-stranded antiparallel β -sheets in loop 2 at the periplasmic side of the membrane (Fig. 7.4.b). This interaction site is stabilized by four hydrogen bonds between residues Gln47 and Arg49 and backbone groups of the respective other β -sheet and a number of hydrophobic contacts. The second contact is formed between TMs VII and IX at the cytoplasmic site of the membrane (Fig. 7.4.c). Most likely, this contact involves a hydrogen bond between Arg204 and Val254 as well as hydrophobic contacts of Arg204 and Leu210 with Trp258. In general, the dimer interface is not densely packed, suggesting a relatively weak binding between the moieties. This is in accordance to the pH dependence of dimer formation and the dissociation of the physiological dimer during 3D crystallization (Screpanti *et al.*, 2006). The interface between the monomers is constituted by one of the two bundles of six helices, the interfacial domain (TMs I, II, VI-IX) (Fig. 7.5.). This interfacial domain contains the pH sensor of NhaA at the periplasmic end of TM IX (Hunte *et al.*, 2005). The other bundle contains the putative ion-translocation pathway and thus forms the translocation domain (TMs III-V, X-XII). Although the two domains are apparently separated, there are interdomain contacts formed by helix I of the interfacial domain with helix XII near the periplasmic side and with helix IVc near the cytoplasmic side of the transporter (Hunte *et al.*, 2005).

Remarkably, recent mutagenesis, cross-linking, and cryo-EM studies support the structural arrangement of the monomers in the EPR structure. Deletion of the two-stranded antiparallel β -sheets is shown to result in a NhaA derivative that exists exclusively as a monomer in the native membrane, suggesting that the β -sheets are essential for dimerization (Rimon *et al.*, 2007). In addition, substitution of each individual residue in the β -sheets by Cys and testing for intermolecular cross-linking and gel mobility on BN-Page reveal that

Gln47 and Arg49 perform the strongest cross-linking and are crucial for stabilization of the dimer (Herz *et al.*, 2009). Furthermore, cross-linking between residues near the cytoplasmic contact side in TM IX implies that Ser246, Pro247, Ala248, Val254, and Leu255 form the interface of the NhaA dimer at this region (Tzuberly *et al.*, 2008). The estimated distances between these residues are in good agreement with the distances predicted by the EPR structure. In particular, for V254 a distance of 8.1 Å is found between the S atoms of the corresponding cysteins in the EPR structure. This is in line with the finding that the rigid cross-linking agent *p*-PDM, which can cross-link only residues that are at least 10 Å apart, causes a change in the pH profile of NhaA activation (Gerchman *et al.*, 2001). Moreover, investigation of the role of these residues for dimerization of NhaA by BN page shows that Cys replacement of the individual residues does not impair the NhaA dimer, suggesting that the β -sheets at the periplasmic side of the lipid bilayer represent the main contact that holds the NhaA dimer (Herz *et al.*, 2009).

Comparison of the EPR structure of NhaA with the recently published cryo-EM based dimer model shows that the dimer geometry is very similar (Appel *et al.*, 2009). Superimposition of both structures yields a RMSD of 1.88 Å (Fig. 9.2.). The main difference between the structures is related to the tips of the β -hairpins that are displaced in the cryo-EM model by ~11 Å out of the membrane plane (Fig. 9.2.b). This difference in the conformation of loop 2 could be explained by a distortion induced by crystal contacts that stabilize the nonphysiological dimer in the 3D crystals. However, as shown by the structural alignment, the distortion is probably limited to the β -sheets.

With respect to the relevance of the dimer for NhaA function it is remarkable that the monomeric form of NhaA, deleted in the β -sheets, exhibits antiporter activity and pH regulation similar to the wild-type, indicating that the functional unit of NhaA is the monomer (Rimon *et al.*, 2007). However, under extreme stress conditions (0.1 M LiCl or 0.7 M NaCl at pH 8.5), the dimer is found to be more beneficial than the monomer in conferring growth resistance. This is thought to be due to the higher stability of the dimer in comparison to the monomer under these conditions (Herz *et al.*, 2009).

In addition to H225R1 and V254R1, also N177R1 in the interfacial domain and A118R1 in the translocation domain of NhaA were tested for pH dependence of the distance distribution. Similar to the results obtained for H225R1 and V254R1, the DEER data of N177R1 and A118R1 do not reveal significant changes with increasing pH, except for variations in the modulation depth. This supports the contention that the structure of the interfacial domain and the dimer interface is conserved during a pH shift from 5.8 to 8.0, although propensity for dimer formation varies slightly. For the DEER measurements of A118R1, this failure to detect any change might be due to the long-distance of approximately 82 Å (Table 7.1.). Because of this large distance, the DEER data may not be sensitive to a

displacement of this residue by up to 10 Å. A conformational change in this region of loop 4 would thus be consistent with the measurements. Indeed, as predicted by MD simulations of NhaA at pH 4.0 and pH 8.0, loop 4 undergoes structural rearrangements that may facilitate a movement of helix IVp (Oikhova *et al.*, 2007). This pH-induced displacement of helix IVp is in agreement with the pH-dependent conformational changes observed in cryo-EM projection maps of 2D crystals incubated at different pH values (Appel *et al.*, 2009). Furthermore, helix IVc and TMs IX, X, and XI of the translocation domain might undergo structural alterations as indicated by cross-linking analysis (Kozachkov *et al.*, 2007), epitope mapping of monoclonal antibodies (Rimon *et al.*, 2008), and MD simulations (Oikhova *et al.*, 2007; Arkin *et al.*, 2007). Based on these findings, it is tempting to speculate that the translocation domain moves with respect to the interfacial domain on changing pH. Such a change in the relative arrangement of the two domains might be the most likely explanation for transmittance of the signal from the pH sensor in the interfacial domain to the translocation pathway.

