

Isabelle Mus-Veteau *Editor*

Membrane Proteins Production for Structural Analysis

 Springer

Membrane Proteins Production for Structural Analysis

Isabelle Mus-Veteau
Editor

Membrane Proteins Production for Structural Analysis

 Springer

Editor

Isabelle Mus-Veteau
Institute of Molecular and Cellular
Pharmacology, UMR-CNRS 7275,
University of Nice-Sophia Antipolis
Valbonne, France

ISBN 978-1-4939-0661-1

ISBN 978-1-4939-0662-8 (eBook)

DOI 10.1007/978-1-4939-0662-8

Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2014940924

© Springer Science+Business Media New York 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Membrane proteins are involved in fundamental biological processes like ion, water, or solute transport, sensing changes in the cellular environment, signal transduction, and control of cell-cell contacts required to maintain cellular homeostasis and to ensure coordinated cellular activity in all organisms. Because of the importance of these proteins to living cells, their dysfunctions are responsible for numerous pathologies like cancer, cystic fibrosis, epilepsy, hyperinsulinism, heart failure, hypertension, and Alzheimer diseases. However, studies on these and other disorders are hampered by a lack of information about the involved proteins. Knowing the structure of these proteins and understanding their molecular mechanism is not only of fundamental biological interest, but also holds great potential for enhancing human health. This is of paramount importance in the pharmaceutical industry, which produces many drugs that bind to membrane proteins and which recognizes the potential of many recently identified G-protein-coupled receptors (GPCRs), ion channels, and transporters, as targets for future drugs.

Fifty percent of all drug targets are GPCRs, which is one of the largest and most diverse membrane protein families. Whereas high-resolution structures are available for a myriad of soluble proteins (more than 42,000 in the Protein Data Bank), atomic structures have so far been obtained for only 424 membrane proteins. Remarkably, this number is growing exponentially with 100 new structures determined in the last two years. However, only ten percent of membrane protein structures are derived from vertebrates. Indeed, the majority of medically and pharmaceutically relevant mammalian membrane proteins are present in tissues at very low concentration, making production of recombinant proteins in heterologous systems suitable for large-scale production a prerequisite for structural studies. For the majority of mammalian membrane proteins, the production of soluble, stable and correctly folded protein is challenging. The breakthrough occurred in 2005 with the two first atomic structures of recombinant mammalian membrane proteins obtained from proteins overexpressed in yeast: the calcium ATPase from sarcoplasmic reticulum SERCA1A and the Kv1.2 voltage-gated potassium channel. Since then, extensive optimization of heterologous expression systems, stabilization tools, and structural analysis methods has begun to bear fruit, and the structure of 37 recombinant mammalian membrane proteins have been solved.

With this book, we compiled some advances in heterologous expression systems, stabilization tools and structural methods that contributed to the growing number of recombinant integral membrane protein structures solved these in the past few years.

It will also facilitate the structural analysis of many other membrane proteins. I want to thank the authors of each chapter for their contribution to this book. I want to thank the authors of each chapter for their contribution to this book, which should be of strong interest for people who wish to produce membrane proteins for structural analysis.

Isabelle Mus-Veteau

Preface

Structural biology of integral membrane proteins has been in the limelight ever since the first 7 Å resolution three-dimensional structure of bacteriorhodopsin was determined by electron crystallography and published in 1975. Since then, there have been incredible advances in our ability to express any membrane protein in heterologous expression systems and purify them in a functional form suitable for crystallization. Some membrane proteins have proven to be amenable for structural analysis. We now have a wealth of information on the structure and function of bacterial ion channels, transporters, respiratory complexes, and photosynthetic assemblies, which has led to the award of a number of Nobel Prizes, highlighting the importance of these proteins in biology and the difficulty in determining their structures. Nevertheless, structure determination of mammalian membrane proteins has proven much more difficult, but in the last five years there have been dramatic advances in our understanding of why these proteins are more difficult than their bacterial counterparts. This has been demonstrated most graphically with the structure determination of G-protein-coupled receptors (GPCRs), where a series of complementary and generic engineering and crystallization methodologies have been developed in different laboratories around the world, making it possible to determine the structure of any GPCR provided that enough authentically folded receptors can be expressed.

Expression of many integral membrane proteins remains challenging. Human membrane proteins often require molecular chaperones to fold correctly in a process that may take hours, and the proteins may be far less stable than their bacterial homologues. Thus the challenge of producing milligrams of correctly folded protein remains. This volume addresses many of the problems associated with producing membrane proteins and more importantly how to purify them in a functional form using stabilizing detergents and detergent mimetics, allowing subsequent biophysical and structural analyses. Every membrane protein behaves in its own unique fashion, with quirks and peccadilloes enough to make each protein a challenge to express, purify, and crystallize. Thus the more tools we have in our toolbox of protocols for handling membrane proteins, the greater chance we have of making even the most wayward membrane protein behave.

Structural biology of membrane proteins is entering a new era. Electron cryo-microscopy was recently used for the first structure determination of an integral membrane protein to 3.4 Å resolution by single particle analysis; in the next few years as technology develops, this will become easier and promises the possibility of determining structures of any protein over about 250 kDa in size without the need for crystallization. The X-ray free electron laser has shown how high-resolution structures can be determined from micron-sized crystals of membrane proteins using only a few hundred micrograms of purified protein. New developments in electron diffraction of sub-micron crystals also show great promise for future structural analyses. Structure-based drug design for GPCRs is a reality, with multiple structures being determined of a single receptor bound to different drug candidates. However, if you cannot express and purify your membrane protein of interest in a biologically relevant state, then these great advances are superfluous. Thus, there will always be the need for improvements in expression systems and for careful biochemical analysis of the proteins produced.

Christopher G. Tate
16th January 2014

Contents

1 Membrane Protein Production for Structural Analysis	1
Isabelle Mus-Veteau, Pascal Demange and Francesca Zito	
2 Membrane Protein Quality Control in Cell-Free Expression Systems: Tools, Strategies and Case Studies	45
Davide Proverbio, Erik Henrich, Erika Orbán, Volker Dötsch and Frank Bernhard	
3 Bacterial Expression and Stabilization of GPCRs	71
Jean-Louis Banères	
4 Membrane Protein Production in <i>Escherichia coli</i>: Overview and Protocols	87
Georges Hattab, Annabelle Y. T. Suisse, Oana Iliaia, Marina Casiraghi, Manuela Dezi, Xavier L. Warnet, Dror E. Warschawski, Karine Moncoq, Manuela Zoonens and Bruno Miroux	
5 <i>Lactococcus lactis</i>: Recent Developments in Functional Expression of Membrane Proteins	107
Sana Bakari, François André, Daphné Seigneurin-Berny, Marcel Delaforge, Norbert Rolland and Annie Frelet-Barrand	
6 Overexpression of Membrane Proteins in <i>Saccharomyces cerevisiae</i> for Structural and Functional Studies: A Focus on the Rabbit Ca²⁺-ATPase Serca1a and on the Yeast Lipid “Flippase” Complex Drs2p/Cdc50p	133
Cédric Montigny, Hassina Azouaoui, Aurore Jacquot, Marc le Maire, Christine Jaxel, Philippe Champeil and Guillaume Lenoir	
7 Amphipols: A General Introduction and Some Protocols	173
Manuela Zoonens, Francesca Zito, Karen L. Martinez and Jean-Luc Popot	

8	New Amphiphiles to Handle Membrane Proteins: “Ménage à Trois” Between Chemistry, Physical Chemistry, and Biochemistry	205
	Grégory Durand, Maher Abla, Christine Ebel and Cécile Breyton	
9	Building Model Membranes with Lipids and Proteins: Dangers and Challenges	253
	James N. Sturgis	
10	Analytical Ultracentrifugation and Size-Exclusion Chromatography Coupled with Light Scattering for the Characterization of Membrane Proteins in Solution	267
	Aline Le Roy, Cécile Breyton and Christine Ebel	
11	Lipidic Cubic Phase Technologies for Structural Studies of Membrane Proteins	289
	Andrii Ishchenko, Enrique Abola and Vadim Cherezov	
12	Micelles, Bicelles, Amphipols, Nanodiscs, Liposomes, or Intact Cells: The Hitchhiker’s Guide to the Study of Membrane Proteins by NMR	315
	Laurent J. Catoire, Xavier L. Warnet and Dror E. Warschawski	
13	Foundations of Biomolecular Simulations: A Critical Introduction to Homology Modeling, Molecular Dynamics Simulations, and Free Energy Calculations of Membrane Proteins	347
	Jérôme Hénin, Marc Baaden and Antoine Taly	
14	Structural Studies of TSPO, a Mitochondrial Membrane Protein	393
	Jean-Jacques Lacapere, Soria Iatmanen-Harbi, Lucile Senicourt, Olivier Lequin, Piotr Tekely, Rudra N. Purusottam, Petra Hellwig, Sebastien Kriegel, Stephanie Ravaud, Céline Juillan-Binard, Eva Pebay Peyroula and Vassilios Papadopoulos	
	Erratum	E1
	Index	423

Contributors

M. Abla Avignon University, Avignon, France

E. Abola Department of Integrated Structural and Computational Biology, The Scripps Research Institute, La Jolla, California, USA

F. André Oxidative Stress and Detoxification Laboratory, UMR-CNRS 8221, Institute of Biology and Technology of Saclay, SB2SM, and Centre for Nuclear Studies and Université Paris-Suds, Gif-sur-Yvette, France

H. Azouaoui Laboratory of Membrane Proteins, Institute of Biology and Technology of Saclay, UMR-CNRS 8221, Centre for Nuclear Studies and Université Paris-Sud, Gif-sur-Yvette, France

M. Baaden Laboratory of Theoretical Biochemistry, Institute of Physico-Chemical Biology, French National Centre for Scientific Research, Université Paris Diderot, Paris, France

S. Bakari Oxidative Stress and Detoxification Laboratory, UMR-CNRS 8221, Institute of Biology and Technology of Saclay, SB2SM, and Centre for Nuclear Studies and Université Paris-Suds, Gif-sur-Yvette, France

J.-L. Banères Institute of Biomolecules Max Mousseron, UMR-CNRS 5247, Faculty of Pharmacy, Université Montpellier 1 and 2, Montpellier, France

Frank Bernhard Centre for Biomolecular Magnetic Resonance, Institute of Biophysical Chemistry, Goethe University Frankfurt am Main, Frankfurt am Main, Germany

C. Breyton Institute of Structural Biology, French National Centre for Scientific Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

M. Casiraghi Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

- L. J. Catoire** Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France
- P. Champeil** Laboratory of Membrane Proteins, Institute of Biology and Technology of Saclay, UMR-CNRS 8221, Centre for Nuclear Studies and Université Paris-Sud, Gif-sur-Yvette, France
- V. Cherezov** Department of Integrated Structural and Computational Biology, The Scripps Research Institute, La Jolla, California, USA
- M. Delaforge** Oxidative Stress and Detoxification Laboratory, UMR-CNRS 8221, Institute of Biology and Technology of Saclay, SB2SM, and Centre for Nuclear Studies and Université Paris-Suds, Gif-sur-Yvette, France
- P. Demange** Institute of Pharmacology and Structural Biology, UMR-CNRS 5089, Université de Toulouse, Toulouse, France
- M. Dezi** Laboratory of Crystallography and NMR Biology, UMR-CNRS 8015, Université Paris Descartes, Paris, France
- V. Dötsch** Centre for Biomolecular Magnetic Resonance, Institute of Biophysical Chemistry, Goethe University Frankfurt am Main, Frankfurt am Main, Germany
- G. Durand** Avignon University, Avignon, France
Institute of Biomolecules Max Mousseron, UMR-CNRS 5247, Montpellier, France
- C. Ebel** Institute of Structural Biology, French National Centre for Scientific Research, Centre for Nuclear Studies, and Université Grenoble Alpes, Grenoble, France
- A. Frelet-Barrand** Oxidative Stress and Detoxification Laboratory, UMR-CNRS 8221, Institute of Biology and Technology of Saclay, SB2SM, and Centre for Nuclear Studies and Université Paris-Suds, Gif-sur-Yvette, France
- G. Hattab** Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, Université Paris Diderot, Paris, France
- P. Hellwig** Laboratory of Vibrational Spectroscopy and Electrochemistry of Biomolecules, UMR-CNRS 7177, Université de Strasbourg, Strasbourg, France
- J. Héning** Laboratory of Theoretical Biochemistry, Institute of Physico-Chemical Biology, French National Centre for Scientific Research, Université Paris Diderot, Paris, France
- E. Henrich** Centre for Biomolecular Magnetic Resonance, Institute of Biophysical Chemistry, Goethe University Frankfurt am Main, Frankfurt am Main, Germany

S. Iatmanen-Harbi BioMolecules Laboratory, UMR-CNRS 7203, Université Pierre et Marie Curie and Ecole Normale Supérieure, Paris, France

O. Ilioaia Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

A. Ishchenko Department of Integrated Structural and Computational Biology, The Scripps Research Institute, La Jolla, California, USA

A. Jacquot Laboratory of Membrane Proteins, Institute of Biology and Technology of Saclay, UMR-CNRS 8221, Centre for Nuclear Studies and Université Paris-Sud, Gif-sur-Yvette, France

C. Jaxel Laboratory of Membrane Proteins, Institute of Biology and Technology of Saclay, UMR-CNRS 8221, Centre for Nuclear Studies and Université Paris-Sud, Gif-sur-Yvette, France

C. Juillan-Binard Institute of Structural Biology, French National Centre for Scientific Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

S. Kriegel Laboratory of Vibrational Spectroscopy and Electrochemistry of Biomolecules, UMR-CNRS 7177, Université de Strasbourg, Strasbourg, France

J.-J. Lacapere BioMolecules Laboratory, UMR-CNRS 7203, Université Pierre et Marie Curie and Ecole Normale Supérieure Paris, France

M. le Maire Laboratory of Membrane Proteins, Institute of Biology and Technology of Saclay, UMR-CNRS 8221, Centre for Nuclear Studies and Université Paris-Sud, Gif-sur-Yvette, France

A. Le Roy Institute of Structural Biology and The European Molecular Biology Laboratory, Integrated Structural Biology Grenoble, French National Centre for Scientific Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

G. Lenoir Laboratory of Membrane Proteins, Institute of Biology and Technology of Saclay, UMR-CNRS 8221, Centre for Nuclear Studies and Université Paris-Sud, Gif-sur-Yvette, France

O. Lequin BioMolecules Laboratory, UMR-CNRS 7203, Ecole Normale Supérieure, Université Pierre et Marie Curie, Paris, France

K. L. Martinez Bio-Nanotechnology Laboratory, Department of Neuroscience and Pharmacology & Nano-Science Center, University of Copenhagen, Copenhagen, Denmark

B. Miroux Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

K. Moncoq Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

C. Montigny Laboratory of Membrane Proteins, Institute of Biology and Technology of Saclay, UMR-CNRS 8221, Centre for Nuclear Studies and Université Paris-Sud, Gif-sur-Yvette, France

I. Mus-Veteau Institute for Molecular and Cellular Pharmacology, UMR-CNRS 7275, University of Nice-Sophia Antipolis, Valbonne, France

E. Orbán Centre for Biomolecular Magnetic Resonance, Institute of Biophysical Chemistry, Goethe University Frankfurt am Main, Frankfurt am Main, Germany

V. Papadopoulos The Research Institute of the McGill University Health Center, Department of Medicine, McGill University, Montreal, QC, Canada

E. Pebay-Peyroula Institute of Structural Biology, French National Centre for Scientific Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

J.-L. Popot Institute of Physico-Chemical Biology, UMR-CNRS 7099, Université Paris Diderot, Paris, France

D. Proverbio Centre for Biomolecular Magnetic Resonance, Institute of Biophysical Chemistry, Goethe University Frankfurt am Main, Frankfurt am Main, Germany

R. N. Purusottam BioMolecules Laboratory, UMR-CNRS 7203, Université Pierre et Marie Curie and Ecole Normale Supérieure, Paris, France

S. Ravaud Institute of Structural Biology, French National Centre for Scientific Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

N. Rolland Cell & Plant Physiology Laboratory, Institute of Sciences Research and Technologies, UMR-CNRS 5168, National Institute of Agronomical Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

D. Seigneurin-Berny Cell & Plant Physiology Laboratory, Institute of Sciences Research and Technologies, UMR-CNRS 5168, National Institute of Agronomical Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

L. Senicourt BioMolecules Laboratory, UMR-CNRS 7203, Université Pierre et Marie Curie and Ecole Normale Supérieure Paris, France

J. N. Sturgis Engineering Laboratory of Macromolecular Systems, UMR-CNRS 7255 and Aix-Marseille University, Marseille, France

A. Y. T. Suisse Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

A. Taly Laboratory of Theoretical Biochemistry, Institute of Physico-Chemical Biology, French National Centre for Scientific Research, Université Paris Diderot, Paris, France