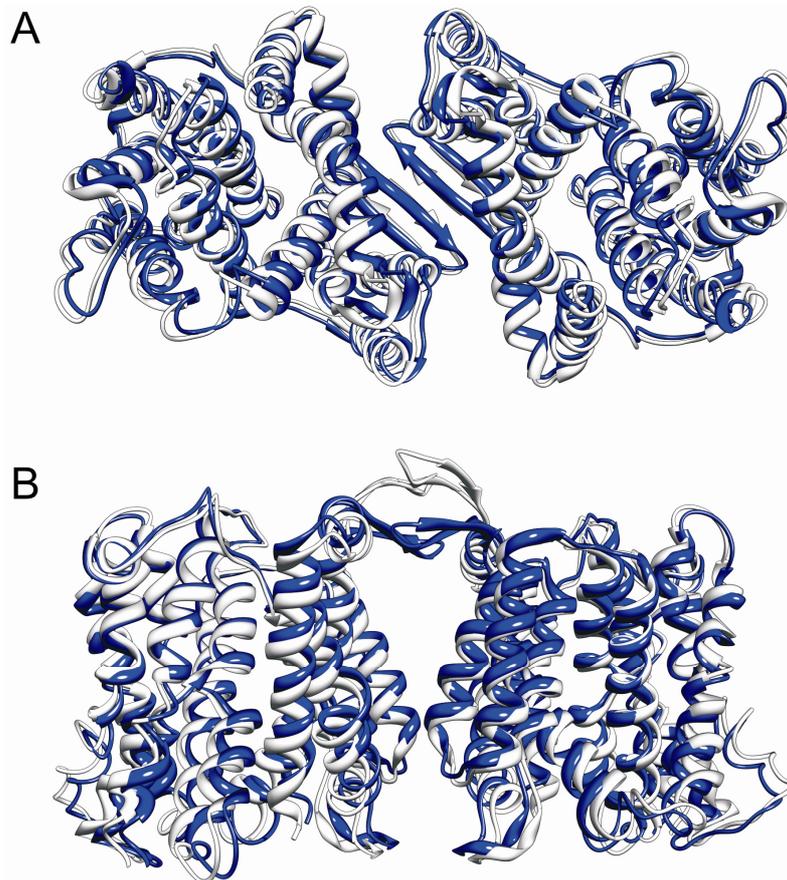


Fig. 9.2.: Superimposition of the EPR structure of the NhaA dimer and the cryo-EM based NhaA dimer model (Appel *et al.*, 2009). The EPR structure (blue) and the cryo-EM based structure (white) are represented as ribbon models. (A) Top view of the structural alignment from the cytoplasm. (B) View of the superimposed structures in plane of the membrane.

Signal transmittance between the domains is also required to explain functional complementation between the conditional lethal mutants H225R and G338S (Gerchman *et al.*, 2001). The replacement H225R does not support growth at alkaline pH, because its activity is shut off above pH 7.5 (Gerchman *et al.*, 1993). Likewise, the substitution G338S is

lethal, but this derivative has lost pH control so that the uncontrolled activity at high pH and high Na⁺ concentrations probably depletes the cells of energy and/or causes over-acidification of the cytoplasm (Rimon *et al.*, 1998). If both mutants are coexpressed, the activity and/or pH regulation of NhaA and accordingly the growth of the cells at alkaline pH in the presence of high Na⁺ concentrations are restored (Gerchman *et al.*, 2001). Such functional complementation between remote sites is only feasible if ion translocation is a process that involves a cooperative, but asymmetric motion of both moieties of the dimer.

This raises the questions of how local structural changes in the two domains of a monomer are coupled and how structural changes are transmitted between the monomers in the dimer. Presumably, the many positively charged residues (Arg203, Arg204, His243, Arg245, Lys249, Arg250, His253, and His256) at or near the dimer interface are important for these processes (Fig. 7.6.). Arg204, His253, and His256 have predicted pK_a values that allow protonation/deprotonation events in the range between pH 5.0-7.5 (Oikhova *et al.*, 2006). Thus, it can be assumed that pH directly influences contacts between the monomers, which is in line with the observed pH dependence of the monomer-dimer equilibrium of NhaA. Among these residues, Arg204 may play a central role. This residue probably forms a hydrogen bond across the interface with the main-chain carbonyl oxygen of Val254 in the other monomer. Val254 in turn is in hydrophobic contact with Arg250 in its own moiety. These interactions are of high importance since they may ensure contact between large clusters of strongly interacting residues within a monomer across the interface. These clusters were identified *in silico* by multiconformation continuum electrostatics (MCCE) analysis of the X-ray structure of NhaA and involve Glu78, Glu82, His243, Lys249, Arg250, Glu252, His253, and His256 of the pH sensor and Asp133, Asp163, Asp164, and Lys300 of the presumed ion-binding site (Oikhova *et al.*, 2006). Thus, the clusters provide evidence for a direct electrostatic coupling between the translocation pathway in the translocation domain and the pH sensor in the interfacial domain.

Besides these electrostatic interactions, three further contact sites exist between the translocation domain and the interfacial domain of a monomer. As mentioned above, His225 in TM VIII of the interfacial domain of NhaA may mediate contact with residues Ala172 and Tyr175 in TM V, which harbors the ion-binding residues Asp163 and Asp164 and belongs to the translocation domain. The importance of this interaction for pH-regulation of NhaA activity is demonstrated by the conditional lethal derivative H225R in which the contact is probably modified or destroyed (Gerchman *et al.*, 2001). The second contact between both domains of the NhaA monomer is provided by residues Ala135-Leu143 in helix IVc, which lines the cytoplasmic part of the ion-translocation pathway, and residues Ala12-Ile19 in helix I. At the third contact site, some residues in the range between Asn64 and Phe72 in TM II interact with residues between Met341 and Phe350 in helix XIp, which is in the periplasmic part of

the ion-translocation channel. Interestingly, displacement of helices IVc and XIp in the antiporter structure at alkaline pH is suggested to lead to an exposure of the ion-binding site to the cytoplasmic cavity (Hunte *et al.*, 2005). This conformation is proposed to allow Na^+/H^+ exchange.