P. Tekely BioMolecules Laboratory, UMR-CNRS 7203, Université Pierre et Marie Curie and Ecole Normale Supérieure, Paris, France

X. L. Warnet Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

D. E. Warschawski Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

F. Zito Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

M. Zoonens Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

List of Abbreviations

3D	three-dimensional
AA	amino acids
ABC	ATP-binding cassette
ABF	adaptive biasing force
Ach	acetylcholine
AChBP	acetylcholine binding protein
ADP	adenosine diphosphate
AMPA	α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid
APO	10, dimethyldecylphosphine oxide
APol	amphipathic polymers
APol	amphipol
ATD	amino-terminal domain
ATP	adenosine triphosphate
AUC	analytical ultracentrifugation
BAD	biotin acceptor domain
BCA	bicinchoninic acid
BR	bacteriorhodopsin
Brij35	polyoxyethylene-(23)-lauryl-ether
Brij58	polyoxyethylene-(20)-cetyl-ether
Brij72	polyoxyethylene-(2)-stearyl-ether
Brij78	polyoxyethylene-(20)-stearyl-ether
Brij98	polyoxyethylene-(20)-oleyl-ether
Bz-ATP	benzoyl-benzoyl-ATP
C12E7	heptaethylene glycol monododecyl ether
C12E8	octaethylene glycol monododecyl ether
CAC	critical aggregation concentration
CD	circular dichroism
CDM	chemically defined medium
cDNA	complementary DNA
CF	cell free
CFE	cell-free expression
CHAPS	3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonat

CHAPSO	cholamidopropyl-dimethylammonio-hydroxy-propanesulfonate
CHS	cholesteryl hemisuccinate
CMC	critical micellar concentration
CPP	critical packing parameter
CXCR	chemokine receptor
CYMAL-5	5-cyclohexyl-1-pentyl- β -D-maltoside
Cymal6	6-cyclohexyl-1-hexyl- β -D-maltoside
DAG	diacylglycerol
DAPol	partially deuterated A8-35 amphipols
DDM	n-dodecyl- β -D-maltopyranoside
DHPC	1,2-diheptanoyl-sn-glycero-3-phosphocholine
diC7PC	1,2-diheptanoyl-sn-glycero-3-phosphocholine
Digit	digitonin
DLS	dynamic light scattering
DM	n-decyl- β -D-maltopyranoside
DMPA	1,2-ditetradecanoyl-sn-glycero-3-phosphate
DMPC	dimyristoyl phosphatidylcholine
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DMPE	1,2-ditetradecanoyl-sn-glycero-3-phosphoethanolamine
DMPG	1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
DNA	deoxyribonucleic acid
DNP	dynamic nuclear polarization
DOPA	1,2-dioleoyl-sn-glycero-3-phosphate
DOPC	1,2-dioleoyl-sn-glycero-3-phosphocholine
DOPE	1,2-dioleoyl-sn-glycero-3-phosphoethanolamine
DOPG	1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
DPC	n-dodecylphosphocholine
C12-PC	n-dodecylphosphocholine
DTPC	ditridecanoyl phosphatidylcholine
<i>E. coli</i>	<i>Escherichia coli</i>
ECD	extracellular domain
ECF	energy coupling factor
EM	electron microscopy
EPR	electronic paramagnetic resonance
ER	endoplasmic reticulum
F6-DigluM	N-1,1-di[(O- β -D-glucopyranosyl)oxymethyl]ethyl-4-thia-7,7,8,8,9,9,10,10,11,11,12,12,12-tridecafluorododecanamide
FAPol	fluorescently-labelled amphipol
FC-12	FOS choline 12 or N-dodecylphosphocholine
Fos14	n-tetradecylphosphocholine
Fos16	n-hexadecylphosphocholine
FS	fluorinated surfactants
F-surfactant	fluorinated surfactant
GABA	gamma aminobutyric acid
GFP	green fluorescent protein

GpA	glycophorin A
GPCR	G protein-coupled receptor
GST	glutathione S-transferase
HDM	n-hexadecyl- β -D-maltopyranoside
HMG-CoA	3-hydroxy-3-methylglutaryl coenzyme A
IB	inclusion body
IEC	ion exchange chromatography
iGluR	ionotropic glutamate receptor
IMAC	immobilized-metal affinity chromatography
IMP	integral membrane protein
IMV	inner membrane vesicles
INV	inverted vesicles
IPTG	isopropyl β -D-1-thiogalactopyranoside
kDa	kilodaltons
LAB	lactic acid bacteria
LBD	ligand-binding domain
LCP	lipidic cubic phase
LDAO	N,N-dimethyldodecylamine N-oxide
C12-DAO	N,N-dimethyldodecylamine N-oxide
LDAO	lauryldimethylamine oxide
LGIC	ligand-gated ion channel
LIC	ligation independent cloning
LMPC	1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac(1-choline)]
LMPG	1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)]
LPC	L- α -lysophosphatidylcholine
LPPG	1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)]
LPS	lipo-polysaccharides
LS	n-lauroyl sarcosine
MAG	mono acyl glycerol
MALS	multi-angle laser light scattering
MBP	maltose binding protein
MCD	methyl- β -D-cyclodextrin
MCS	multi-cloning site
MD	molecular dynamics
MFS	major facilitator superfamily
PBSA	Poisson Boltzmann surface accessibility
MNG	maltose-neopentyl glycol
MOMP	major outer membrane protein
MP	membrane proteins
MR	molecular replacement
MS	microsomes
MSP	membrane scaffold protein
nAChR	nicotinic acetylcholine receptor
NADPH	nicotinamide adenine dinucleotide phosphate reduced
NApol	neutral amphipathic polymers

NaPol	non-ionic amphipols
NBD	nucleotide binding domain
NCS-ATP	8-thiocyano-ATP
NG	n-nonyl- β -D-glucopyranoside
NICE	NIsin controlled gene expression
Ni-NTA	Ni ²⁺ -nitrilotriacetic acid
NMA	normal mode analysis
NMDA	N-methyl-D-aspartate
NMR	nuclear magnetic resonance
nOG	n-octyl- β -D-glucopyranoside
Nvoy	NV10 polymer
OG	octyl- β -D-glucopyranoside
OG or C8-G	n-octyl- β -D-glucoside
OGCP	oxoglutarate carrier protein
OMV	outer membrane vesicles
ONGPG	octyl glucose neopentyl glycol
OTG	n-octyl- β -D-glucopyranoside
P2X	ATP-gated receptor
PC	L- α -phosphatidylcholinePE, L- α -phosphatidylethanolamine
PC-APol	phosphocholine-based amphipol
PCC- α -M	propylcyclohexyl cyclohexyl- α -D-maltoside
PDB	Protein Data Bank
PDC	protein detergent complex
PE	phosphatidylethanolamine
perDAPol	perdeuterated A8-35 amphipols
PL	polar lipid extract
pLGIC	pentameric ligand-gated ion channel
PMF	potential of mean force
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
POPG	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
PS	phosphatidylserine
RDC	residual dipolar coupling
RI	refractive index
RNA	ribonucleic acid
RS	Stokes radius
SANS	small-angle neutron scattering
SAPol	sulfonated amphipol
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate poly acrylamide gel electrophoresis
SEC	size exclusion chromatography
SEC/MALS	size exclusion chromatography coupled to static and dynamic light scattering, absorbance and refractive index detections
SERCA	sarco/endoplasmic reticulum Ca ²⁺ -ATPase
SF	selectivity filter

SFX	serial femtosecond crystallography
SMD	steered molecular dynamics
SPR	surface plasmon resonance
ssNMR	solid state NMR
SV	sedimentation velocity
T4L	T4 lysozyme
TEM	transmission electronic microscopy
TEV	tobacco etch virus
TGN	trans-Golgi network
TL	total lipid extract
TM	transmembrane
TMP	total membrane proteins
TMRM	tetramethyl-rhodamine maleimide
TNP	trinitrophenyl
tOmpA	transmembrane domain of the outer membrane protein A from <i>Escherichia coli</i>
TRIS	tris(hydroxymethyl)amidomethane
TROSY	transverse relaxation optimized spectroscopy
Trp	tryptophan
TRX	thioredoxin
TX-100	Triton X-100
UDM	undecyl- β -D-maltoside
UV	ultraviolet
VBEx	vector backbone exchange
WALP	tryptophan, alanine and leucine peptide
WT	wild type
XFEL	X-ray free-electron laser
ZIREX	zinc regulated expression system
β -DDM or C12-M	n-dodecyl- β -D-maltoside

About the Editor

Dr. Isabelle Mus-Veteau is a biochemist and biophysicist specialist in membrane protein characterization. She obtained her PhD in microbiology and cell biology at the University of Marseille in France. She has held a French National Centre for Scientific Research (CNRS) tenure position at the Institute of Molecular and Cellular Pharmacology (IPMC, Sophia Antipolis near Nice, France), where she supervises projects on the characterization of the Hedgehog receptor Patched. She was a member of the CNRS tenure position recruitment committee and is currently a member of the executive committees of the French Biophysical Society and of the Membrane Group Society. She organized two international summer schools on membrane protein production for structural analysis and two international congresses on membrane biophysics.

Chapter 1

Membrane Protein Production for Structural Analysis

Isabelle Mus-Veteau, Pascal Demange and Francesca Zito

1.1 Introduction

Integral membrane proteins (IMPs) account for roughly 30% of all open reading frames in fully sequenced genomes (Liu and Rost 2001). These proteins are of main importance to living cells. They are involved in fundamental biological processes like ion, water, or solute transport, sensing changes in the cellular environment, signal transduction, and control of cell–cell contacts required to maintain cellular homeostasis and to ensure coordinated cellular activity in all organisms. IMP dysfunctions are responsible for numerous pathologies like cancer, cystic fibrosis, epilepsy, hyperinsulinism, heart failure, hypertension, and Alzheimer diseases. However, studies on these and other disorders are hampered by a lack of information about the involved IMPs. Thus, knowing the structure of IMPs and understanding their molecular mechanism not only is of fundamental biological interest but also holds great potential for enhancing human health. This is of paramount importance in the pharmaceutical industry, which produces many drugs that bind to IMPs, and recognizes the potential of many recently identified G-protein-coupled receptors (GPCRs), ion channels, and transporters, as targets for future drugs. GPCR, which account for 50% of all drug targets, is one of the largest and most diverse IMP families encoded by more than 800 genes in the human genome (Fredriksson et al. 2003; Lundstrom 2006). However, whereas high-resolution structures are available for a myriad of soluble proteins (more than 42,000 in the Protein Data Bank,

I. Mus-Veteau (✉)

Institute for Molecular and Cellular Pharmacology, UMR-CNRS 7275, University of Nice-Sophia Antipolis, Valbonne, France
e-mail: mus-veteau@ipmc.cnrs.fr

P. Demange

Institute of Pharmacology and Structural Biology, UMR-CNRS 5089, Université de Toulouse, Toulouse, France

F. Zito

Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099, Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France

PDB), atomic structures have so far been obtained for only 474 IMPs, with 150 new structures determined in 2012 and 2013 (see <http://blanco.biomol.uci.edu/mp-struct/>), but only 10% of the unique IMP structures are derived from vertebrates. The first mammalian IMPs were crystallized on to their natural abundance, circumventing all the difficulties associated with overexpression (ATP synthase, Stock et al. 1999; rhodopsin, Palczewski et al. 2000; and Calcium ATPase, Toyoshima et al. 2000). However, the majority of medically and pharmaceutically relevant IMPs are present in tissues at very low concentration, making production of recombinant IMPs in heterologous systems suitable for large-scale production a prerequisite for structural studies. In 2005, the two first atomic structures of recombinant mammalian IMPs were obtained from proteins overexpressed in yeast: the calcium ATPase from sarcoplasmic reticulum SERCA1A in *Saccharomyces cerevisiae* (Jidenco et al. 2005) and the Kv1.2 voltage-gated potassium channel in *Pichia pastoris* (Long et al. 2005). Since then, extensive optimization of heterologous expression systems (Mus-Veteau 2010) has begun to bear fruit, and early 2014 the structure of 37 recombinant mammalian IMPs were determined, of which 20 belong to GPCRs (Table 1.1; Venkatakrishnan et al. 2013). Impressive progress has been made in the 2012 and 2013 with more than 21 new structures of recombinant mammalian IMPs determined, 13 belonging to GPCRs. The large majority of these structures were obtained from IMPs overexpressed in Sf9 insect cells using recombinant baculovirus (21 proteins over the 37, 14 being GPCRs). From IMPs produced in the yeast ten structures were obtained from *Pichia pastoris* and one from *Saccharomyces cerevisiae*, three were determined from proteins produced in the bacteria *Escherichia coli*, two from proteins expressed in mammalian cells. Concurrently, with the advances in recombinant mammalian IMP production, improvement of the stabilization strategies of IMPs in solution has contributed to the growing number of IMP structures solved. Indeed, purification of IMPs requires the use of detergents to extract IMPs from membrane and to maintain them in aqueous solution (in complex with detergents and lipids). Many mammalian IMPs are unstable in detergent solution, and finding suitable detergent and conditions that ensure protein homogeneity, functionality, stability, and crystallization is often a limiting and crucial step (Tate 2010). New surfactants able to maintain IMPs in solution with less denaturing effect have been synthesized and are currently under development (Chap. 7 by Zoonens et al. and Chap 8 by Durand et al. in this volume). Among these new surfactants, Maltose-neopentyl glycol (MNG; Chae et al. 2010) allowed the crystallization and structure determination of δ (Manglik et al. 2012) and μ (Granier et al. 2012) opioid receptors. Another strategy that has been shown to be highly efficient to stabilize IMPs in solution is protein engineering. For the majority of the structures solved, N- and C-terminus, which are usually flexible or not ordered structures that can prevent crystallization, have been truncated (Table 1.1). For GPCRs which are highly instable in solution, truncation of N- and C-terminal domains were not sufficient, additional strategies were necessary to stabilize these proteins: (1) the replacement of a flexible loop by a stable soluble protein domain such as T4 lysozyme (T4L) or apocytochrome b_{562} RIL has been a successful strategy for the determination of a dozen of IMP structures, (2) thermostabilizing point mutations, (3) engineering

disulfide bridges and *N*-glycosylation sites, or (4) a mixture of these strategies, e.g., serotonin receptor and smoothed structures were obtained using apocytochrome b_{562} RIL and thermostabilizing point mutations (Wacker et al. 2013; Wang et al. 2013); chemokine receptor CXCR4 (Wu et al. 2010), dopamine D3 receptor (Chien et al. 2010), and neurotensin receptor (White et al. 2012) structures were obtained using T4L and thermostabilizing point mutations; and the nociceptin/orphanin receptor (NOP) structure was solved using both apocytochrome b_{562} RIL and T4L fusions (Table 1.1; Thompson et al. 2012). Often, it was necessary to use ligands, agonists, or antagonists to enhance receptor stability. Improvement in structural analysis techniques for IMPs also contributed to the exponential growth of the number of mammalian IMP atomic structures observed these past few years. X-ray crystallography is the technique which let the resolution of the most structures, and advances in micro-crystallography have allowed obtaining higher-resolution diffraction from smaller crystals (Moukhametzianov et al. 2008). Vapor diffusion with hanging drops is the most commonly used crystallization strategy. However, use of lipidic cubic phase (LCP; Caffrey 2009) and new detergents such as MNG (Chae et al. 2010) has improved the likelihood of obtaining crystals. This strategy was applied for most of the GPCR structures. Electron diffraction allowed the structure resolution at atomic level of only one mammalian recombinant IMP, the aquaporin AQP4 (Hiroaki et al. 2005; Tani et al. 2009). Recent progresses in solution and solid-state nuclear magnetic resonance (ssNMR) methods (Maslennikov and Choe, 2013) have permitted the determination of the atomic structure of three mammalian recombinant IMPs: the chemokine receptor CXCR1 (Park et al. 2012), the mitochondrial uncoupling protein UCP2 (Berardi et al. 2011), and the phospholamban (Oxenoid and Chou 2005; Verardi et al. 2011).

Table 1.1 reports the strategies that allowed determination of recombinant mammalian IMPs. Lots of tools and strategies in the field of heterologous expression systems, stabilization, and structural analyses are still under development. This chapter introduces the tools developed in the past few years to increase the number of atomic structure of recombinant mammalian IMPs.

1.2 Production of Recombinant IMPs

It is clear that the first and probably the narrowest bottleneck in IMP expression is the production of abundant quantity of material. It is true, indeed, that the majority of structures for vertebrate IMPs were solved from native material and not from recombinant ones (Stock et al. 1999; Palczewski et al. 2000; Toyoshima et al. 2000). On the other hand, the larger part of medically and pharmaceutically relevant MPs is found at very low concentration, thus rendering overexpression of recombinant MPs essentials for large-scale production for structural studies. In the recent years, the panel of possibilities for overexpression of IMPs has become larger and larger, from *E. coli* to insect and mammalian cells passing by yeast systems and *in vitro* production, in order to create the “right expression system” for each protein. Indeed,

IMPs are very different in structure and physical–chemical properties, thus making it difficult to predict the good approach. In any case, each system has pros and cons, and the choice is often empirical, especially with regard to the levels of functional protein expression. In other words, the more we have the better it is.

The most widely used organism is still *E. coli* (for a review, see Sahdev et al. 2008), and since it presents disadvantages of improper folding with inclusion body (IB) formation, up-to-date efforts are concentrated to create new strains able to allow improved control on protein expression. The best-known strains are probably those of Miroux (Miroux and Walker 1996; see also for a review, Chap. 4 by Hattab et al. in this volume) that display internal membrane proliferation in which all the overexpressed proteins are located (Arechaga 2000). The results from Walker’s laboratory still constitute the basement for engineering new strains. This is the case of the Lemo21(DE3) strain (Schlegel et al. 2012) in which the modulation of the activity of the T7 RNAP by the T7 lysozyme is the key to optimize the ratio of IMPs properly inserted in the cytoplasmic membrane to noninserted proteins. In this strain, maximizing the yields of IMPs is accompanied by reduction of the harmful effects of MP overexpression, resulting in stable overexpression. Moreover, IMPs produced in Lemo21(DE3) can be used for functional and structural studies demonstrating that the overexpressed material is not only inserted in the cytoplasmic membrane but also properly folded.

Another approach, which seems promising for eukaryotic IMPs, relies on new fusion protein expression of the amphipathic Mystic protein as a cargo to drive IMPs to the membrane (Roosild et al. 2005). This strategy was used to express the chloroplast ATP/ADP transporter from *Arabidopsis thaliana* (NTT1) and characterize its transport properties (Deniaud et al. 2011). NTT1 fused to Mystic has a very low transport activity, which can be recovered after *in vivo* Mystic fusion cleavage. Therefore, if one considers the high yield of mature NTT1 obtained via the Mystic fusion approach, this becomes a valid approach for obtaining quantities of pure and active proteins that are adequate for structural studies.

Besides *E. coli*, there are other bacteria that seem to be well adapted to IMP overexpression, such as *Lactococcus lactis* and *Rhodobacter sphaeroides*. The first one is a nonpathogenic and noninvasive lactic acid Gram-positive bacterium (for a review, see Junge et al. 2008; Frelet-Barrand et al. 2010a, 2010b; see also Chap. 5 in this volume). The recombinant proteins are expressed under the control of the Nisin-inducible promoter (NICE system), and the access to new technology as the gateway one renders possible efficient cloning strategies. Recently, it has been shown that the Mystic fusion can be used to facilitate high-yield production also in *L. lactis* (Xu et al. 2013).

The prokaryotic systems we have described above are the ones giving the largest amount of recombinant proteins, even if often not in a functional form, since post-translational modifications are needed. Many eukaryotic IMPs are indeed unstable during purification and detergent manipulation even if the barrier of the poor overexpression can be overcome with generic strategies (Bill et al. 2011). Moreover, there are bacteria able to overcome this step, as an *E. coli* strain that allows glyco-

sylation (Chen et al. 2012), but nevertheless it is often necessary to turn to eukaryotes to assure the best ratio of produced versus functional protein.

The tendency is now to use systems which are closer and closer to the native environment of the protein we want to express: yeast, plant, insect, and mammalian cells seem to trace an approach route to the most effective strategy.

Yeasts are able to perform various post-translational modifications including proteolytic processing of signal sequences, disulfide-bond formation, acylation, prenylation, phosphorylation, and certain types of glycosylation that are crucial for activity and folding. Although the lipid composition of yeast membranes is similar to higher eukaryotes, the absence of specific sterols, e.g., cholesterol for mammalian proteins, might affect protein functionality. Ergosterol, the predominant sterol in yeast, compensates for some mammalian proteins, but for full activity, presence of cholesterol might be essential (Tate et al. 1999). *S. cerevisiae* has been largely employed for eukaryotic MPs due to the well-known genetic composition, the large number of available strains, and the facility to tune expression via inducible promoters. Nevertheless, *S. cerevisiae* is not easy to handle in fermenter condition, and for this reason, attention has been focused on other yeasts such as *P. pastoris* that can reach high cell density and have the same “good heterologous MP producer potential” as *S. cerevisiae*. Quite a lot of efforts are now concentrated to decipher the physiological response of yeasts to MPs’ overexpression. To this end, Bonander et al. (2005) studied the impact of pH and temperature on the expression of Fps1p, a glycerol uptake/efflux facilitator, and they demonstrated that optimal conditions for growing are not always profitable for functional, overexpressed MPs. Optimizing culture conditions to the end of MPs’ production seems to be a real opportunity for both *P. pastoris* and *S. cerevisiae* (for a review Bonander and Bill 2012).

Going upward to the “best adapted” expression system, it is easy to run into baculovirus/insect cell expression system, which is a kind of compromise between bacterial and mammalian cells. This system is well adapted for eukaryotic proteins because of similar codon usage, better expression levels, and fewer truncated proteins. Moreover, insect cells allow post-translational modifications that are closer to those of mammalian cells than those produced by bacteria or even yeast (Jarvis and Finn 1995). Briefly, the baculovirus system relies on the infection of insect cell lines (usually Sf9, Sf2) by recombinant viruses encoding the gene(s) of interest. Improvements in recombinant baculovirus generation have been implemented over the past two decades (Condreay and Kost 2007), including a system (BacMam, In-vitrogen) which allows baculovirus-based expression in mammalian cells.

Recently, structures of mammalian IMPs as bovine rhodopsin (Standfuss et al. 2007; Standfuss et al. 2011; Deupi et al. 2012) and the human ammonia transporter RhCG (Gruswitz et al. 2010) have been determined upon overexpression in mammalian cells (Table 1.1). Different approaches using transient or stables cell lines could be used, the first being relatively rapid to settle and the second much longer. Moreover, the choice of the promoter—inducible or constitutive—together with the choice of the cell line is critical for the good issue of the study (Andrell and Tate 2013).

A paradigm of the importance to find the right expression system has been described for the serotonin transporter (SERT; Tate et al. 2003). The most detailed

study on different expression systems on a unique protein, indeed, is the one on SERT. This protein needs glycosylation to be correctly folded in the presence of calnexin (Tate et al. 1999), and has a strict requirement for cholesterol. For evident reasons, *E. coli* and yeast are inappropriate to overexpress SERT, and insect cells, even though they have the entire requirement for a correct expression, led to an inactive protein (Tate et al. 1999). The same group turned the effort to mammalian systems (Tate et al. 2003), and they succeeded with the tetracycline-inducible system (Andrell and Tate 2013).

Last but not the least, cell-free (CF) systems are evolving as a valid alternative in IMP expression. The most classical extracts from *E. coli* and wheat germ are used routinely for structural approaches as they tolerate additives for the co-translational solubilization of CF system-expressed IMPs. The palette of molecules for protein solubilization in CF systems is now very varied and is increased by the possibility of compounds mixture (Park et al. 2011; Bazzacco et al. 2012; Junge et al. 2010; Ma et al. 2011). A valid alternative to tensioactifs is the insertion in lipid bilayer or nanodiscs (Periasamy et al. 2012; Roos et al. 2012). In the past years, CF system-produced MPs have been used for structural studies, particularly NMR spectroscopy as the case of the C-terminal fragment of human presenilin-1, a subunit of the γ -secretase complex (Sobhanifar et al. 2010b). CF systems also allow the expression and the structural evaluation of membrane complexes such as ATP synthase, which results in a fully assembled complex in the CF system in the presence of detergents (Matthies et al. 2011). CF system-expressed MPs also gave some promising results in crystallization: the human voltage-dependent anion channel-1 structure was solved at low resolution (Deniaud et al. 2010) and the structure of a eukaryotic rhodopsin was solved after *in vitro* production in the presence of a mixture of lipids and detergents (Wada et al. 2011). CF systems are then a powerful approach to produce difficult proteins such IMPs, and their development passes through efficient CF extract sources, which are essential for the preparative-scale CF system production of post-translationally modified proteins (for details, see Chap. 2 by Proverbio et al. in this volume).

In conclusion overexpression of IMPs is, in a sense, fighting against evolution since most IMPs have not evolved to be abundant a few thousand copies per cell. Is that the reason why we are obliged to circumvent the bottleneck and look for new methodology strategies? As we said before, “the more we have the better it is” seems to be a good sentence...

1.3 Stabilization of Solubilized IMPs for Structural Analysis

Aqueous solubilization, necessary for structural analysis, generally requires a detergent to shield the large lipophilic surfaces displayed by IMPs. Unfortunately, IMPs tend to denature, aggregate, or remain unstable in detergents. The poor stability of the detergent-solubilized IMP in a form that is amenable for crystal formation

Table 1.1 Recombinant mammalian IMPs for which the structure has been solved

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
<i>G</i> protein-coupled receptors							
<i>Rhodopsin</i> (two structures)							
N2C/D282C mutant		<i>Bos taurus</i>	COS cells	Disulfide bond between the N-terminus and loop E3 (enhanced thermostability: 10 °C)	C8E4	X-ray: vapor diffusion and sitting drops 3.4 Å	Standfuss et al. (2007)
M257Y mutant	Constitutively active meta-II state, in complex with G α CT	<i>B. taurus</i>	HEK2935-GnTI-cells	Constitutively active mutation and agonist peptide G α CT	OG	X-ray: vapor diffusion and sitting drops 3.30 Å	Deupi et al. (2012)
<i>β1 adrenergic receptor</i> (5 structures)							
β 1AR36-m23: removal of flexible regions at the N and C terminus, and in the cytoplasmic loop 3, 8 point mutations	Bound to antagonist cyanopindolol	Turkey	Insect cells infected by baculovirus	Six point mutations (enhanced thermostability: 21 °C) and antagonist	OG	X-ray: vapor diffusion/hanging drops 2.7 Å	Warne et al. (2008)
β 1AR36-m23	Bound to agonists carmotero/isoprenalin	Turkey	Insect cells infected by baculovirus	Six point mutations (enhanced thermostability: 21 °C) and agonist	Hega-10 and CHS	X-ray: vapor diffusion/hanging drops 2.50 Å	Warne et al. (2011)
β 1AR36-m23	Inactive state bound to antagonist carazolol	Turkey	Insect cells infected by baculovirus	Six point mutations (enhanced thermostability: 21 °C) and antagonist	Hega-10	X-ray: vapor diffusion/hanging drops 3.00 Å	Moukhamet-zianov et al. (2011)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
β 1AR36-m23	Bound to carvedilol (a biased agonist and beta blocker)	Turkey	Insect cells infected by baculovirus	Six point mutations (enhanced thermo-stability: 21 °C) and antagonist	Hega-10	X-ray: vapor diffusion/sitting drops 2.30 Å	Warne et al. (2012)
β 1 adrenergic receptor oligomer	Ligand-free basal state in a lipid membrane-like environment	Turkey	Insect cells infected by baculovirus		DDM + lipids POPC/POPE/ POPG/ cholesterol	X-ray: vapor diffusion/hanging drops 3.50 Å	Huang et al. (2013)
<i>β2 adrenergic receptor (seven structures)</i>							
β 2AR365-Fab5 complex	Bound to an inverse agonist	<i>Homo sapiens</i>	Insect cells infected by baculovirus	Increasing the polar surface by interaction with Fab5 from monoclonal antibody	Bicelles composed of the lipid DMPC and the detergent CHAPSO	X-ray: vapor diffusion/hanging drops 3.4/3.7 Å	Rasmussen et al. (2007)
Methylated β 2AR365-Fab5 complex		<i>H. sapiens</i>	Insect cells infected by baculovirus	Interaction with Fab5	DDM + CHS	X-ray and NMR spectroscopy 3.4 Å	Bokoch (2010)
β 2AR-T4L: T4L replaces third intracellular loop	Bound to the partial inverse agonist carazolol	<i>H. sapiens</i>	Insect cells infected by baculovirus	Cholesterol		X-ray with cholesterol-doped lipid cubic phase (LCP) 2.4 Å	Cherezov et al. (2007)
β 2AR-E122W-T4L: C-terminus truncated	Bound to cholesterol and partial inverse agonist timolol	<i>H. sapiens</i>	Insect cells infected by baculovirus	Cholesterol and the partial inverse agonist timolol		X ray with LCP 2.8 Å	Hanson et al. (2008)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
β 2AR-T4L	In complex with a "novel inverse agonist"	<i>H. sapiens</i>	Insect cells infected by baculovirus			X-ray with LCP 2.84 Å	Wacker et al. (2010)
β 2AR-T4L	Agonist-bound active state	<i>H. sapiens</i>	Insect cells infected by baculovirus	Nanobody and agonist BI-167107		X-ray: LCP with 7.7 MAG containing 10% cholesterol 3.50 Å	Rasmussen et al. (2011)
β 2AR-T4L-H93C with agonist covalently linked by a disulfide bond	Agonist-bound active state	<i>H. sapiens</i>	Insect cells infected by baculovirus	Agonist covalently linked		X-ray with cholesterol-doped monoolenin cubic phase 3.50 Å	Rosenbaum et al. (2011)
β 2AR-Gs protein complex with T4L fused to the amino terminus of the β 2AR	Active state ternary complex: agonist/monomeric β (2)AR/nucleotide-free Gs heterotrimer	<i>H. sapiens</i>	Insect cells infected by baculovirus	Nanobody (Nb35) that binds at the interface between the G α and G β subunits		X ray: LCP with MAG7.7 3.20 Å	Rasmussen et al. (2011)
<i>A_{2A} adenosine receptor (six structures)</i> A_{2A} AR-T4L- Δ C: T4L inserted between TM V and VI	In complex with a high-affinity subtype-selective antagonist ZM241385	<i>H. sapiens</i>	Insect cells infected by baculovirus			X-ray: LCP with monoolenin, and cholesterol 2.6 Å	Jaakola et al. (2008)
A_{2A} AR-T4L- Δ C	With bound agonist (UK-432097)	<i>H. sapiens</i>	Insect cells infected by baculovirus			X-ray: LCP with monoolenin, and cholesterol 2.71 Å	Xu et al. (2011)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
A _{2A} AR-GL31 with four thermostabilising point mutations (L48A, A54L, T65A and Q89A3) and N-glycosylation site mutation N154A	Intermediate conformation between the inactive and active states with bound adenosine	<i>H. sapiens</i>	Insect cells infected by baculovirus	Thermostabilising mutations	OG + CHS	X-ray: vapor diffusion 3.00 Å	Lebon et al. (2011)
A _{2A} -StaR2: A _{2A} R-GL31, 8 thermostabilising mutations, truncation of the C-terminus	Inactive state conformation in complex with caffeine	<i>H. sapiens</i>	Insect cells infected by baculovirus	Thermostabilising mutations A54 L, T88A, R107A, K122A, L202A, L235A, V239A, S277A	NG	X-ray: vapor diffusion 3.60 Å	Doré et al. (2011)
A _{2A} AR ^{N154Q} with N-glycosylation site mutation N154A	In complex with inverse-agonist and antibody	<i>H. sapiens</i>	<i>Pichia pastoris</i>	Fab-fragment and antagonist (ZM241385)	OG	X-ray: vapor diffusion/hanging drops 2.70 Å	Hino et al. (2012)
A _{2A} AR-BRIL-ΔC: third intracellular loop replaced by apocytochrome b(562)RIL	In complex with a high-affinity subtype-selective antagonist ZM241385	<i>H. sapiens</i>	Insect cells infected by baculovirus	Apocytochrome b562RIL (BRIL)	OG	X ray: LCP with monoolein and cholesterol 1.80 Å	Liu et al. (2012a)
CXCR1 chemokine receptor	Active receptor reconstituted in phospholipid bilayers and bound to interleukin-8 (IL-8)	<i>H. sapiens</i>	<i>Escherichia coli</i>	Refolding in DMPC proteoliposomes		Magic angle spinning (MAS) and oriented-sample (OS) solid-state NMR	Park et al. (2012)
¹³ C/ ¹⁵ N-labeled full-length GST-CXCR1							