Based on the EPR structure of the NhaA dimer, it is now imaginable that during activation of NhaA, a movement of helix IVc results in a shift of the cytoplasmic end of TM I toward the interface. On the periplasmic site of the protein, the proposed movement of helix XIp is presumably related to a displacement of the N-terminal end of helix II toward the contact site of the two monomers. Activation would thus correspond to a clockwise rotation of the interface, or keeping the interfacial domains fixed, to an anticlockwise rotation of the two translocation domains with respect to the interface, when seen from the periplasmic side. The latter idea agrees well with the displacement of the translocation domain seen in the EPR structure compared to the cryo-EM structure published by Williams (2000) (Fig. 7.5.). However, in contrast to this observation, superimposition of the more recent cryo-EM based dimer model of NhaA (Appel *et al.*, 2009) and the EPR structure does not reveal any significant changes in both domains of the monomer (Fig. 9.2.). Furthermore, it has to be said that the EPR structure as well as the cryo-EM structures were obtained at $\text{pH} \leq 5.8$ where the wild-type protein is still inactive. Clearly, a high-resolution structure at pH above 6.5 is required to proof or refute the hypothesis of cooperative structural rearrangements in the two moieties of the dimers that appear to be the basis for the ion translocation and/or pH regulation of NhaA.

In conclusion, the new established method based on DEER distance measurements and modeling of the conformational distribution of the spin label side chain is shown to provide structures of dimers or complexes of biomacromolecules with essentially the same resolution that can be obtained for the structures of the constituents. It is also applicable even if the interactions between the molecules are rather weak, since it does not require crystallization. Thus, the structures of membrane proteins can be determined in the native environment of a lipid bilayer. For NhaA, a high-resolution structure of the putative physiological dimer was obtained with a resolution of 0.6 Å. It reveals two points of contact between the two moieties that were confirmed by recent cross-linking and cryo-EM studies. Finally, the structure suggests cooperative conformational motions in the two monomers of the dimer, important for ion translocation and/or pH-regulation of NhaA.

9.6. DeerAnalysis2006 - A comprehensive software package for analyzing of pulsed ELDOR data

Extraction of distance distributions from dipolar evolution functions measured by pulsed ELDOR techniques such as the four-pulsed DEER experiment requires numerically

stable mathematical algorithms that can handle data for different classes of samples. Furthermore, experimental imperfections such as the limited excitation bandwidth of microwave pulses need to be considered (Maryasov and Tsvetkov, 2000; Milov *et al.*, 2004; Banham *et al.*, 2008). The software package DeerAnalysis2006 features a comprehensive set of tools for such data analysis within a common user interface. In comparison to DeerAnalysis2004 (Jeschke *et al.*, 2002; Jeschke *et al.*, 2004a; Jeschke *et al.*, 2004b), a number of new approaches are incorporated for the derivation and use of experimental background functions as well as for Tikhonov regularization with L curve computation, model-based computation of distance distributions, excitation bandwidth correction, spin counting, and concentration measurements.

The determination and use of experimental background functions was tested on two singly spin-labeled PutP derivatives and their mixture reconstituted into liposomes. A fit of the experimental DEER data of S41R1 by background functions assuming homogeneous distribution confined to two ($D = 2$) and three dimensions ($D = 3$) shows that the RMSD with $D = 2$ is significantly better than with $D = 3$ (Fig. 8.8.a, b). This finding fully agrees with the 2D confinement assumed for a membrane protein embedded in the bilayer of a liposome. The RMSD again improves when the logarithm of $V(t)$ is fitted by a fifth-order polynomial, suggesting that S41R1 is not located at the surface but in the interior of the protein (Fig. 8.8.c). Similar observations are made for S423R1, although the improvement by the polynomial compared to $D = 2$ is not drastic in this case, which probably means that this site is closer to the surface of PutP. Taken together, the results indicate that polynomial fits are well suited for the derivation of experimental background functions from singly labeled samples. This is important, as the direct use of the noisy experimental data sets of the singly labeled samples would introduce significant statistical errors. Furthermore, in contrast to the experimental data sets, utilization of smooth fit functions allows to account for differences in protein concentration.

The quality of the experimentally determined background models was tested on the mixture of both singly labeled proteins. By fitting the individual background polynomials obtained for S41R1 and S423R1 to $V(t)$ of the mixture, the RMSD significantly improves as compared to a fit with analytical background models. The best fit is obtained by using the weighted sum of both experimentally derived background functions, suggesting that the sum is a better approximation for $B(t)$ of the mixture than each individual background function. As the $B(t)$ of the mixture is well reproduced by the experimental background function, one can expect that the approach is well suited for the background correction of the corresponding doubly labeled mutant. Indeed, the only fit parameter c depends only slightly on the selected range for the background fit, i.e., the procedure of fitting $B(t)$ at later times and extrapolating

it back to early times works well for the mixture of single mutants. The same procedure can also be applied for complexes of two different macromolecules, but not for dimers or higher homooligomers.