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
<i>CXCR4 chemokine receptor</i>							
Replacement of the third cytoplasmic loop (ICL3) with T4L and thermostabilizing L125W mutation	Complexed with IT1t antagonist	<i>H. sapiens</i>	Insect cells infected by baculovirus	Thermostabilizing L125W mutation		X-ray: LCP with monolein, and cholesterol 2.5 Å	Wu et al. (2010)
<i>Dopamine D3 receptor</i>							
D3R-T4L: replacement of the third cytoplasmic loop (ICL3) with T4L and thermostabilizing mutation L119W	Complexed with D2/D3-selective antagonist	<i>H. sapiens</i>	Insect cells infected by baculovirus	Thermostabilizing mutation L119W and antagonist eticlopride		X-ray: LCP with monolein, and cholesterol 2.89 Å	Chien et al. (2010)
<i>Histamine H1 receptor</i>							
H1R-T4L: replacement of the third cytoplasmic loop with T4L and deletion of 19 N-terminal residues-terminal residues	Complexed with antagonist doxepin	<i>H. sapiens</i>	<i>P. pastoris</i>			X-ray: LCP with monolein, and cholesterol 3.10 Å	Shimamura et al. (2011)
<i>Sphingosine 1-phosphate receptor</i>							
S1P1-T4L: replacement of the third cytoplasmic loop with T4L	Complexed with an antagonist sphingolipid mimic	<i>H. sapiens</i>	Insect cells infected by baculovirus		DDM + CHS	X-ray: LCP with monolein, and cholesterol 3.35 Å	Hanson et al. (2012)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
<i>M2 muscarinic acetylcholine receptor</i>							
M2-T4L: replacement of the third cytoplasmic loop with T4L, mutation of the N-linked glycosylation sites	Bound to an antagonist	<i>H. sapiens</i>	Insect cells infected by baculovirus	MNG	MNG	X-ray: LCP with monoolein, and cholesterol 3.00 Å	Haga et al. (2012)
<i>M3 muscarinic acetylcholine receptor</i>							
M3-T4 L: replacement of the third cytoplasmic loop with T4L	In complex with tiotropium (Spiriva), a potent muscarinic inverse agonist	<i>Rattus norvegicus</i>	Insect cells infected by baculovirus	MNG and inverse agonist tiotropium	MNG	X-ray: LCP with monoolein, and cholesterol 3.40 Å	Kruse et al. (2012)
<i>κ-opioid receptor</i>							
hKOR-T4 L: replacement of the third cytoplasmic loop with T4L	In complex with the antagonist JD1ic	<i>H. sapiens</i>	Insect cells infected by baculovirus			X-ray: LCP with monoolein, and cholesterol 2.90 Å	Wu et al. (2012)
<i>μ-opioid receptor</i>							
μOR-T4L: replacement of the third cytoplasmic loop with T4L, carboxy terminus truncated	Bound to a morphinan antagonist	<i>Mus musculus</i>	Insect cells infected by baculovirus	MNG + CHS	MNG + CHS	X-ray: LCP with monoolein, and cholesterol 2.80 Å	Manglik et al. (2012)
<i>δ-opioid receptor</i>							
δ-OR-T4L: replacement of the third cytoplasmic loop with T4L	In complex with the antagonist naltrindol	<i>M. musculus</i>	Insect cells infected by baculovirus	MNG + CHS + antagonist naltrindol	MNG + CHS	X-ray: LCP with monoolein, and cholesterol 3.40 Å	Granier et al. (2012)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
<i>Noiceptin/orphanin FQ (N/OFO) receptor</i> BRIL-NOP-T4L: replacement of the third cytoplasmic loop with T4L, of N-terminus with b562RIL (BRIL) and truncation of 31 C-terminal residues-terminal residues	In complex with the peptide mimetic antagonist C-24	<i>H. sapiens</i>	Insect cells infected by baculovirus	Replacement of N-terminus with b ₅₆₂ RIL (BRIL) and addition of antagonist C-24		X-ray: LCP with monoolein, and cholesterol 3.01 Å	Thompson et al. (2012)
<i>NTS1 neurotensin receptor</i> NTS1-GW5-T4L: replacement of the third cytoplasmic loop with T4L and 5 thermostabilizing point mutations	Bound to the C terminal portion (NT8-13) of the endogenous agonist neurotensin	<i>R. norvegicus</i>	Insect cells infected by baculovirus	Six stabilizing mutations: A86L, E166A, G215A, L310A, F358A, V360A, and addition of the agonist NT, MNG + CHS		X-ray: LCP with monoolein, and cholesterol 2.80 Å	White et al. (2012)
<i>Protease-activated receptor 1 (PAR1)</i> PAR1-T4L: replacement of the third cytoplasmic loop with T4L, mutation of the N-linked glycosylation sites in ECL2 and truncation of the N- and C-terminus	Bound with antagonist vorapaxar	<i>H. sapiens</i>	Insect cells infected by baculovirus	Antagonist vorapaxar	DDM, CHS and sodium cholate	X-ray: LCP with monoolein, and cholesterol 2.20 Å	Zhang et al. (2012)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
<i>5-HT1B serotonin receptor</i>							
SHT1B-BRIL: b562 RIL (BRIL) replaces cytoplasmic loop 3, truncation of the N-terminus to remove all glycosylation sites, and point mutation L138W	With bound ergotamine	<i>H. sapiens</i>	Insect cells infected by baculovirus	Thermostabilizing mutation L138W and addition of ergotamine	DDM + CHS	X-ray: LCP with monoolein, and cholesterol 2.70 Å	Wacker et al. (2013)
<i>5-HT2B serotonin receptor</i>							
5-HT2B-BRIL: b562 RIL (BRIL) replaces cytoplasmic loop 3, truncation of the N- and C-terminus, and introduction of M144W mutation	With bound ergotamine	<i>H. sapiens</i>	Insect cells infected by baculovirus	Thermostabilizing mutation M144W and addition of ergotamine	DDM + CHS	X-ray: LCP with monoolein, and cholesterol 2.70 Å	Wacker et al. (2013)
<i>Smoothed (SMO) receptor</i>							
BRIL-ΔCRD-SMO-ΔC: b562 RIL and truncation of N and the C terminus	With bound antagonist LY2940680	<i>H. sapiens</i>	Insect cells infected by baculovirus	Antagonist LY2940680	DDM + CHS	X-ray: LCP with monoolein, and cholesterol 2.45 Å	Wang et al. (2013)
<i>Channels: potassium and sodium ion-selective Two-pore domain potassium channel</i>							
K2P1.1 (TWIK-1)		<i>H. sapiens</i>	<i>P. pastoris</i>			3.40 Å	Miller and Long (2012)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
<i>Voltage-gated potassium channel</i>							
Kv1.2	Open in complex with an oxidoreductase beta subunit	<i>R. norvegicus</i>	<i>P. pastoris</i>			2.9 Å	Long et al. (2005) Chen et al. (2010)
<i>Kv1.2/Kv2.1 voltage-gated potassium channel chimera</i>							
WT and F233W sensor paddle transferred from Kv2.1 to Kv1.2, His10 tag at the amino terminus. Co-expressed with the rat β2.1 subunit	The pore is open and the voltage sensors adopt a membrane depolarized conformation	<i>R. norvegicus</i>	<i>P. pastoris</i>		Cymal-6, Cymal-7, CHAPS and lipids (POPC, POPE and POPG)	X-ray: vapor diffusion/hanging drops 2.4 Å and 2.9 Å	Long et al. (2007); Tao et al. (2010)
<i>Two-pore domain potassium channel K2P4.1 (TRAAK)</i>							
Codon-optimized, two consensus N-linked glycosylation sites mutated (N104Q, N108Q) and truncation of the C-terminal EGFP and 10xHistag in C-terminus		<i>R. norvegicus</i>	<i>P. pastoris</i>		DDM + Fos-choline 12	X-ray: vapor diffusion/hanging drops 3.80 Å	Brohawn et al. (2012, 2013)
<i>Inward-rectifier potassium channel</i>							
Kir2.2 Synthetic gene from 38 to 369 with GFP and 1D4 antibody recognition sequence in C-terminus	Without substrate In complex with PIP2	<i>Gallus gallus</i>	<i>P. pastoris</i>		DM	X-ray: vapor diffusion/hanging drops 3.1 Å and 3.31 Å	Tao et al. (2009); Hansen et al. (2011)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
<i>G protein-gated inward rectifying potassium channel GIRK2</i>							
Kir3.2 Deletion of N- and C-termini, EGFP and 10xHistag in C-terminus Wild-type and constitutively active mutant	Close and open conformation in complex with PIP2 In complex with β G-protein subunits: pre-open state intermediate between closed and open conformation	<i>M. musculus</i>	<i>P. pastoris</i>		DM	X-ray: vapor diffusion/hanging drops 3.60 Å 3.45 Å	Whorton and Mackinnon (2011, 2013)
<i>Other ion channels</i>							
<i>GluA2 Glutamate receptor (AMPA-subtype)</i>							
GluA2crist: Deletion of C- and N-terminus, mutation of 3 N-glycosylation sites, mutation of 4 residues from loop 1 by Ala, 2 point mutations (R586Q and C589A)	In complex with the competitive antagonist ZK 200775 With bound glutamate	<i>R. norvegicus</i>	Insect cells infected by baculovirus	Two point mutations (R586Q and C589A) to stabilize the tetrameric state and reduce nonspecific aggregation	C11Thio and synthetic lipids (POPC/POPE/POPG)	X-ray: vapor diffusion/hanging drops 3.60 Å	Sobolevsky et al. (2009)
<i>ASIC1 acid-sensing ion channel</i> Δ ASIC1: N- and C-terminus deletions	With bound chloride ion	<i>G. gallus</i>	Insect cells infected by baculovirus			X-ray: vapor diffusion/hanging drops 1.9 Å	Jasti et al. (2007)

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
ASIC1mf: amino terminal fusion with GFP and 8xHistag	Desensitized state, minimal functional channel	<i>G. gallus</i>	Insect cells infected by baculovirus		DDM	X-ray: vapor diffusion/hanging drops 3.0 Å	Gonzales et al. (2009)
PcTx1-ΔASIC1(26-463) with His10- tag in N-terminus	In complex with psalmotoxin 1 (PcTx1)	<i>G. gallus</i>	Insect cells infected by baculovirus		DDM	X-ray: vapor diffusion/hanging drops 2.99 Å 2.80 Å	Dawson et al. (2012); Bacongus and Gouaux (2012)
<i>Channels: aquaporins and glyceroporins</i>							
AQP4 aquaporin water channel							
AQP4 S180D mutant		<i>R. norvegicus</i>	Insect cells infected by baculovirus			Electron Diffraction 3.2 Å 2.80 Å	Hiroaki et al. (2006) Tani et al. (2009)
<i>M1 isoform full-length</i> AQP4 with N-terminal 8xHis followed by a flag tag		<i>H. sapiens</i>	<i>P. pastoris</i>		OG	X-ray: vapor diffusion/hanging drops 1.8 Å	Ho et al. (2009)
<i>HsAQP5: aquaporin water channel</i>							
HsAQP5		<i>H. sapiens</i>	<i>P. pastoris</i>		NG	X-ray: vapor diffusion/hanging drops 2.0 Å	Horsefield et al. (2008)
<i>Channels: Urea Transporter</i>							
UT-B: with 8-Histag in C-terminus		<i>B. taurus</i>	Insect cells infected by baculovirus		OG	X-ray: vapor diffusion/hanging drops 2.36 Å	Levin et al. (2012)
Channels: Amt/Rh proteins							

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
Rh C glycoprotein ammonia transporter		<i>H. sapiens</i>	HEK293 cells		OG	X-ray: vapor diffusion/hanging drops 2.10 Å	Gruswitz et al. (2010)
<i>Connexin gap junction</i> Connexin 26 Cx26; GJB2		<i>H. sapiens</i>	Insect cells infected by baculovirus			3.5 Å	Maeda et al. (2009)
<i>Intramembrane proteases</i> CAAX protease ZMP-STE24. E336A mutant	In complex with C-terminus tetrapeptide from prelamin A	<i>H. sapiens</i>	Insect cells infected by baculovirus			3.40 Å	Quigley et al. (2013)
<i>Antiporters</i> <i>UCP2 mitochondrial uncoupling protein 2</i> Residues 14–309, C-terminal His6 tag. ¹⁵ N-, ¹³ C-, ² H labeled protein	In complex with GDP	<i>M. musculus</i>	<i>E. coli</i>	GDP	DMPc cardio-lipiphy-tanoyl lipids and DPC		Berardi et al. (2011)
<i>ATP-binding cassette (ABC) transporters</i> <i>P-glycoprotein</i> Apo-Pgp with 3 N-linked glycosylation sites mutated and 6xHistag in C-terminus	Drug-bound: nucleotide-free inward-facing conformation competent to bind drugs	<i>M. musculus</i>	<i>P. pastoris</i>		DDM ± QZ59-RRR or QZ59-SSS	X ray: vapor diffusion/hanging drops 3.8 Å	Aller et al. (2009)
<i>ABCB10 mitochondrial ABC transporter</i>							

Table 1.1 (continued)

Recombinant protein	Conformation	Source	Expression system	Stabilization strategy	Detergent used	Structural method and resolution	Reference
Apo-ABC10: deletion of the N-terminal residues, with both N- and C-terminal His-tags	With bound ATP analogs. Open inward conformation	<i>H. sapiens</i>	Insect cells infected by baculovirus		DDM and CHS, X-ray: vapor diffusion/hanging drops 2.85 Å		Shintre et al. (2013)
<i>P-type ATPases</i>							
<i>Calcium ATPase from sarcoplasmic reticulum</i>							
SERCA1a-BAD: SERCA1a fused to a biotin acceptor domain	With bound Ca ²⁺ and AMPPCP	Rabbit	<i>Saccharomyces cerevisiae</i>		C12E8 and POPC, in presence of Ca ²⁺ and AMPPCP	X-ray: vapor diffusion/hanging drops 3.3 Å	Jidenko et al. (2005)
<i>Phospholamban</i>							
PLN in fusion with the maltose-binding protein (MBP) and ¹⁵ N, ¹³ C-labeled PLN	Unphosphorylated PLN pentamer in DPC micelles In T state	<i>H. sapiens</i>	<i>E. coli</i>		POPC/POPE lipid bilayer	Solution NMR and solid-state NMR	Oxenoid and Chou (2005); Verardi et al. (2011)

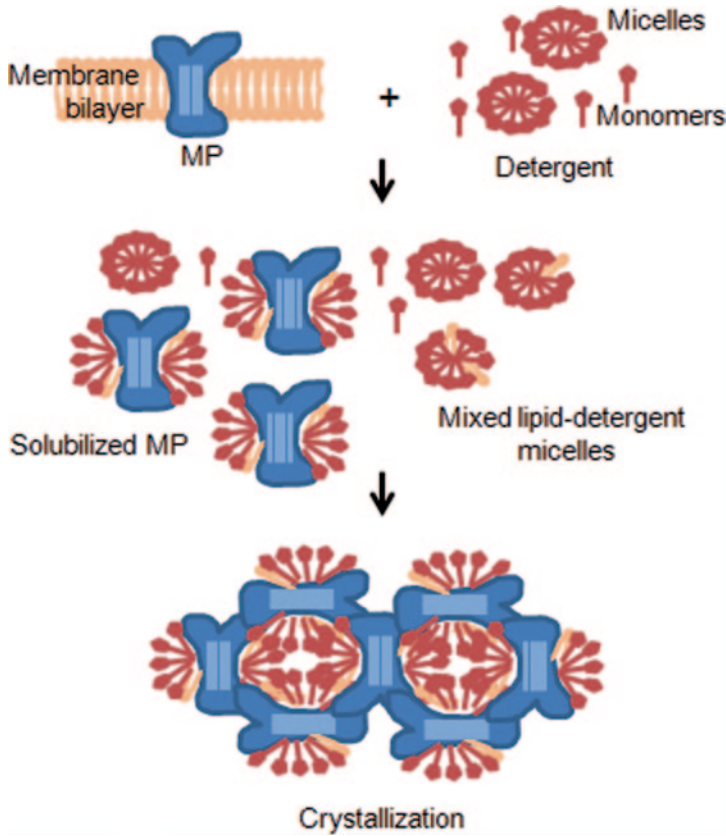


Fig. 1.1 Solubilization and crystallization of membrane proteins (MPs) in detergent solution

is a major obstacle to IMP structure determination. Scaffolds of the crystal lattice are found predominantly between the exposed polar surfaces of proteins, while the transmembrane parts remain buried from the detergent micelle (Fig. 1.1). Proteins with large extra-membranous domains are favored, and detergents that assemble into small micelles such as octyl- β -D-glucopyranoside (OG) or dimethyldodecylamineoxide (DDAO) are preferred in crystal trials (Prive 2007). In the past few years, structures of some recombinant “native” mammalian IMPs have been solved in OG, e.g., the aquaporin AQP4 (Ho et al. 2009), the urea transporter UTB (Levin et al. 2012), and the RhC glycoprotein ammonia transporter (Gruswitz et al. 2010); in nonyl- β -D-glucopyranoside (NG), e.g., aquaporin AQP5 (Horsefield et al. 2008); and in decylmaltoside (DM), e.g., potassium channels Kir2.2 (Tao et al. 2009) and Kir3.2 (Table 1.1; Whorton and Mackinnon 2011). Dodecyl- β -D-maltoside (DDM), a detergent that forms larger micelles, has also given rise to the structure of some recombinant “native” mammalian IMPs such as the acid-sensing ion channel 1 (Jasti et al. 2007), the potassium channel TRAAK (Brohawn et al. 2012, 2013), and the mitochondrial ABC transporter ABCB10 (Table 1.1; Shintre et al. 2013). However,

these detergents are not good for structure determination of “native” recombinant GPCRs. Indeed, the hydrophilic regions of these proteins, which are essential to form crystal contacts, are relatively small and often occluded by large detergent micelles such as those formed by DDM, and detergents forming smaller micelles are often denaturing for GPCRs. Moreover, the dynamic character of these proteins led to conformational heterogeneity that prevents the formation of well-ordered crystals. Thus, there are several options for promoting crystal formation of such IMPs: (1) increasing the conformational homogeneity of the protein by locking it in a single conformation, (2) increasing the hydrophilic area of the protein so that large micelle detergents can be used, (3) increasing the thermostability of the protein in detergent solution, and/or (4) finding new surfactants that are able to maintain the protein conformation stable in solution and are suitable for crystallization (Fig. 1.2).