The efficiency of model-free and model-based computation of distance distribution $P(r)$ from $F(t)$ is demonstrated by determining the shape of the narrow asymmetric distance distribution of a shape-persistent biradical (Godt *et al.*, 2000). For model-free analysis the Tikhonov regularization with L curve computation of the regularization parameter α was used, which is the main approach for model-free conversion of $F(t)$ to $P(r)$ in DeerAnalysis2006. In contrast to the self-consistency criterion (Weese, 1992) implemented in the past version, DeerAnalysis2004, the L curve criterion is more transparent and less susceptible to failure due to noisy data (Chiang *et al.*, 2005). For model-based analysis, the continuous form of the worm-like chain model from polymer physics was utilized, which depends on only two parameters, the contour length L of the polymer and the persistence length l_p that quantifies flexibility (Wilhelm and Frey, 1996). Comparison of the distance distribution $P(r)$ obtained by Tikhonov regularization and model-based data analysis suggests that the Tikhonov regularization slightly oversmooths the distribution (Fig. 8.6.d). This is also indicated by the fit residuals that reveal more oscillatory deviations at late times for the model-free fit of the experimental form factor than for the model-based fit (Fig. 8.6.a, c). Nevertheless, the distributions provided by Tikhonov regularization and the wormlike chain model show similar asymmetry. This demonstrates that monomodal distance distributions with narrow peaks can be reliably determined from experimental form factors by Tikhonov regularization with L curve computation, even if no information about the expected resolution of the distance distribution is available. Although, fitting of the form factor $F(t)$ by a model distribution generally improves the reliability of the analysis (Jeschke and Polyhach, 2007), some deviations are found in the model-based fit at early times, implying that the distribution does not exactly have the shape suggested by the model (Fig. 8.6.c). This can be explained by the neglected conformational distribution of the nitroxide endlabers that introduces additional broadening of the distance distribution. Thus, in an improved model this conformational distribution should be taken into account to reduce such deviations. DeerAnalysis2006 allows for implementing of user-defined models with up to six parameters. The quality of such a model can be checked by comparing its RMSD to the one of Tikhonov regularization with optimum regularization parameter. Furthermore, it is advisable to visually inspect whether the shapes of $P(r)$ obtained by both approaches agree.

The precision of the data analysis procedures was further tested on a broad distance distribution in surface-labeled gold nanoparticles (Chechik *et al.*, 2004). For these nanoparticles a good model for the distance distribution $P(r)$ exists that has two parameters,

the mean diameter $\langle d_s \rangle$ of the sphere and the standard deviation σ_d of this diameter (Sect. 8.3.3.). The distance distribution extends from $r = 0$ to the effective sphere diameter. Therefore, the data were also utilized for testing whether excitation bandwidth correction can recover suppressed contributions to the distance distribution at short distances. The distance profiles obtained by Tikhonov regularization with and without bandwidth correction and by model-based analysis with bandwidth correction are very similar at distances larger than 2 nm (Fig. 8.7.c). However, at distances shorter than about 2 nm, significant deviations are observed. Compared with the model-based fit, a neglect of the excitation bandwidth correction in Tikhonov regularization leads to an underestimate of the contribution of short distances. Surprisingly, Tikhonov regularization with excitation bandwidth correction appears to overcompensate this suppression with respect to the model-based fit (Fig. 8.7.c). Since the exact shape of the distance distribution cannot be recovered by adjusting the excitation bandwidth parameter σ_{exc} (Eq. 8.5.), the overcompensation is tentatively assigned to contributions from distances below 1.3 nm. In this range, the assumptions of a constant oscillation frequency and a constant time shift fail. Nevertheless, the results demonstrate that excitation bandwidth correction does improve agreement with the model based distribution. As shown by test measurements on a series of spin-labeled peptides the excitation bandwidth correction is applicable down to distances of about 1.6 nm (Banham *et al.*, 2008).

The model-based fit is shown to be much more robust with respect to limited excitation bandwidth than Tikhonov regularization. For the spin-labeled gold nanoparticles, $F(t)$ is dominated by contributions from distances larger than 2 nm, so that the two-parameter fit is hardly affected by these effects. The values for $\langle d_s \rangle$ and σ_d obtained with and without excitation bandwidth correction are quite similar and the RMSD of the fit of the experimental form factor does not change. Since these values agree well with previous transmission electron microscopy analysis of the diameter of the gold core and CW EPR measurements of the distance between the label and its point of attachment to the surface (Ionita *et al.* 2005), this implies that the model is well suited for reliable analyzing the distance distribution in the gold nanoparticles.

Taken together, the model-free analysis by Tikhonov regularization with L curve computation provides an efficient means for reliable determination of the shape of $P(r)$ for narrow distance distributions. However, as shown by the broad distance distribution, Tikhonov regularization gives reasonable results only for contributions above 2 nm, while distances below 2 nm are strongly suppressed if the excitation bandwidth limitation is not considered. The analysis of distances close to the lower boundary of applicability of the DEER experiment can be significantly improved by excitation bandwidth correction. Model-based computation of $P(r)$ stabilizes the solution against the suppression of short distances due to limited excitation bandwidth. The quality of the model for $P(r)$ can be inspected by

comparing the RMSD of the model-based fit to the RMSD for Tikhonov regularization with an optimum regularization parameter.

The number $\langle n \rangle$ of spins per nanoobject can be determined from the modulation depth Δ of the background corrected form factor of a given sample (Milov *et al.*, 1984; Chapter 6). Quantification of $\langle n \rangle$ requires knowledge of the modulation depth factor λ , which is influenced by the position of the shape of the pump pulse and the position of the pump frequency with respect to the center of the mode (Jeschke *et al.*, 2006). Thus, reliable computation of $\langle n \rangle$ requires determination of λ by an experimental calibration with known samples. As shown in Chapter 6, a calibration using five biradicals and one triradical, λ could be determined with a correlation coefficient of 0.9965. The default value of λ in DeerAnalysis2006 is based on these data sets. However, as this parameter strongly depends on the measurement conditions, it has to be recalibrated with at least one biradical sample for each spectrometer and probe head.

In addition to the determination of the number of spins per nanoobject, also measurements of local concentrations have to be calibrated. For a homogeneous 3D distribution of nanoobjects, the density parameter is given by $k = xc$, where c is the local concentration of spins. Because the factor x depends on the modulation depth factor λ , x is best determined by calibration. In the DeerAnalysis2006 package, an example data set is provided that was acquired with a 2.5 mM TEMPOL solution in toluene. However, for any new combination of spectrometer and resonator a recalibration is necessary with a sample of known concentration.

In summary, DeerAnalysis2006 provides several modifications and extensions that enhance its reliability and functionality with respect to the previous version DeerAnalysis2004. Its graphical user interface allows for determination of distance distributions, spin numbers, and spin concentrations from experimental dipolar evolution functions. Furthermore, the reliability and precision of the extracted information can be tested. Finally, user-supplied models for spin-to-spin distance distribution within a certain class of nanoobjects can be added to an existing library and be fitted with a universal algorithm.

9.7. References

- Appel, M., Hizlan, D., Vinothkumar, K. R., Ziegler, C., and Kühlbrandt, W. (2009) Conformations of NhaA, the Na⁺/H⁺ exchanger from *Escherichia coli*, in the pH-activated and ion-translocating states. *J. Mol. Biol.* 388: 659-672.
- Arkin, I. T., Xu, H., Jensen, M. Ø., Arbely, E., Bennett, E. R., Bowers, K. J., Chow, E., Dror, R. O., Eastwood, M. P., Flitman-Tene, R., Gregersen, B. A., Klepeis, J.

- L., Kolossváry, I., Shan, Y., and Shaw, D. E.** (2007) Mechanism of Na⁺/H⁺ antiporting. *Science* 317: 799-803.
- Banham, J. E., Baker, C. M., Ceola, S., Day, I. J., Grant, G. H., Groenen, E. J., Rodgers, C. T., Jeschke, G., and Timmel, C. R.** (2008) Distance measurements in the borderline region of applicability of CW EPR and DEER: a model study on a homologous series of spin-labelled peptides. *J. Magn. Reson.* 191: 202-218.
- Chechik, V., Wellsted, H. J., Korte, A., Gilbert, B. C., Caldararu, H., Ionita, P., and Caragheorghopol A.** (2004) Spin-labelled Au nanoparticles. *Faraday Discuss.* 125: 279-291.
- Chiang, W.-Y., Borbat, P. P., and Freed, J. H.** (2005) The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J. Magn. Reson.* 172: 279-295.
- Crippen, G. M., and Havel, T. F.** (1988) *Distance geometry and molecular conformation*, Research Studies Press Ltd., Taunton.
- Dahl, S. G., Kristiansen, K., and Sylte, I.** (2002) Bioinformatics: from genome to drug targets. *Ann. Med.* 34: 306-312.
- Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T., and MacKinnon R.** (2002) X-ray structure of a CIC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature* 415: 287-294.
- Faham, S., Watanabe, A., Besserer, G. M., Cascio, D., Specht, A., Hirayama, B. A., Wright, E. M., and Abramson, J.** (2008) The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na⁺/sugar symport. *Science* 321: 810-814.
- Gerchman, Y., Olami, Y., Rimon, A., Taglicht, D., Schuldiner, S., and Padan, E.** (1993) Histidine-226 is part of the pH sensor of NhaA, a Na⁺/H⁺ antiporter in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 90: 1212-1216.
- Gerchman, Y., Rimon, A., and Padan, E.** (1999) A pH-dependent conformational change of NhaA Na⁺/H⁺ antiporter of *Escherichia coli* involves loop VIII-IX, plays a role in the pH response of the protein, and is maintained by the pure protein in dodecyl maltoside. *J. Biol. Chem.* 274: 24617-24624.
- Gerchman, Y., Rimon, A., Venturi, M., and Padan, E.** (2001) Oligomerization of NhaA, the Na⁺/H⁺ antiporter of *Escherichia coli* in the membrane and its functional and structural consequences. *Biochemistry* 40: 3403-3412.
- Godt, A., Franzen, C., Veit, S., Enkelmann, V., Pannier, M., and Jeschke, G.** (2000) EPR probes with well defined, long distances between two or three unpaired electrons. *J. Org. Chem.* 65: 7575-7582.
- Herz, K., Rimon, A., Jeschke, G., and Padan, E.** (2009) β-sheet-dependent dimerization is essential for the stability of NhaA Na⁺/H⁺ antiporter. *J. Biol. Chem.* 284: 6337-6347.

- Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E., and Michel, H.** (2005) Structure of a Na⁺/H⁺ antiporter and insights into mechanism of action and regulation by pH. *Nature* 435: 1197-1202.
- Ionita, P., Caragheorgheopol, A., Gilbert, B. C., and Chechik, V.** (2005) Dipole-dipole interactions in spin-labeled au nanoparticles as a measure of interspin distances. *J. Phys. Chem. B* 109: 3734-3742.
- Jardetzky, O.** (1966) Simple allosteric model for membrane pumps. *Nature* 211: 969-970.
- Jeschke, G., Koch, A., Jonas, U., and Godt, A.** (2002) Direct conversion of EPR dipolar time evolution data to distance distributions. *J. Magn. Reson.* 155: 72-82.
- Jeschke, G., Panek, G., Godt, A., Bender, A., and Paulsen, H.** (2004a) Data analysis procedures for pulse ELDOR measurements of broad distance distributions. *Appl. Magn. Reson.* 26: 223-244.
- Jeschke, G., Bender, A., Paulsen, H., Zimmermann, H., and Godt, A.** (2004b) Sensitivity enhancement in pulse EPR distance measurements. *J. Magn. Reson.* 169: 1-12.
- Jeschke, G., Zimmermann, H., and Godt, A.** (2006) Isotope selection in distance measurements between nitroxides. *J. Magn. Reson.* 180: 137-146.
- Jeschke, G., and Polyhach, Y.** (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* 9: 1895-1910.
- Kozachkov, L., Herz, K., and Padan, E.** (2007) Functional and structural interactions of the transmembrane domain X of NhaA, Na⁺/H⁺ antiporter of *Escherichia coli*, at physiological pH. *Biochemistry* 46: 2419-2430.
- Maryasov, A. G., and Tsvetkov, Y. D.** (2000) Formation of the pulsed electron-electron double resonance signal in the case of a finite amplitude of microwave fields. *Appl. Magn. Reson.* 18: 583-605.
- Milov, A. D., Ponomarev, A. B., and Tsvetkov, Y. D.** (1984) Electron-electron double resonance in electron spin echo: model biradical systems and the sensitized photolysis of decalin. *Chem. Phys. Lett.* 110: 67-72.
- Milov, A. D., Naumov, B. D., and Tsvetkov, Y. D.** (2004) The effect of microwave pulse duration on the distance distribution function between spin labels obtained by PELDOR data analysis. *Appl. Magn. Reson.* 26: 587-599.
- Nie, Y., Ermolova, N., and Kaback, H. R.** (2007) Site-directed alkylation of LacY: effect of the proton electrochemical gradient. *J. Mol. Biol.* 374: 356-364.
- Olami, Y., Rimon, A., Gerchman, Y., Rothman, A., and Padan, E.** (1997) Histidine 225, a residue of the NhaA-Na⁺/H⁺ antiporter of *Escherichia coli* is exposed and faces the cell exterior. *J. Biol. Chem.* 272: 1761-1768.
- Olkhova, E., Hunte, C., Screpanti, E., Padan, E., and Michel, H.** (2006) Multiconformation continuum electrostatics analysis of the NhaA Na⁺/H⁺ antiporter of *Escherichia coli* with functional implications. *Proc. Natl. Acad. Sci. U. S. A.* 103: 2629-2634.