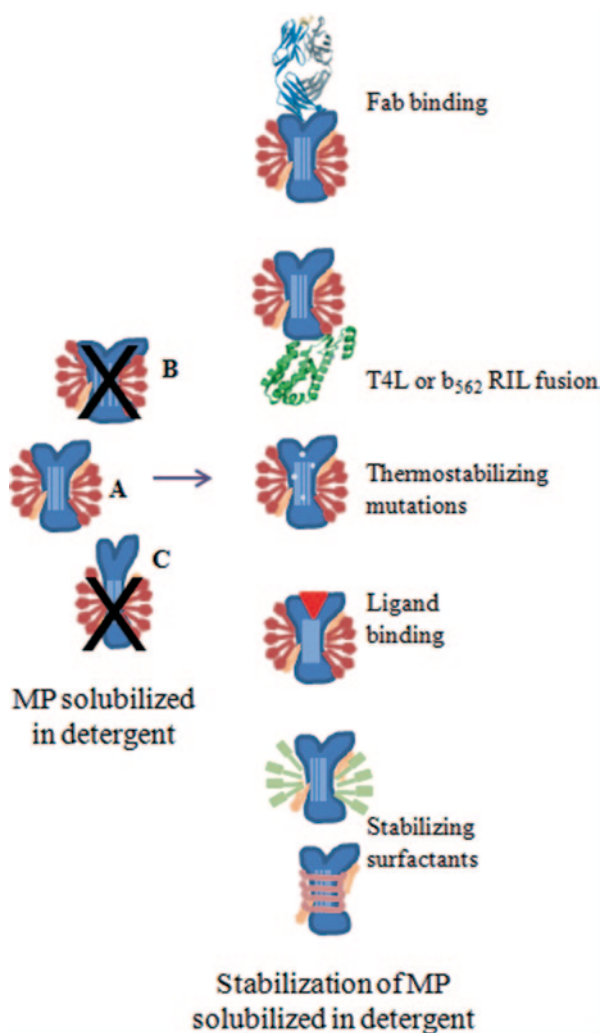
1.3.1 Increasing the Conformational Homogeneity of the Protein by Locking it in a Single Conformation

IMPs such as GPCRs, transporters, and channels are dynamic proteins that exist in several functionally distinct conformation states (active/inactive, open/closed, etc.). “Crystallogensis typically traps the most stable low energy states, making it difficult to obtain high-resolution structures of other less stable but biologically relevant functional states” (Steyaert and Kobilka 2011). Almost all recombinant GPCR structures have been solved thanks to the binding of a high-affinity ligand, agonist, inverse agonist, or antagonist (Table 1.1). The ligand probably locks the receptor in a single conformation and stabilizes it during crystallization. However, agonist alone is not sufficient to stabilize a fully active conformation, and the structures obtained are almost from inactive conformation (Steyaert and Kobilka 2011).

Locking of a single conformation can be obtained by applying mutagenesis. Structure of rhodopsin active conformation has been solved using constitutively active mutants (Deupi et al. 2012).

Co-crystallization with monoclonal antibody fragments (Fab) that bind to a single conformational state of the protein with high affinity and specificity locks it in a specific conformation and reduces the flexibility of the loop regions, can enhance the formation of well-diffracting crystals, and gives rise to the determination of the structure of different conformations of the protein (Fig. 1.2). Co-crystallization of β_2 adrenergic receptor and A_{2A} adenosine receptor with Fab fragments allowed the structure determination of an inactive conformation (Rasmussen et al. 2007; Hino et al. 2012), while active conformation of β_2 adrenergic receptor in complex with a heterotrimeric G protein was solved using co-crystallization with nanobodies (Rasmussen et al. 2011). These small fragments (13 kDa) corresponding to the variable domains of the llama antibody heavy chains are characterized by outstanding properties in terms of production, stability, and recognition of epitopes buried and inaccessible to conventional monoclonal antibodies (Steyaert and Kobilka 2011; Hamers-Casterman et al. 1993; Muyldermans et al. 2001).

Fig. 1.2 Strategies for the stabilization of membrane proteins solubilized in detergent. *Left:* solubilization destabilizes membrane proteins (MPs). Presence of nonfunctional conformations in solution which prevent protein crystallization. *Right:* different stabilization strategies to preserve functional conformation in solution



1.3.2 Increasing the Hydrophilic Area of the Protein

Co-crystallization with Fab or nanobodies not only favors a given conformation but also increases the hydrophilic surface of the protein-promoting crystal formation. The probability of obtaining crystals of IMPs can also be enhanced by the fusion of stable soluble protein domain increasing the hydrophilic area. In many IMPs, regions such as N- and C-terminus and loops are highly flexible and possess large stretches of polar residues unstructured and not suitable for crystal lattice contact formation. This is the case of GPCRs. The replacement of a flexible region by a stable soluble protein domain that increases the hydrophilic area of the GPCRs has had a lot of success in determining structures of GPCRs. The replacement of

the third intracellular loop, which is highly flexible in many GPCRs, by T4L has allowed the resolution of the structure of 11 of the 19 GPCRs published to date (Table 1.1, Fig. 1.2). T4L is a very stable and highly crystallizable soluble protein. Combined with the use of LCP as the crystallization matrix, this is by far the most effective strategy to solve GPCR structure. However, a number of GPCRs were not amenable to this approach due to deleterious effects on the expression or stability of the chimeric protein. Ray Stevens and collaborators reported recently the identification and development of novel GPCR fusion proteins to facilitate GPCR crystallization (Chun et al. 2012). Of the five new fusion proteins considered in their study, thermostabilized apocytochrome b_{562} RIL showed great utility in the crystallization of several GPCRs with superior characteristics relative to T4L that has been used previously (Chun et al. 2012). The replacement of the third intracellular loop of A_{2A} adenosine receptor with apocytochrome b_{562} RIL gave rise to the highest-resolution structure of A_{2A} adenosine receptor (Liu et al. 2012, 2012a), and the replacement of the N-terminal domain of smoothed receptor with apocytochrome b_{562} RIL allowed the resolution of the first structure of this receptor (Wang et al. 2013).

1.3.3 Increasing the Thermostability of the Protein in Detergent Solution

1.3.3.1 Using Additives to Detergent

Several successful GPCR structure determination efforts have shown that the addition of cholesterol analogs is often critical for maintaining GPCR stability. Thompson and coworkers (2011) have shown that sterols such as cholesteryl hemisuccinate (CHS), which induces the formation of a bicelle-like micelle architecture when mixed with DDM detergent, substantially increases the stability of the NOP receptor ORL-1, a member of the opioid GPCR family, in a mixed micelle environment.

1.3.3.2 Using Protein Engineering

A single point mutation can increase the stability of an IMP as shown for KcsA, M13 coat protein, diacylglycerol kinase (DAGK), and bacteriorhodopsin (Bowie 2001). Rhodopsin was significantly stabilized by engineering a new disulfide bond (Standfuss et al. 2007). Thermostabilizing mutations can be additive when combined and highly stable IMPs can be engineered (Zhou and Bowie 2000; Lau et al. 1999). These mutations favor a particular state of the receptor, which enhances the probability of obtaining high-quality diffracting crystals for structure determination (see Tate 2012 for a review). Tate and coworkers developed systematic scanning mutagenesis and used a ligand-binding assay to isolate mutants highly stable in detergent solution (Fig. 1.2). This approach allowed the determination of the structure of several GPCRs carrying combination of thermostabilizing point mutations like

adrenergic receptor $\beta 1$ (Warne et al. 2008, 2011, 2012; Moukhametziyanov et al. 2011), adenosine receptor A_{2A} (Lebon et al. 2011), and neurotensin receptor NTS1 (White et al. 2012; Shibata et al. 2013). This strategy allowed the structure determination of GPCRs bound to low-affinity agonists or ligands which, with native protein, would lead to incomplete occupancy of the receptor and conformational heterogeneity in crystallization trials preventing the formation of well-diffracting crystals. In the case of $\beta 1$ adrenergic receptor and A_{2A} adenosine receptor, the structures of thermostabilized mutants are identical to those of receptors in which the third intracellular loop was replaced by T4L, suggesting that thermostabilizing mutations predominantly affect receptor dynamics rather than the structure (Tate 2012).

For GPCRs, the use of several stabilizing strategies was necessary for the structure determination. The structures of smoothed and serotonin receptors were obtained using both apocytochrome b_{562} RIL and thermostabilizing point mutations (Wang et al. 2013; Wacker et al. 2013). The structures of chemokine receptor CXCR4 (Wu et al. 2010), dopamine D3 receptor (Chien et al. 2010), and neurotensin receptor (White et al. 2012) were obtained using T4L and thermostabilizing point mutations; and the NOP receptor structure was solved using both apocytochrome b_{562} RIL and T4L fusions (Thompson et al. 2012). In addition, often, it was necessary to truncate flexible N- and C-terminal regions of the protein.

Although these methods have been very successful and have resulted in an almost exponential growth in the number of mammalian IMPs and more particularly in GPCRs solved structures, the inherent limitations with such recombinant methods are that (1) truncations of a loop and/or of the N- and the C-terminal domains provide limited understanding of the structure and function of these regions, and (2) the locking of a conformational state using antibodies, replacement of a flexible loop by a stable soluble protein domain such as the T4L or the apocytochrome b_{562} -RIL, or the insertion of thermostabilizing mutations may affect the conformation and the structure of the protein. Thus, it is important to also develop new surfactants enhancing IMPs' stability and homogeneity, and the probability of obtaining well-diffracting crystals (Fig. 1.2).

1.3.4 New Surfactants Enhancing MP Stability and Homogeneity in Solution

To solubilize biological membranes, detergents need to be dissociating: they compete with lipid–lipid and lipid–protein interactions and disperse membrane components in the form of detergent-solubilized particles. The dissociating character of the detergent is frequently responsible for the destabilization and irreversible inactivation of IMPs after extraction (see Breyton et al. 2010 for a review). Several new classes of amphiphiles have been developed in the last decade in order to enhance IMPs' stability in solution and the probability of obtaining well-diffracting crystals such as the MNG from Samuel Gellman (Chae et al. 2010), the maltoside surfac-

tants from Wolfram Welte (Hovers et al. 2011), the fluorinated surfactants (FS) from Bernard Pucci (Breyton et al. 2004, Chap. 8 by Durand et al. in this volume), and the amphipathic polymers “amphipols” from Jean-Luc Popot (Popot et al. 2011, Chap 7 by Zoonens et al. in this volume). To date, of these different new classes of amphiphiles, only MNG has allowed obtaining crystals of sufficient quality for structure determination of mammalian IMPs.

1.3.4.1 Maltose-Neopentyl Glycol

This class of amphiphiles is built around a central quaternary carbon atom, which enables the incorporation of two hydrophilic and two lipophilic subunits, and is intended to place subtle restraints on conformational flexibility (Chae et al. 2010). In case of MNG, central quaternary carbon atom was derived from neopentyl glycol and hydrophilic groups were derived from maltose. The structures of four GPCRs were obtained using MNG. However, these four GPCRs were engineered by fusion with the soluble lysozyme domain T4L; M3 muscarinic acetylcholine receptor (Kruse et al. 2012), μ -opioid receptor (Manglik et al. 2012) and δ -opioid receptor (Granier et al. 2012), or with thermostabilizing mutations; NTS1 neurotensin receptor (White et al. 2012), and no structure of nonengineered IMP has been obtained using MNG yet. MNG is also suitable for NMR structural characterization of IMPs as shown by the recent study on the role of ligands on the equilibriums between functional states of the β 2 adrenergic receptor (Kim et al. 2013).

1.3.4.2 Maltoside Surfactants

Welte and co-workers hypothesized that conformationally restricted groups in the hydrophobic part lead to mild surfactants and increase protein stability in solution. They synthesized maltoside surfactants with rigid, saturated or aromatic hydrocarbon groups as hydrophobic parts by combinations of cyclohexyl rings and aromatic rings, and maltosyl as polar head group because of its stabilizing effects (Xie and Timasheff 1997) and its compactness (Hovers et al. 2011). The use of PCC- α -M propylcyclohexyl cyclohexyl- α -D-maltoside allowed the determination of the structure of the cytochrome b_6f complex from *Chlamydomonas reinhartii* (Hovers et al. 2011). Although, this surfactant was shown to efficiently stabilize β 1 adrenergic and smoothened receptors, no sufficiently well-diffracting crystal of these mammalian IMPs has been obtained using this surfactant yet.

1.3.4.3 Fluorinated Surfactants

These surfactants possess the same general structure as classical detergents, i.e., a hydrophilic head group and a hydrophobic tail, but the latter, rather than being a

hydrogenated aliphatic chain, is a fluorocarbon chain (Pavia et al. 1992, Chabaud et al. 1998). Several IMPs such as bacteriorhodopsin and cytochrome b_6/f have been shown to be more stable once transferred to FSs than in the presence of detergents (Breyton et al. 2004). A similar effect was shown on the human smoothened GPCR (Nehmé et al. 2010). In contrast to detergents, FSs are not able to extract IMPs from membranes; they do not interact with lipids and do not compete with protein–lipid interaction (Rodnim et al. 2008). This character, in combination with the relative stiffness of their chains compared to detergents, is favorable for the maintenance of MP integrity and function (Talbot et al. 2009). FSs were also described as possible chemical chaperones (Rodnim et al. 2008), and were used with success for the CF production of bacterial IMPs (Park et al. 2007, 2011) and also mammalian IMPs such as the mitochondrial uncoupled protein UCP1 (Blesneac et al. 2012). Recent advances in the use of these amphiphiles for IMPs structural analysis are reviewed in Chap. 8 by Durand et al. in this volume.

1.3.4.4 Amphipathic Polymers

Amphipathic polymers (APols) have been designed to form compact and stable complexes with IMPs (Breyton et al. 2010; Popot et al. 2011). They are small (9–10 kDa), with dense hydrophobic chain distribution that ensures high affinity for the protein's transmembrane surface, high solubility in water to keep MPs soluble up to tens of grams per liter, and high flexibility to adapt to the irregularities and small radius of curvature of protein's transmembrane regions. A8-35, by far the most extensively studied APol, is composed of a relatively short polyacrylate chain (~70 acrylate residues), some of the carboxylates of which have been grafted with octylamine (~17 of them) or isopropylamine (~28 units). The ~25 acid groups that have remained free are charged in aqueous solution, which makes the polymer highly water soluble, while the octylamide moieties render it highly amphipathic (Tribet et al. 1996; Popot et al. 2011). Compare to the IMPs solubilization with detergents, the IMP/APols complex is highly stable and allows the absence of any polymer molecule free. Thus, the IMP in complex with APols behaves almost like a conventional water-soluble protein, which makes it easy to handle for functional and structural studies (Fig. 1.2). During the last decade, APols permitted major breakthroughs regarding overexpression, purification, and stabilization of IMPs, opening very exciting perspectives for structural and dynamic investigations of these proteins (Popot et al. 2011). Neutral APols (NAPols) have been shown to maintain the native folding and the activity of some GPCRs in solution (Rahmeh et al. 2012), but also to permit native folding recovery by *in vitro* refolding through strategies based on the expression of the receptors in IBs (Banères and Mouillac 2012). Recent advances in the use of APols for IMPs structural analysis are reviewed in Chap. 7 (Zoonens et al.) in this volume.

1.4 Structural Biology of IMPs

As discussed earlier, IMPs are difficult to study due to their hydrophobic properties which generate a large number of critical steps during their expression and purification for structural analysis. Despite these difficulties, the elucidation of high-resolution structures of IMPs is increasing. This progress in structural biology of IMPs results from the conjunction of biotechnological advances for expression, sample preparation (as discussed earlier), and technical improvement of structural methods for structure determination (crystallography, NMR spectroscopy, electron microscopy).