- Olkhova, E., Padan, E., and Michel, H.** (2007) The influence of protonation states on the dynamics of the NhaA antiporter from *Escherichia coli*. *Biophys. J.* 92: 3784-3791.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E.** (2004) UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25: 1605-1612.
- Pirch, T., Quick, M., Nietschke, M., Langkamp, M., and Jung, H.** (2002) Sites important for Na⁺ and substrate binding in the Na⁺/proline transporter of *Escherichia coli*, a member of the Na⁺/solute symporter family. *J. Biol. Chem.* 277: 8790-8796.
- Pirch, T., Landmeier, S., and Jung, H.** (2003) Transmembrane domain II of the Na⁺/proline transporter PutP of *Escherichia coli* forms part of a conformationally flexible, cytoplasmic exposed aqueous cavity within the membrane. *J. Biol. Chem.* 278: 42942-42949.
- Quick, M., Tebbe, S., and Jung, H.** (1996) Ser57 in the Na⁺/proline permease of *Escherichia coli* is critical for high-affinity proline uptake. *Eur. J. Biochem.* 239: 732-736.
- Quick, M., and Jung, H.** (1997) Aspartate 55 in the Na⁺/proline permease of *Escherichia coli* is essential for Na⁺-coupled proline uptake. *Biochemistry* 36: 4631-4636.
- Ren, Q., and Paulsen, I. T.** (2005) Comparative analysis of fundamental differences in membrane transport capabilities in prokaryotes and eukaryotes. *PLoS Comp. Biol.* 1: 190-201.
- Ressl, S., Terwisscha van Scheltinga, A. C., Vorrhein, C., Ott, V., and Ziegler, C.** (2009) Molecular basis of transport and regulation in the Na⁺/betaine symporter BetP. *Nature* 458: 47-52.
- Rimon, A., Gerchman, Y., Olami, Y., Schuldiner, S., and Padan, E.** (1995) Replacements of histidine 226 of NhaA-Na⁺/H⁺ antiporter of *Escherichia coli*. Cysteine (H226C) or serine (H226S) retain both normal activity and pH sensitivity, aspartate (H226D) shifts the pH profile toward basic pH, and alanine (H226A) inactivates the carrier at all pH values. *J. Biol. Chem.* 270: 26813-26817.
- Rimon, A., Gerchman, Y., Kariv, Z., and Padan, E.** (1998) A point mutation (G338S) and its suppressor mutations affect both the pH response of the NhaA-Na⁺/H⁺ antiporter as well as the growth phenotype of *Escherichia coli*. *J. Biol. Chem.* 273: 26470-26476.
- Rimon, A., Tzuber, T., and Padan, E.** (2007) Monomers of the NhaA Na⁺/H⁺ antiporter of *Escherichia coli* are fully functional yet dimers are beneficial under extreme stress conditions at alkaline pH in the presence of Na⁺ or Li⁺. *J. Biol. Chem.* 282: 26810-26821.
- Rimon, A., Hunte, C., Michel, H., and Padan, E.** (2008) Epitope mapping of conformational monoclonal antibodies specific to NhaA Na⁺/H⁺ antiporter: structural and functional implications. *J. Mol. Biol.* 379: 471-481.

- Screpanti, E., Padan, E., Rimon, A., Michel, H., and Hunte, C.** (2006) Crucial steps in the structure determination of the Na⁺/H⁺ antiporter NhaA in its native conformation. *J. Mol. Biol.* 362: 192-202.
- Screpanti, E., and Hunte, C.** (2007) Discontinuous membrane helices in transport proteins and their correlation with function. *J. Struct. Biol.* 159: 261-267.
- Singh, S. K., Yamashita, A., and Gouaux, E.** (2007) Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. *Nature* 448: 952-956.
- Taglicht, D., Padan, E., and Schuldiner, S.** (1991) Overproduction and purification of a functional Na⁺/H⁺ antiporter coded by *nhaA* (*ant*) from *Escherichia coli*. *J. Biol. Chem.* 266: 11289-11294.
- Tzuber, T., Rimon, A., and Padan, E.** (2004) Mutation E252C increases drastically the K_m value for Na⁺ and causes an alkaline shift of the pH dependence of NhaA Na⁺/H⁺ antiporter of *Escherichia coli*. *J. Biol. Chem.* 279: 3265-3272.
- Tzuber, T., Rimon, A., and Padan, E.** (2008) Structure-based functional study reveals multiple roles of transmembrane segment IX and loop VIII-IX in NhaA Na⁺/H⁺ antiporter of *Escherichia coli* at physiological pH. *J. Biol. Chem.* 283: 15975-15987.
- Venturi, M., Rimon, A., Gerchman, Y., Hunte, C., Padan, E., and Michel, H.** (2000) The monoclonal antibody 1F6 identifies a pH-dependent conformational change in the hydrophilic NH₂ terminus of NhaA Na⁺/H⁺ antiporter of *Escherichia coli*. *J. Biol. Chem.* 275: 4734-4742.
- Weese, J.** (1992) A reliable and fast method for the solution of Fredholm integral equations of the first kind based on Tikhonov regularization. *Comput. Phys. Commun.* 69: 99-111.
- Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E. P., Rutherford, N. G., Hadden, J. M., O'Reilly, J., Ma, P., Saidijam, M., Patching, S. G., Hope, R. J., Norbertczak, H. T., Roach, P. C., Iwata, S., Henderson, P. J., and Cameron, A. D.** (2008) Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* 322: 709-713.
- Wilhelm, J., and Frey, E.** (1996) Radial distribution function of semiflexible polymers. *Phys. Rev. Lett.* 77: 2581-2584.
- Williams, K. A., Geldmacher-Kaufer, U., Padan, E., Schuldiner, S., and Kühlbrandt, W.** (1999) Projection structure of NhaA, a secondary transporter from *Escherichia coli*, at 4.0 Å resolution. *EMBO J.* 18: 3558-3563.
- Williams, K. A.** (2000) Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature* 403: 112-115.
- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E.** (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* 437: 215-223.

- Yamato, I., and Anraku, Y.** (1988) Site-specific alteration of cysteine 281, cysteine 344, and cysteine 349 in the proline carrier of *Escherichia coli*. *J. Biol. Chem.* 263: 16055-16057.
- Yamato, I., and Anraku, Y.** (1993) Na⁺/substrate symport in prokaryotes. In *Alkali cation transport systems in prokaryotes* (Bakker, E. P., ed.) pp. 53-76, CRC Press, Inc., Boca Raton, FL.
- Zhou, A., Wozniak, A., Meyer-Lipp, K., Nietschke, M., Jung, H., and Fendler, K.** (2004) Charge translocation during cosubstrate binding in the Na⁺/proline transporter of *E. coli*. *J. Mol. Biol.* 343: 931-942.
- Zhou, Z., Zhen, J., Karpowich, N. K., Goetz, R. M., Law, C. J., Reith, M. E., and Wang, D. N.** (2007) LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. *Science* 317: 1390-1393.

ACKNOWLEDGMENTS

First of all I would like to thank my supervisor Prof. Dr. Heinrich Jung for giving me the opportunity to work on these challenging and exciting projects in a highly interdisciplinary manner. I am deeply grateful for his support, encouragement, guidance, and for the trust he has put in me.

I am extremely grateful to Prof. Dr. Gunnar Jeschke for introducing me to the fascinating world of pulse-EPR. I would also like to thank him for his interest and motivating enthusiasm in the projects and for his crucial contributions, without which this work would not have been possible.

Special thanks go to my thesis committee, especially to Prof. Dr. Dirk Schüler for being the second examiner of this thesis.

Many thanks to Prof. Dr. Etana Padan for providing all the *nhaA* single-Cys mutants, for her support, and helpful discussions.

I would like to thank Prof. Dr. Kirsten Jung for her constant support and her interest in my work.

I want to express my gratitude to Prof. Dr. Heinz-Jürgen Steinhoff for providing access to his instrumental facilities in Osnabrück, for important suggestions and helpful discussions.

Dr. Yevhen Polyhach is gratefully acknowledged for his help with setting up the Elexsys spectrometer in Konstanz and for his essential contributions to this work. Thanks for the highly enjoyable trip to the Rhine Falls and for interesting scientific and non-scientific discussions.

Sincere thanks to Prof. Dr. Wolfgang Spiess and Dr. Dariush Hinderberger for providing measurement time on my favorite Elexsys spectrometer in Mainz.

Many thanks to Christian Bauer for technical support at the MPI in Mainz and highly enjoyable conversations.

I am very grateful to Dr. Enrica Bordignon, Dr. Johann Klare, Dr. Maxim Yulikov, Klaus-Peter Vogel, and Dr. Henrik Brutlach for great help with the spectrometers in Osnabrück or Konstanz.

I owe a tremendous debt of gratitude to Christiane Langkamp for generously sharing her small apartment for numerous weeks in which I stayed in Mainz.

I like to thank Dr. Markus Langkamp for critical reading of this thesis.

Moreover, I want to thank all present and former members of the H. and K. Jung labs for creating such a friendly and open working atmosphere. Special thanks to Michi for remembering me that eating is important for survival, for his attempts to teach me the Bavarian dialect, for all the discussions about science, politics, sports, and social life and

many more. I also like to thank him, Käptn and Christoph for the nice time we have spent together on concerts, “Käptn’s dinner”, and other marvelous activities.

I dedicate many thanks to my friends for their encouragement and moral support. I am particularly indebted to Thomas, Robert, Oliver, Higher, Tina, Veronika, Aleks, Malte, and The Unknown Stuntmen for our many fun times and helping me along the way. I apologize for the last months I was not available.

Warmest thanks to my wonderful Julia, who always stood by me, not only in good but also in bad times. She is a truly beautiful person whose presence in my life allows me to look forward to each new day.

Finally and most importantly I would like to thank my family, in particular my parents for believing in me and providing their love, care, and support every step of the way.

CURRICULUM VITAE**Personal Data**

Name: Daniel Hilger
Date of birth: 14.03.1978
Place of birth: Georgsmarienhütte
Family status: Single
Profession: Biologist (Diplom-Biologe)

University

since 08/2004 PhD thesis in the research group of Prof. Dr. Heinrich Jung at the Department of Biology, Microbiology, Ludwig-Maximilians University of Munich
Topic: Electron paramagnetic resonance spectroscopic analyses of membrane transport proteins.

10/1998-06/2004 Biology studies (Diploma) at the University of Osnabrück

06/2004 Graduation (Diploma)

Diploma thesis in the research group of Prof. Dr. Heinrich Jung at the Department of Biology, Microbiology, University of Osnabrück
Title: Topology and conformational dynamics of Na⁺-coupled secondary transporters of *Escherichia coli* – a protein chemical and EPR-spectroscopic analysis.