Actually, most IMP structures have been resolved by X-ray crystallography, and it is still the method of choice to elucidate large IMPs. Nevertheless, the increasing sensitivity, the development of new pulse sequences, and the production of specific labeled proteins make the use of NMR spectroscopy possible not only for structure determination but also for the characterization of the intrinsic dynamics of the protein and ligand interaction.

1.4.1 Recent Progresses in Crystallogenesis and Crystallography of IMPs

Crystallization process of IMPs is still a challenging project. Advances in protein crystallization (high-throughput screening conditions, chemical synthesis of new detergents, and chemical additives) result in a better understanding of the crystallogenesis process. *In surfo* and *in meso* phase crystallizations of IMPs are the most commonly used methods for IMPs.

1.4.1.1 In Surfo Phase Crystallization

Micellar Systems

In surfo crystallization of IMPs is the most common and easy to use strategy for the first trials for the crystallization of IMPs. In the Brookhaven data bank, the majority of deposited structures have been determined using crystals grown from detergent-solubilized proteins by traditional vapor diffusion experiments (Chayen 2005). The recent advances in the synthesis of new detergents offer a large panel suitable for the purification and crystallization of IMPs (Kang et al. 2013). The choice of the detergent is crucial and depends on many different parameters including extraction and solubility efficiency, protein stability, and retention of function. Accounting the crystal structures of IMPs listed on the webpage <http://blanco.biomol.uci.edu/mpstruc/>, alkyl maltopyranosides or glucopyranosides are the most successful detergents followed by amine oxides and polyoxyethylene glycol (Parker and Newstead 2012). Actual screens dedicated to the crystallization of

IMPs are commercially available and are very useful for the first crystallization trials. The stability of the protein during the crystallization process often needs additional lipids. Lipids have become more and more important during the crystallization process and can play a role as molecular chaperone for IMPs. Cholesterol is necessary for the crystallization of GPCRs, and it was shown that cholesterol molecules formed the interface between the protein molecules in the physiological state (Cherezov et al. 2007). However, not all IMPs can be maintained in native conformation when solubilized with conventional detergents. Recently, new synthesis of detergents has been proposed to overcome the stability problem occurring during extraction of IMPs and to favor their crystallization by limiting the size of the micelles surrounding the protein. These new detergents, discussed before, are based on DDM. By adding two hydrophilic and two lipophilic subunits linked by a central quaternary carbon, MNG is a better detergent compared to DDM for extraction, stabilization, and crystallization of IMPs (Chae et al. 2010). Another detergent called facial or tandem facial amphiphile has been developed by (Zhang et al. 2007) to favor the stability of solubilized IMP by forming small micelles compared to DDM and are found to be more suitable for the crystallization and NMR spectrometry of IMPs.

Nonmicellar Systems

Once IMPs are extracted from their natural environment, they are vulnerable. Detergents are a relatively poor substitute for the bilayer and IMPs are often unstable outside the membrane. Slight perturbation of MP structure may lead to denaturation and aggregation after detergent extraction. Furthermore, obtaining crystals means screening various biochemical and/or biophysical conditions (pH, ionic strength, additives, temperature, etc.) which may alter the tiny equilibrium for the micelle/protein complex. One way to stabilize IMPs is to reconstitute them in mimicking phospholipid bilayer membranes like bicelles or nanodiscs.

The bicelles are lipid bilayers limited in size. The size is determined by the ratio between a long chain of phospholipids, generally phosphatidyl choline (like dimyristoyl phosphatidylcholine (DMPC) or ditridecanoyl phosphatidylcholine (DTPC), and short lipids like cholate or short phospholipids like dihexanoyl-phosphatidylcholine (DHPC) that form the rim. Bicelle is an attractive lipidic medium for mimicking phospholipid membranes. In the case of bacteriorhodopsin (Faham and Bowie 2002), clear density for a CHAPSO molecule inserted between protein subunits is seen within the layers, indicating that an important interaction between lipids and protein has been preserved within the bicelle by the cholesterol-like detergent. Thus, this method has the advantage to be a bilayer-based method, to preserve integrity of the protein, and the ability to diffuse in three dimensions to form a three-dimensional (3D) lattice. Moreover, the ability to grow crystals at room temperature (below the phase transition temperature) significantly expands the applicability of bicelle crystallization (Ujwal and Bowie 2011). Proteins in bicelle can be handled like proteins in detergent. Bicelle crystallization trials can be performed

like the standard detergent-based protocol including robotics and all commercially available screens.

Nanodiscs have been extensively used for solubilizing IMPs and present similar advantages as bicelles in terms of size and stability (Bayburt and Sligar 2010). The nanodisc is a noncovalent assembly of phospholipids (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine, POPC; dipalmitoylphosphatidylcholine, DPPC, and dimyristoylphosphatidylcholine, DMPC) and a genetically engineered “membrane scaffold protein” (MSP) based on the sequence of human serum apolipoprotein AI (Bayburt and Sligar 2010). The molecular ratio of phospholipids to MSP is crucial for right self-assembling particle formation. The phospholipid associates as a bilayer domain while two molecules of MSP wrap around the edges of the discoidal structure in a belt-like configuration, one MSP covering the hydrophobic alkyl chains of each leaflet. The MSPs were engineered into the synthetic gene optimized for expression in *E. coli* and include various affinity tags (6His, FLAG, Cys, etc.). The size of nanodiscs can be adjusted depending on the length of MSP (Hagn et al. 2013). Nanodiscs are self-formed from a mixture of detergent/phospholipid micelles and MSP upon removal of the detergent (Ritchie et al. 2009). Nanodiscs are then an ideal model membrane system with defined size and phospholipid composition. Despite their properties, the use of nanodiscs as mimicking membrane for crystallization is still quite limited and efforts must be accomplished to increase the packing contact necessary for crystallization. The small isotropic nanodiscs are more useful to study specific lipids/proteins and proteins/proteins interactions for a better understanding of MP function (El Moustaine et al. 2012).

1.4.1.2 *In Meso* Phase Crystallization: The LCP Method

Crystallization of IMPs *in meso* phase has emerged as the most powerful method, in particular for GPCRs structure determination. This method for crystallization of IMPs was first originally described by Landau and Rosenbusch (1996) using bacteriorhodopsin. Actually, the advances of a robotic system for the crystallization of IMPs *in meso* phase result in more than two thirds of GPCR’s crystallographic structures (Cherezov 2004). The *meso* phase is a bicontinuous lipidic phase formed spontaneously by mixing monoacylglycerols (MAGs) and water at a given ratio (Caffrey and Cherezov 2009a). The cubic phase can be doped by essential lipids like cholesterol or phospholipids for the stabilization of the protein of interest. LCP is composed of highly curved lipid bilayers and is connected by a water channel. The MP solubilized in detergent is added to the LCP, the lipids present at high concentration will replace the detergent molecules, and thus the protein will be reconstituted into the lipid bilayer. IMPs are able to diffuse freely in lipids and therefore make contact with each other for the nucleation and crystallization process. LCPs are viscous and not easy to handle for crystallization trials. But during the past few years, many tools and instruments have become available, making LCP for IMP crystallization a routine application (Caffrey and Cherezov 2009). Moreover,

the development of methods for measuring thermal stability and diffusion of MPs embedded in LCP seek to identify precisely the best condition for crystallization.

1.4.2 *Data Collection and Structure Determination from MP Crystals*

Data collection of IMP crystals is still very challenging. In the *in surfo* phase, the crystals usually have a high solvent content owing to the detergent micelle, which covers the hydrophobic part of the protein. Therefore, the crystals of IMPs are often fragile, difficult to handle, and suffer from anisotropic X-ray scattering and radiation sensitivity. In a LCP, the crystals generally contain less solvent, which makes them less radiation sensitive, but the crystals are often much smaller. In addition, crystal quality can vary considerably, even between crystals from the same drop. This means that the complete X-ray data of an IMP crystal require screening a large number of crystals at the synchrotron. The presence of automatic sample changers at synchrotron beamlines has helped to solve this problem, enabling many crystals to be screened quickly and efficiently (Blow 2008). Moreover, the microfocus beamlines at the synchrotron (Bowler et al. 2010) make data collection possible from the under-sized crystals or on focused patches of the best-ordered regions of larger crystals. Focused X-ray beams with low background scatter and beam sizes of less than 10 μm remarkably improve the resolution and the data statistics from small crystals. The disadvantage of these microfocused beamlines is to increase the radiation damages of crystals; however, this can be overcome by merging data from several crystals.

Phase determination of IMP crystals can also be very challenging. However, with the increasing number of solved structures, molecular replacement (MR) has nonetheless become the most successful method to obtain phase information (Bill et al. 2011). Actually, most of GPCR structures recently published have been solved using MR. Co-crystallization of the protein of interest with a protein domain of well-known structure (e.g., T4L fused in the intracellular loop of GPCRs) can also be used as a search model to obtain phase information. However, if the structural fold is not known, experimental phase determination is required. Classical crystal-phase determination methods used for soluble proteins can be applied such as selenomethionine labeled protein or heavy metal derivatives.

Serial femtosecond crystallography (SFX) using X-ray free-electron laser (XFEL) radiation is an emerging method for 3-D structure determination using crystals ranging from a few micrometers to a few hundred nanometers in size. This method relies on X-ray pulses that are sufficiently intense to produce high-quality diffraction but short enough to prevent substantial radiation damage (Chapman et al. 2011). X-ray pulses of only 70-fs duration terminate before any chemical damage processes have time to occur, leaving primarily ionization and X-ray-induced thermal motion as the main sources of radiation damage (Holton 2009). SFX therefore promises to break the correlation between sample size, damage, and resolution in structural biology.

1.4.3 Recent Progresses in NMR Spectroscopy of IMPs

According to the structure of IMPs deposited in PDB and by comparing the two lists of X-ray structures (<http://blanco.biomol.uci.edu/mpstruc/>) and NMR structures (<http://www.drorlist.com/nmr/MPNMR.html>), the larger part of IMP structures have been solved by X-ray crystallography; nevertheless, recent advances in NMR spectroscopy, both in liquid and solid states, have made comprehensive studies of larger IMPs more accessible. NMR structures of IMPs contribute to 10–15% of the overall structures (Nietlispach and Gautier 2011). This contribution is due to the recent advances that appeared not only in NMR spectroscopy but also in biotechnology for labeling and stabilizing IMPs in large amount for structural biology (as mentioned previously).

One advantage of NMR spectroscopy over the X-ray structure determination is the measurement of intrinsic dynamics of the protein in native-like conditions (Warschawski et al. 2011). NMR spectroscopy can be used with IMPs in several environments: in isotropic environments, in bicelles, or in anisotropic environments such as proteins embedded in phospholipid membranes in physiological conditions (Warschawski et al. 2011). According to the sample preparation, NMR in solution or ssNMR spectroscopy will be used. More than 90% of the IMPs structures determined by NMR have been elucidated in solution. But the recent progress in ssNMR offers great potentiality, as illustrated in studies of large IMPs like the phospholamban (Traaseth et al. 2009; Verardi et al. 2011), the influenza proton channel (Cady et al. 2009; Cady et al. 2010; Sharma et al. 2010), the *Yersinia enterocolitica* adhesin A (Shahid et al. 2012), and the structure of a GPCR (Park et al. 2012). However, structure determination by NMR spectroscopy of IMPs remains a challenging project. The major limitations of NMR structure determination still persist because of sample preparation, in particular preparation of large amount of isotope-labeled proteins, and the choice of the mimicking environment to keep the proteins stable during the long-time NMR experiments.

1.4.3.1 Strategy for Isotopic Labeling of IMPs

Most of IMP structures determined by NMR spectroscopy have been overexpressed in *E. coli*. This prokaryotic expression system is the most appropriate expression system for fully ^{15}N -, ^{13}C -, and ^2H -labeled IMPs. As discussed previously, a large variety of expression plasmids and strains are available for heterologous expression of IMPs for this purpose (Freigassner et al. 2009; Kainosho et al. 2006; Schlegel et al. 2012; Vaiphei et al. 2011). Yeast and mammalian cell expression systems are also being developed as alternate sources of isotope-labeled proteins (Egorova-Zachernyuk et al. 2011; Fan et al. 2011; Sarramegna et al. 2003). CF systems based on *E. coli*, wheat germ, or insect cell extracts (see Chap. 2 from Proverbio et al. in this volume) are an alternative for expression of uniformly ^{15}N , ^{13}C -labeled IMPs as

well as perdeuteration (Etezady-Esfarjani et al. 2007; Sobhanifar et al. 2010a). This strategy of expression overcomes the limitations occurred *in vivo* expression systems, and is very useful for the selective labeling of amino acids. CF system expression was recently used in combination with a sequence-optimized combinatorial dual-labeling approach to achieve rapid backbone assignment for a two- α -helical, a two-transmembrane, and a four-transmembrane histidine kinase receptors (Hefke et al. 2011; Maslennikov et al. 2010).

Additionally, it is noteworthy that fluorine labeling of IMPs could be very useful in NMR spectroscopy, for example, to define conformational state in IMPs. A covalent modification of cysteines with ^{19}F -labeled compound was used to describe conformational changes in β 2-adrenergic receptor (Liu et al. 2012a). Moreover, incorporation of unnatural fluorinated amino acid in the DAGK was used to evaluate the dynamics of this MP in *N*-dodecylphosphocholine (DPC) micelles and in natural *E. coli* membrane (Shi et al. 2012; Shi et al. 2011).

1.4.3.2 Membrane-Mimicking Environment for IMP Structure Determination by NMR

For Solution NMR Spectroscopy

One of the challenges for solution NMR spectroscopy of IMPs is the identification of conditions that can mimic the native lipid bilayer environment while maintaining the sample in a stable, folded state with a total complex size of ~ 100 kDa or less (Kim et al. 2009; Sanders and Sonnichsen 2006). Micelle-forming detergent is the most common way to solubilize IMPs for NMR structural studies. A large panel of detergent molecules is actually available. Screening condition is required for finding a suitable detergent and right biochemical conditions to ensure the solubilization of IMPs in native conformation. Detergents that have been used successfully for solution-state NMR spectroscopy of IMPs include *N,N*-dimethyldodecylamine *N*-oxide (LDAO), sodium dodecyl sulfate (SDS), CYFOS-7, β -OG, and DPC (see <http://www.drorlist.com/nmr/MPNMR.html>). For an appropriate balance between solubilization and stabilization, mix-micelle has been used. For example, NMR spectra of the KvAP voltage-dependent K^+ channel were found to be optimal in the mixture composed of a ratio of 2:1 DPC/LDAO, in which DPC alone yielded exchange-broadened NMR spectra, while LDAO yielded sharp spectra but short lifetimes (Shenkarev et al. 2010). In some cases, detergent micelle can be doped by addition of natural phospholipids as was required for UCP2 (Berardi et al. 2011).

Bicelles have emerged as a common medium for use in NMR studies of IMPs. Many recent structures have been solved by using isotropic bicelles composed of DMPC phospholipids and small acyl chain phosphatidyl choline DHPC (Bocharov et al. 2008; Bocharov et al. 2007; Mineev et al. 2010; Sharma et al. 2010; Shenkarev et al. 2013; Unnerst ale et al. 2011). Bicelles provide medium that allows both solid state and solution NMR. Playing with the q value, corresponding to the ratio be-

tween the concentration of DHPC (short acyl chain) and DMPC (long acyl chain), we can increase the size of bicelles which become no more isotropic and can then be used for either ssNMR or crystallography (De Angelis et al. 2006; De Angelis and Opella 2007; Rasmussen et al. 2007).

Other non-detergent molecules have been developed for solubilization of IMPs and are very useful for NMR spectroscopy (Dahmane et al. 2009; Popot et al. 2011). As mentioned before, APols are short amphipathic polymers that can substitute for detergents to keep IMPs water soluble and stabilized. IMPs solubilized with APols have been used in NMR spectroscopy to define the conformation of the leukotriene B4 bound to its receptor (Catoire et al. 2010).

Nanodiscs are new membrane mimetic media, which are closest artificial media to the natural phospholipidic membrane with size suitable for solution NMR studies (Ritchie et al. 2009). The major advantage of nanodiscs is the absence of any detergents, which are known to affect the stability of the IMP by interacting with the extra-membrane regions. Originally developed for the solubilization of functionally active IMPs, they have since been used for ssNMR (Kijac et al. 2007), and more recently, solution NMR applications (Raschle et al. 2009). In a recent study, design of novel nanodiscs with more limited size shows 30% reduced apparent correlation time compared to the classical one. Using this property, NMR structures of the IMP outer membrane protein-X (OMP-X) have been characterized and compared with those obtained in several micelle compositions (Hagn et al. 2013).