Community Service

09/1997-09/1998 Community service at the Heilpädagogische Hilfe, Osnabrück

Education

06/1997 University-entrance diploma (Abitur)

Publications

Hilger, D., Jung, H., Padan, E., Wegener, C., Vogel, K.-P., Steinhoff, H.-J., and Jeschke, G. (2005) Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na⁺/H⁺ antiporter of *E. coli*. *Biophys. J.* 89: 1328-1338.

Jung, H., Pirch, T., and Hilger, D. (2006) Secondary transport of amino acids in prokaryotes. *J. Membr. Biol.* 213: 119-133.

Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H. (2006) DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Reson.* 30: 473-498.

Hilger, D., Polyhach, Y., Padan, E., Jung, H., and Jeschke, G. (2007) High-resolution structure of a Na⁺/H⁺ antiporter dimer obtained by pulsed electron paramagnetic resonance distance measurements. *Biophys. J.* 93: 3675-3683.

Hilger, D., Böhm, M., Hackmann, A., and Jung, H. (2008) Role of Ser-340 and Thr-341 in transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli* in ligand binding and transport. *J. Biol. Chem.* 283: 4921-4929.

Raba, M., Baumgartner, T., Hilger, D., Klempahn, K., Härtel, T., Jung, K., and Jung, H. (2008) Function of transmembrane domain IX in the Na⁺/proline transporter PutP. *J. Mol. Biol.* 382: 884-893.

Hilger, D., Polyhach, Y., Jung, H., and Jeschke, G. (2009) Backbone structure of transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli*. *Biophys. J.* 96: 217-225.

Hilger, D., and Jung, H. (2009) Protein chemical and EPR spectroscopic approaches to monitor membrane protein structure and dynamics. In *Bacterial Signaling* (Krämer, R., and Jung, K., eds.) pp. 247-263, Wiley-VCH, Weinheim.

Posters and Conferences

Hilger, D., Jeschke, G., Wegener, C., Steinhoff, H.-J., and Jung, H.. Structure and dynamics of the sodium/proline transporter PutP of *E. coli*: a protein chemical and EPR spectroscopic analysis. *Annual Fall Meeting of GBM, Münster* (2004)

Hilger, D., Jeschke, G., Wegener, C., Steinhoff, H.-J., Padan, E., and Jung, H.. Functional dynamics of the Na⁺/H⁺ antiporter NhaA of *E. coli* probed by EPR spectroscopy. *Annual Fall Meeting of GBM, Münster* (2004)

Hilger, D., Jeschke, G., Wegener, C., Nietschke, M., Steinhoff, H.-J., and Jung, H.. Structure and dynamics of the sodium/proline transporter PutP of *E. coli*: a protein chemical and EPR spectroscopic analysis. *Annual Conference of the Association for General and Applied Microbiology (VAAM), Braunschweig* (2004)

Hilger, D., Jeschke, G., Wegener, C., Vogel, K.-P., Steinhoff, H.-J., Padan, E., and Jung, H.. Effect of pH on the monomer-dimer equilibrium of NhaA probed by EPR-spectroscopy. *1st. Joint German/British Bioenergetics Conference in Cooperation with the GBM study group Bioenergetics "Mechanisms of Bioenergetic Membrane Proteins: Structures and Beyond"*, Naurod (2005)

Hilger, D., Jeschke, G., Steinhoff, H.-J., Böhm, M., Pirch, T., and Jung, H.. Transmembrane domains II and IX play a pivotal role in sodium/proline symport by PutP of *E. coli*. *1st. Joint German/British Bioenergetics Conference in Cooperation with the GBM study group Bioenergetics "Mechanisms of Bioenergetic Membrane Proteins: Structures and Beyond"*, Naurod (2005)

Hilger, D., Pirch, T., Böhm, M., and Jung, H.. The substrate translocation pathway of the sodium/proline transporter PutP of *E. coli*. *Annual Conference of the Association for General and Applied Microbiology (VAAM), Göttingen* (2005)

Hilger, D., Polyhach, Y., Jung, H., and Jeschke G.. Structure determination of TM IX of PutP of *Escherichia coli* by four-pulsed DEER. *Symposium "Membrane Proteins 2006"*, Naurod (2006)

Hilger, D., Polyhach, Y., Padan, E., Jung, H., and Jeschke, G.. Structure determination of membrane proteins by site-directed spin labeling in conjunction with pulsed electron paramagnetic resonance. *EMBO Workshop "New Methods in Membrane Protein Research"*, Stockholm (2007)

Hilger, D., Polyhach, Y., Padan, E., Jung, H., and Jeschke, G.. Structure determination of membrane proteins by site-directed spin labeling in conjunction with pulsed electron paramagnetic resonance. *Munich Interact, PhD-Symposium*, München (2007)

Talks

Hilger, D., Polyhach, Y., Padan, E., Jung, H., and Jeschke, G.. High-resolution structure of the dimer of the Na⁺/H⁺ antiporter NhaA of *Escherichia coli* based on site-directed spin labeling and pulsed EPR distance measurements. *Annual Conference of the Association for General and Applied Microbiology*, Osnabrück (2007)

Hilger, D., Polyhach, Y., Raba, M., Jeschke, G., and Jung, H.. Structure and function of transmembrane domain IX of the Na⁺/proline transporter PutP of *Escherichia coli*. *Joint Annual Conference of the Association for General and Applied Microbiology (VAAM) and the German Society for Biochemistry and Molecular Biology (GBM)*, Frankfurt (2008)

EHRENWÖRTLICHE VERSICHERUNG

Hiermit versichere ich ehrenwörtlich, dass die vorliegende Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt wurde. Des Weiteren erkläre ich, dass ich nicht anderweitig ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen. Die vorliegende Dissertation liegt weder ganz, noch in wesentlichen Teilen einer anderen Prüfungskommission vor.

Daniel Hilger, München