For ssNMR Spectroscopy

In ssNMR, the absence of a direct correlation between molecular size and sample line widths allows the use of bigger membrane media such as large bicelles or lipid bilayers, which provide a more native environment for functional proteins. The homogeneity of the sample leads to improved line widths and therefore spectral resolution. Recently, the ssNMR structure of the CXCR1 GPCR fully ^{15}N , ^{13}C -labeled was determined after reconstitution in phospholipid liposomes (Park et al. 2012). Biological membranes consist of highly complex lipid mixtures of varying compositions which can adaptively adjust to changes in the physical properties of the membrane.

1.4.3.3 Recent Developments in NMR

In Solution NMR

Transverse relaxation optimized spectroscopy (TROSY)-based experiments have expanded the applicability of 3-D structures of large protein complexes including MPs solubilized in mimicking-membrane environment (Konrat et al. 1999; Salzmann et al. 1999; Yang and Kay 1999). 3-D-TROSY version of the HNCA, HN(CO)CA, HN(CA)CB, HNCO, HN(CA)CO, etc., are used for the complete

resonance assignment. To avoid unfavorable relaxation properties of protons in large proteins, all the non-exchangeable carbon-bound protons are replaced by deuterium. Sparse proton density in perdeuterated amide HN back-exchange proteins leads to few long-range nOes. To overcome this problem, a powerful method to reintroduce a perdeuterated sample is to biosynthetically incorporate protonated methyl groups of leucine, valine, and isoleucine by growing samples in minimal media with selectively labeled α -ketoisovalerate and α -ketobutyrate (Goto et al. 1999). This method was successfully applied for the structural determination of numerous IMPs from the β -barrel family at first with OMP-X in DHPC (Hilty et al. 2002), Kp OMP-A (Renault et al. 2009), and VDAC (Hiller et al. 2008). Other methyl protons can be targeted, such as those of alanine or methionine, by adding protonated amino acid into the culture media. This specific labeling approach is very useful for NMR structure determination of large MPs but also provides large application to study the intrinsic dynamics of very large proteins like the 670-kDa 20S core-particle proteasome (Religa et al. 2010; Ruschak et al. 2010), the DDM-solubilized KcsA channel (Imai et al. 2010), and β 2 adrenergic receptor (Bokoch et al. 2010).

Structure determinations by NMR spectroscopy are issued from the chemical shifts and nOes observed in 2-D, 3-D, or 4-D spectra. Distance restraints obtained by measuring nOes are often not enough to define precisely a structure of protein. Long-range restraints are necessary to obtain a more accurate NMR structure. Residual dipolar coupling (RDC) and paramagnetic relaxation enhancement (PRES) have been developed and nicely detailed in reviews (Kim et al. 2009; Qureshi and Goto 2012) and in Chap. 12 by Catoire et al. in this volume.

In ssNMR

ssNMR spectroscopy becomes a more and more attractive method for structure determination of IMPs in their native environment. The recent advances described (see for reviews Baldus 2006 and Renault et al. 2010) offer a large panel of recent developments and applications for structural and dynamic information of the large MP complex in phospholipid membrane. Moreover, technical improvements for ssNMR spectroscopy: design of the new probe heads to give access to ultrahigh-speed ssNMR and dynamic nuclear polarization (DNP) will promote ssNMR for structural biological studies of IMPs in native membranes.

References

- Andréll J, Tate CG (2013) Overexpression of membrane proteins in mammalian cells for structural studies. *Mol Membr Biol* 30:52–63
- Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, Harrell PM, Trinh YT, Zhang Q, Urbatsch IL, et al (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 323:1718–1722

- Arechaga I, Miroux B, Karrasch S, Huijbregts R, de Kruijff B, Runswick MJ, and Walker, JE (2000) Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase. *FEBS Lett* 482:215–219
- Baldus M (2006) Molecular interactions investigated by multi-dimensional solid-state NMR. *Curr Opin Struct Biol* 16:618–623
- Baconguis I, Gouaux E (2012) Structural plasticity and dynamic selectivity of acid-sensing ion channel-spider toxin complexes. *Nature* 489:400–405
- Banères JL, Mouillac B (2012) Handling G-protein-coupled receptors: expression, purification and in vitro stabilization. *Med Sci (Paris)* 28:837–844
- Bayburt TH, Sligar SG (2010) Membrane protein assembly into Nanodiscs. *FEBS Lett* 584:1721–1727
- Bazzacco P, Billon-Denis E, Sharma KS, Catoire LJ, Mary S, Le Bon C, Point E, Banères JL, Durand G, Zito F, Pucci B, Popot JL (2012) Nonionic homopolymeric amphipols: application to membrane protein folding, cell-free synthesis, and solution nuclear magnetic resonance. *Biochemistry* 51:1416–1430
- Berardi MJ, Shih WM, Harrison SC, Chou JJ (2011) Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature* 476:109–113
- Bill RM, Henderson PJ, Iwata S, Kunji ER, Michel H, Neutze R et al (2011) Overcoming barriers to membrane protein structure determination. *Nat Biotechnol* 29:335–340
- Blesneac I, Ravaud S, Juillan-Binard C, Barret LA, Zoonens M, Polidori A, Miroux B, Pucci B, Pebay-Peyroula E (2012) Production of UCP1 a membrane protein from the inner mitochondrial membrane using the cell free expression system in the presence of a fluorinated surfactant. *Biochim Biophys Acta* 1818:798–805
- Blow N (2008) Structural genomics: inside a protein structure initiative center. *Nat Meth* 5:203–207
- Bocharov EV, Pustovalova YE, Pavlov KV, Volynsky PE, Goncharuk MV, Ermolyuk YS, Karpunin DV, Schulga AA, Kirpichnikov MP, Efremov RG et al (2007) Unique dimeric structure of BNip3 transmembrane domain suggests membrane permeabilization as a cell death trigger. *J Biol Chem* 282:16256–16266
- Bocharov EV, Mineev KS, Volynsky PE, Ermolyuk YS, Tkach EN, Sobol AG, Chupin VV, Kirpichnikov MP, Efremov RG, Arseniev AS (2008) Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *J Biol Chem* 283:6950–6956
- Bokoch MP, Zou YZ, Rasmussen SGF, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi HJ, Thian FS, Kobilka TS et al (2010) Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature* 463:108–121
- Bonander N, Bill RM (2012) Optimising yeast as a host for recombinant protein production. *Methods Mol Biol* 866:1–9
- Bonander N, Hedfalk K, Larsson C, Mostad P, Chang C, Gustafsson L, Bill RM (2005) Design of improved membrane protein production experiments: quantitation of the host response. *Protein Sci* 14:1729–1740
- Bowie JU (2001) Stabilizing membrane proteins. *Curr Opin Struct Biol* 11:397–402
- Bowler MW, Guijarro M, Pettidmange S, Baker I, Svensson O, Burghammer M, Mueller-Dieckmann C, Gordon EJ, Flot D, McSweeney SM et al (2010) Diffraction cartography: applying microbeams to macromolecular crystallography sample evaluation and data collection. *Acta Crystallogr Sect D Biol Crystallogr* 66:855–864
- Breyton C, Chabaud E, Chaudier Y, Pucci B, Popot JL (2004) Hemifluorinated surfactants: a non-dissociating environment for handling membrane proteins in aqueous solutions? *FEBS Lett* 564:312–318
- Breyton C, Pucci B, Popot JL (2010) Amphipols and fluorinated surfactants: two alternatives to detergents for studying membrane proteins in vitro. *Methods Mol Biol* 601:219–245
- Brohawn SG, del Marmol J, MacKinnon R (2012) Crystal Structure of the Human K2P TRAAK, a Lipid- and Mechano-Sensitive K⁺ Ion Channel. *Science* 335:436–441

- Brohawn SG, Campbell EB, MacKinnon R (2013) Domain-swapped chain connectivity and gated membrane access in a Fab-mediated crystal of the human TRAAK K⁺ channel. *Proceedings of the National Academy of Sciences of the United States of America* 110:2129–2134
- Cady SD, Mishanina TV, Hong M (2009) Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding. *J Mol Biol* 385:1127–1141
- Cady SD, Schmidt-Rohr K, Wang J, Soto CS, DeGrado WF, Hong M (2010) Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* 463:689–692
- Caffrey M, Cherezov V (2009) Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc* 4:706–731
- Catoire LJ, Damian M, Giusti F, Martin A, van Heijenoort C, Popot JL, Guittet E, Baneres JL (2010) Structure of a GPCR ligand in its receptor-bound state: leukotriene B₄ adopts a highly constrained conformation when associated to human BLT₂. *J Am Chem Soc* 132:9049–9057
- Chabaud E, Barthelemy P, Mora N, Popot JL, Pucci B (1998) Stabilization of integral membrane proteins in aqueous solution using fluorinated surfactants. *Biochimie* 80:515–530
- Chae PS, Rasmussen SGF, Rana RR, Gotfryd K, Chandra R, Goren MA, Kruse AC, Nurva S, Lolland CJ, Pierre Y et al (2010) Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* 7:1003–1090
- Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U et al (2011) Femtosecond X-ray protein nanocrystallography. *Nature* 470:73–81
- Chayen NE (2005) Methods for separating nucleation and growth in protein crystallisation. *Prog Biophys Mol Biol* 88:329–337
- Chen R (2012) Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnol Adv* 30:1102–1107
- Cherezov V, Peddi A, Muthusubramaniam L, Zheng YF, Caffrey M (2004) A robotic system for crystallizing membrane and soluble proteins in lipidic mesophases. *Acta Crystallogr D Biol Crystallogr* 60:1795–1807
- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK et al (2007) High-resolution crystal structure of an engineered human beta(2)-adrenergic G protein-coupled receptor. *Science* 318:1258–1265
- Chien EY, Liu W, Zhao Q, Katritch V, Han GW, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, Stevens RC (2010) Structure of the human dopamine D₃ receptor in complex with a D₂/D₃ selective antagonist. *Science* 330:1091–1095
- Chun E, Thompson AA, Liu W, Roth CB, Griffith MT, Katritch V, Kunken J, Xu F, Cherezov V, Hanson MA, Stevens RC (2012) Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* 20:967–976
- Condreay JP, Kost TA (2007) Baculovirus expression vectors for insect and mammalian cells. *Curr Drug Targets* 8:1126–1131
- Dahmane T, Damian M, Mary S et al (2009) Amphipol-assisted in vitro folding of G protein-coupled receptors. *Biochemistry* 48:6516–6521
- Dahmane T, Giusti F, Catoire LJ, Popot JL (2011) Sulfonated amphipols: synthesis, properties, and applications. *Biopolymers* 95:811–823
- Dawson RJ, Benz J, Stohler P, Tetaz T, Joseph C, Huber S, Schmid G, Hugin D, Pflimlin P, Trube G et al (2012) Structure of the acid-sensing ion channel 1 in complex with the gating modifier Psalmotoxin 1. *Nat Commun* 3
- De Angelis AA, Opella SJ (2007) Bicelle samples for solid-state NMR of membrane proteins. *Nat Protoc* 2:2332–2338
- De Angelis AA, Howell SC, Opella SJ (2006) Assigning solid-state NMR spectra of aligned proteins using isotropic chemical shifts. *J Magn Reson* 183:329–332
- Deniaud A, Liguori L, Blesneac I, Lenormand JL, Pebay-Peyroula E (2010) Crystallization of the membrane protein hVDAC1 produced in cell-free system. *Biochim Biophys Acta: Biomembr* 1798:1540–1546

- Deniaud A, Bernaudat F, Frelet-Barrand A, Juillan-Binard C, Vernet T et al. (2011) Expression of a chloroplast ATP/ADP transporter in *E. coli* membranes: behind the Mistic strategy. *Biochim Biophys Acta* 1808: 2059–2066
- Deupi X, Edwards P, Singhal A, Nickle B, Oprian D, Schertler G et al (2012) Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proc Natl Acad Sci U S A* 109:119–124
- Dore AS, Robertson N, Errey JC, Ng I, Hollenstein K, Tehan B, Hurrell E, Bennett K, Congreve M, Magnani F et al (2011) Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthenes XAC and caffeine. *Structure* 19:1283–1293
- Egorova-Zachernyuk TA, Bosman G, DeGrip WJ (2011) Uniform stable-isotope labeling in mammalian cells: formulation of a cost-effective culture medium. *Appl Microbiol Biotechnol* 89:397–406
- El Moustaine D, Granier S, Doumazane E, Scholler P, Rahmeh R, Bron P, Mouillac B, Baneres JL, Rondard P, Pin JP (2012) Distinct roles of metabotropic glutamate receptor dimerization in agonist activation and G-protein coupling. *Proc Natl Acad Sci U S A* 109: 16342–16347
- Etezady-Esfarjani T, Hiller S, Villalba C, Wuthrich K (2007) Cell-free protein synthesis of perdeuterated proteins for NMR studies. *J Biomol NMR* 39:229–238
- Faham S, Bowie JU (2002) Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J Mol Biol* 316:1–6
- Fan Y, Shi LC, Ladizhansky V, Brown LS (2011) Uniform isotope labeling of a eukaryotic seven-transmembrane helical protein in yeast enables high-resolution solid-state NMR studies in the lipid environment. *J Biomol NMR* 49:151–161
- Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol* 63:1256–1272
- Freigassner M, Pichler H, Glieder A (2009) Tuning microbial hosts for membrane protein production. *Microb Cell Fact* 8:69
- Frelet-Barrand A, Boutigny S, Kunji ERS, Rolland N (2010a) Membrane protein expression in *Lactococcus lactis*. *Methods Mol Biol* 601:67–85
- Frelet-Barrand A, Boutigny S, Moyet L, Deniaud A, Seigneurin-Berny D et al (2010b) *Lactococcus lactis*, an alternative system for functional expression of peripheral and intrinsic *Arabidopsis* membrane proteins. *PLoS ONE* 7(3):e32325
- Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ 1) methyl-protonated N-15-, C-13-, H-2-labeled proteins. *J Biomol NMR* 13:369–374
- Gonzales EB, Kawate T, Gouaux E (2009) Pore architecture and ion sites in acid-sensing ion channels and P2X receptors. *Nature* 460:599–604
- Granier S, Manglik A, Kruse AC, Kobilka TS, Thian FS, Weis WI, Kobilka BK (2012) Structure of the δ -opioid receptor bound to naltrindole. *Nature* 485:400–404
- Gruswitz F, Chaudhary S, Ho JD, Schlessinger A, Pezeshki B, Ho CM et al (2010) Function of human Rh based on structure of RhCG at 2.1 Å. *Proc Natl Acad Sci U S A* 107:9638–9643
- Haga K, Kruse AC, Asada H, Yurugi-Kobayashi T, Shiroishi M, Zhang C, Weis WI, Okada T, Kobilka BK, Haga T, Kobayashi T (2012) Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* 482:547–551
- Hagn F, Eitzkorn M, Raschle T, Wagner G (2013) Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *J Am Chem Soc* 135:1919–1925
- Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R (1993) Naturally occurring antibodies devoid of light chains. *Nature* 3636428:446–448
- Hansen SB, Tao X, MacKinnon R (2011) Structural basis of PIP2 activation of the classical inward rectifier K⁺ channel Kir2.2. *Nature* 477:495–498

- Hanson MA, Roth CB, Jo E, Griffith MT, Scott FL, Reinhart G, Desale H, Clemons B, Cahalan SM, Schuerer SC, Sanna MG, Han GW, Kuhn P, Rosen H, Stevens RC (2012) Crystal structure of a lipid G protein-coupled receptor. *Science* 335:851–855
- Hefke F, Bagaria A, Reckel S, Ullrich S, Dötsch V, Glaubitz C, Güntert P (2011) Optimization of amino acid type-specific ¹³C and ¹⁵N labeling for the backbone assignment of membrane proteins by solution- and solid-state NMR with the UPLABEL algorithm. *J Biomol NMR* 49:75–84
- Hiller S, Garces RG, Malia TJ, Orekhov VY, Colombini M, Wagner G (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science* 321:1206–1210
- Hilty C, Fernandez C, Wider G, Wuthrich K (2002) Side chain NMR assignments in the membrane protein OmpX reconstituted in DHPC micelles. *J Biomol NMR* 23:289–301
- Hino T, Arakawa T, Iwanari H, Yurugi-Kobayashi T, Ikeda-Suno C, Nakada-Nakura Y, Kusano-Arai O, Weyand S, Shimamura T, Nomura N, Cameron AD, Kobayashi T, Hamakubo T, Iwata S, Murata T (2012) G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. *Nature* 482:237–240
- Hiroaki Y, Tani K, Kamegawa A, Gyobu N, Nishikawa K, Suzuki H, Walz T, Sasaki S, Mitsuoka K, Kimura K, et al (2006) Implications of the aquaporin-4 structure on array formation and cell adhesion. *J Mol Biol* 355:628–639
- Holton JM (2009) A beginner's guide to radiation damage. *J Synchrotron Radiat* 16:133–142
- Horsefield R, Norden K, Fellert M, Backmark A, Tornroth-Horsefield S, van Scheltinga ACT, Kvassman J, Kjellbom P, Johanson U, Neutze R (2008) High-resolution x-ray structure of human aquaporin 5. *Proceedings of the National Academy of Sciences of the United States of America* 105:13327–13332
- Hovers J, Potschies M, Polidori A, Pucci B, Raynal S, Bonneté F, Serrano-Vega MJ, Tate CG, Picot D, Pierre Y, Popot JL, Nehmé R, Bidet M, Mus-Veteau I, Busskamp H, Jung KH, Marx A, Timmins PA, Welte W (2011) A class of mild surfactants that keep integral membrane proteins water-soluble for functional studies and crystallization. *Mol Membr Biol* 28:171–181
- Imai S, Osawa M, Takeuchi K, Shimada I (2010) Structural basis underlying the dual gate properties of KcsA. *Proc Natl Acad Sci U S A* 107: 6216–6221
- Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens, RC (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* 322:1211–1217
- Jarvis DL, Finn EE (1995) Biochemical analysis of the N-glycosylation pathway in baculovirus-infected lepidopteran insect cells. *Virology* 212:500–511
- Jasti J, Furukawa H, Gonzales EB, Gouaux E (2007) Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH. *Nature* 449:316–+
- Jidenko M, Nielsen RC, Sorensen TL et al (2005) Crystallization of a mammalian membrane protein overexpressed in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 102:11687–11691
- Junge F, Schneider B, Reckel S, Schwarz D, Dötsch V et al (2008) Large-scale production of functional membrane proteins. *Cell Mol Life Sci* 65:1729–1755
- Junge F, Luh LM, Proverbio D, Schäfer B, Abele R, Beyermann M, Dötsch V, Bernhard F (2010) Modulation of G-protein coupled receptor sample quality by modified cell-free expression protocols: a case study of the human endothelin A receptor. *J Struct Biol* 172 :94–106
- Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Mei Ono A, Güntert P (2006) Optimal isotope labelling for NMR protein structure determinations. *Nature* 440:52–57
- Kang HJ, Lee C, Drew D (2013) Breaking the barriers in membrane protein crystallography. *Int J Biochem Cell Biol* 45:636–644
- Kijac AZ, Li Y, Sligar SG, Rienstra CM (2007) Magic-angle spinning solid-state NMR spectroscopy of nanodisc-embedded human CYP3A4. *Biochemistry* 46:13696–13703
- Kim HJ, Howell SC, Van Horn WD, Jeon YH, Sanders CR (2009) Recent advances in the application of solution NMR spectroscopy to multi-span integral membrane proteins. *Prog Nucl Magn Reson Spectrosc* 55:335–360

- Kim TH, Chung KY, Manglik A, Hansen AL, Dror RO, Mildorf TJ, Shaw DE, Kobilka BK, Prosser RS (2013) The role of ligands on the equilibria between functional states of a G protein-coupled receptor. *J Am Chem Soc* 135:9465–9474
- Konrat R, Yang DW, Kay LE (1999) A 4D TROSY-based pulse scheme for correlating (¹H)-¹⁵N, ¹³C-¹³(α), C-¹³(β) chemical shifts in high molecular weight, ¹⁵N, ¹³C, ¹H-2 labeled proteins. *J Biomol NMR* 15:309–313
- Kruse AC, Hu J, Pan AC, Arlow DH, Rosenbaum DM, Rosemond E, Green HF, Liu T, Chae PS, Dror RO, Shaw DE, Weis WI, Wess J, Kobilka BK (2012) Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* 482:552–556
- Landau EM, Rosenbusch JP (1996) Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci U S A* 93: 14532–14535
- Lau FW, Nauli S, Zhou Y, Bowie JU (1999) Changing single side-chains can greatly enhance the resistance of a membrane protein to irreversible inactivation. *J Mol Biol* 290:559–564
- Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AG, Tate CG (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* 474:521–525
- Levin EJ, Cao Y, Enkavi G, Quick M, Pan Y, Tajkhorshid E, Zhou M (2012) Structure and permeation mechanism of a mammalian urea transporter. *Proc Natl Acad Sci USA* 109:11194–11199
- Liu J, Rost B (2001) Comparing function and structure between entire proteomes. *Protein Sci* 10:1970–1979
- Liu W, Chun E, Thompson AA, Chubukov P, Xu F, Katritch V, Han GW, Roth CB, Heitman LH, IJzerman AP, Cherezov V, Stevens RC (2012a) Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* 337:232–236
- Liu JJ, Horst R, Katritch V, Stevens RC, Wuthrich K (2012b) Biased signaling pathways in beta(2)-adrenergic receptor characterized by F-19-NMR. *Science* 335:1106–1110
- Nietlispach D, Gautier A (2011) Solution NMR studies of polytopic alpha-helical membrane proteins. *Current Opinion in Structural Biology* 21:497–508
- Long SB, Campbell EB, Mackinnon R (2005) Crystal structure of a mammalian voltage-dependent Shaker family K⁺ channel. *Science* 309:897–903
- Long SB, Tao X, Campbell EB, MacKinnon R (2007) Atomic structure of a voltage-dependent K⁺ channel in a lipid membrane-like environment. *Nature* 450:376–382
- Lundstrom K (2006) Structural genomics for membrane proteins. *Cell Mol Life Sci* 63:2597–2607
- Ma Y, Muench D, Schneider T, Sahl HG, Bouhss A, Ghoshdastider U, Wang J, Doetsch V, Wang X, Bernhard F (2011) Preparative scale cell-free production and quality optimization of Mray homologues in different expression modes. *J Biol Chem* 286:38844–38853
- Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, Granier S (2012) Crystal structure of the μ -opioid receptor bound to a morphinan antagonist. *Nature* 485:321–326
- Maslennikov I, Klammt C, Hwang E, Kefala G, Okamura M, Esquivies L, Mors K, Glaubitz C, Kwiatkowski W, Jeon YH et al. (2010) Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis. *Proc Natl Acad Sci U S A* 107: 10902–10907
- Maslennikov I, Choe S (2013) Advances in NMR structures of integral membrane proteins. *Current Opinion in Structural Biology* 23:555–562
- Matthies D, Haberstock S, Joos F, Dötsch V, Vonck J, Bernhard F, Meier T (2011) Cell-free expression and assembly of ATP synthase. *J Mol Biol* 413 :593–603
- Miller AN, Long SB (2012) Crystal structure of the human two-pore domain potassium channel K2P1. *Science* 335:432–436
- Mineev KS, Bocharov EV, Pustovalova YE, Bocharova OV, Chupin VV, Arseniev AS (2010) Spatial structure of the transmembrane domain heterodimer of ErbB1 and ErbB2 receptor tyrosine kinases. *J Mol Biol* 400:231–243
- Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260:289–298

- Moukhametdzianov R, Burghammer M, Edwards PC, Petitdemange S, Popov D, Fransen M, McMullan G, Schertler GF, Riekel C (2008) Protein crystallography with a micrometre-sized synchrotron-radiation beam. *Acta Crystallogr D Biol Crystallogr* 64:158–166
- Moukhametdzianov R, Warne T, Edwards PC, Serrano-Vega MJ, Leslie AG, Tate CG, Schertler GF (2011) Two distinct conformations of helix 6 observed in antagonist-bound structures of a beta1-adrenergic receptor. *Proc Natl Acad Sci USA* 108:8228–8232
- Mus-Veteau I (2010) Heterologous expression of membrane proteins for structural analysis. *Methods Mol Biol* 601: 1–16. (In Mus-Veteau (Ed.) *Methods in molecular biology, heterologous expression of membrane proteins: methods and protocols*, Humana Press)
- Muyldermans S (2001) Single domain camel antibodies: current status. *J Biotechnol* 74:277–302
- Nehme R, Joubert O, Bidet M, Lacombe B, Polidori A, Pucci B, Mus-Veteau I (2010) Stability study of the human G-protein coupled receptor, Smoothened. *Biochim Biophys Acta* 1798: 1100–1110
- Oxenoid K, Chou JJ (2005) The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc Natl Acad Sci USA* 102:10870–10875
- Nietlispach D, Gautier A (2011) Solution NMR studies of polytopic alpha-helical membrane proteins. *Current Opinion in Structural Biology* 21:497–508
- Palczewski K, Kumasaka T, Hori T et al (2000) Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* 289:739–745
- Park KH, Berrier C, Lebaupain F, Pucci B, Popot JL, Ghazi A, Zito F (2007) Fluorinated and hemifluorinated surfactants as alternatives to detergents for membrane protein cell-free synthesis. *Biochem J* 403:183–187
- Park KH, Billon-Denis E, Dahmane T, Lebaupain F, Pucci B, Breyton C, Zito F (2011) In the cauldron of cell-free synthesis of membrane proteins: playing with new surfactants. *Nat Biotechnol* 28:255–261
- Park SH, Das BB, Casagrande F, Tian Y, Nothnagel HJ, Chu M, Kiefer H, Maier K, De Angelis AA, Marassi FM et al (2012) Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature* 491:779–783
- Parker JL, Newstead S (2012) Current trends in α -helical membrane protein crystallization: an update. *Protein Sci* 21:1358–1365
- Pavia AA, Pucci B, Riess JG, Zarif L (1992) New perfluoro alkylated telomeric non ionic surfactants synthesis physicochemical and biological properties. *Makromol Chem* 193:2505–2517
- Periasamy A, Shadiac N, Amalraj A, Garajova S, Nagarajan Y, Waters S, Mertens HDT, Hrmova M (2012) Cell-free synthesis of membrane (1,3)- β -d-glucan (curdlan) synthase: co-translational insertion in liposomes and reconstitution in nanodiscs. *Biochim Biophys Acta: Biomembr* 1828:743–757
- Popot JL, Althoff T, Bagnard D, Baneres JL, Bazzacco P, Billon-Denis E, Catoire LJ, Champeil P, Charvolin D, Cocco MJ et al. (2011) Amphipols From A to Z. *Annu Rev Biophys* 40:379–408 (Rees DC, Dill KA, Williamson JR, eds.)
- Prive GG (2007) Detergents for the stabilization and crystallization of membrane proteins. *Methods* 41:388–397
- Quigley A, Dong YY, Pike AC, Dong L, Shrestha L, Berridge G, Stansfeld PJ, Sansom MS, Edwards AM, Bountra C et al (2013) The structural basis of ZMPSTE24-dependent laminopathies. *Science* 339:1604–1607
- Qureshi T, Goto NK (2012) Contemporary methods in structure determination of membrane proteins by solution NMR. In Zhu G (ed) *NMR of proteins and small biomolecules*. Springer, Heidelberg, pp 123–185
- Rahmeh R, Damian M, Cottet M, Orcel H, Mendre C, Durroux T, Sharma KS, Durand G, Pucci B, Trinquet E, Zwier JM, Deupi X, Bron P, Banères JL, Mouillac B, Granier S (2012) Structural insights into biased G protein-coupled receptor signaling revealed by fluorescence spectroscopy. *Proc Natl Acad Sci U S A* 109:6733–6738
- Raschle T, Hiller S, Yu T-Y, Rice AJ, Walz T, Wagner G (2009) Structural and functional characterization of the integral membrane protein VDAC-1 in lipid bilayer nanodiscs. *J Am Chem Soc* 131:17777–17779

- Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VRP, Sanishvili R, Fischetti RF et al (2007) Crystal structure of the human beta(2) adrenergic G-protein-coupled receptor. *Nature* 450:383–384
- Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Par-don E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK (2011) Crystal structure of the β 2 adrenergic receptor-Gs protein complex. *Nature* 477:549–555
- Religa TL, Sprangers R, Kay LE (2010) Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science* 328:98–102
- Renault M, Saurel O, Czaplicki J, Demange P, Gervais V, Lohr F, Reat V, Piotto M, Milon A (2009) Solution state NMR structure and dynamics of KpOmpA, a 210 residue transmembrane domain possessing a high potential for immunological applications. *J Mol Biol* 385:117–130
- Renault M, Cukkemane A, Baldus M (2010) Solid-state NMR spectroscopy on complex biomol-ecules. *Angew Chem Int Ed* 49:8346–8357
- Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, Sligar SG (2009) Chapter 11 Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. In: Nejat D (ed) *Methods in Enzymology*, vol 464. Academic, Waltham, pp 211–231
- Rodnin MV, Posokhov YO, Contino-Pepin C, Brettmann J, Kyrichenko A, Palchevskyy SS, Pucci B, Ladokhin AS (2008) Interactions of fluorinated surfactants with diphtheria toxin T-domain: testing new media for studies of membrane proteins. *Biophys J* 94:4348–4357
- Roos C, Zocher M, Müller D, Münch D, Schneider T, Sahl HG, Scholz F, Wachtveitl J, Ma Y, Pro-verbio D et al. (2012) Characterization of co-translationally formed nanodisc complexes with small multidrug transporters, proteorhodopsin and with the *E. coli* MraY translocase. *Biochim Biophys Acta: Biomembr* 1818:3098–3106
- Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R et al (2005) NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* 307:1317–1321
- Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, DeVree BT, Sunahara RK, et al (2011) Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature* 469:236–240
- Ruschak AM, Religa TL, Breuer S, Witt S, Kay LE (2010) The proteasome antechamber maintains substrates in an unfolded state. *Nature* 467:868–871
- Sahdev S, Khattar SK, Saini KS (2008) Production of active eukaryotic proteins through bacter-ial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem* 307:249–264
- Salzmann M, Wider G, Pervushin K, Wuthrich K (1999) Improved sensitivity and coherence selec-tion for $[15N, 1H]$ -TROSY elements in triple resonance experiments. *J Biomol NMR* 15:181–184
- Sanders CR, Sonnichsen F (2006) Solution NMR of membrane proteins: practice and challenges. *Magn Reson Chem* 44:S24–S40
- Sarramegna V, Talmont R, Demange P, Milon A (2003) Heterologous expression of G-protein-coupled receptors: comparison of expression systems from the standpoint of large-scale pro-duction and purification. *Cell Mol Life Sci* 60:1529–1546
- Schlegel S, Löfblom J, Lee C, Hjelm A, Klepsch M, Strous M, Drew D, Slotboom DJ, de Gier J-W (2012) Optimizing membrane protein overexpression in the *Escherichia coli* strain Lemo21(DE3). *J Mol Biol* 423:648–659
- Shahid SA, Bardiaux B, Franks WT, Krabben L, Habeck M, van Rossum B-J, Linke D (2012) Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. *Nat Methods* 9:1212–1217
- Sharma M, Yi M, Dong H, Qin H, Peterson E, Busath DD, Zhou H-X, Cross TA (2010) Insight into the mechanism of the influenza A proton channel from a structure in a lipid bilayer. *Sci-ence* 330:509–512
- Shenkarev ZO, Paramonov AS, Lyukmanova EN, Shingarova LN, Yakimov SA, Dubinnyi MA, Chupin VV, Kirpichnikov MP, Blommers MJJ, Arseniev AS (2010) NMR structural and dy-

- namical investigation of the isolated voltage-sensing domain of the potassium channel KvAP: implications for voltage gating. *J Am Chem Soc* 132:5630–5637
- Shenkarev ZO, Paramonov AS, Lyukmanova EN, Gizatullina AK, Zhuravleva AV, Tagaev AA, Yakimenko ZA, Telezhinskaya IN, Kirpichnikov MP, Ovchinnikova TV et al (2013) Peptaibol antimoebin I: spatial structure, backbone dynamics, interaction with bicelles and lipid-protein nanodiscs, and pore formation in context of barrel-stave model. *Chem Biodivers* 10:838–863
- Shi P, Wang H, Xi Z, Shi C, Xiong Y, Tian C (2011) Site-specific ^{19}F NMR chemical shift and side chain relaxation analysis of a membrane protein labeled with an unnatural amino acid. *Protein Sci* 20:224–228
- Shi P, Li D, Chen H, Xiong Y, Wang Y, Tian C (2012) In situ ^{19}F NMR studies of an *E. coli* membrane protein. *Protein Sci* 21:596–600
- Shibata Y, Gvozdenovic-Jeremic J, Love J, Kloss B, White JF, Grisshammer R, Tate CG (2013) Optimising the combination of thermostabilising mutations in the neurotensin receptor for structure determination. *Biochim Biophys Acta* 1828:1293–1301
- Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, Kobayashi T, Stevens RC, Iwata S (2011) Structure of the human histamine H1 receptor complex with doxepin. *Nature* 475:65–70
- Shintre CA, Pike ACW, Li Q, Kim JI, Barr AJ, Goubin S, Shrestha L, Yang J, Berridge G, Ross J et al (2013) Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proceedings of the National Academy of Sciences of the USA* 110: 9710–9715
- Sobhanifar S, Reckel S, Junge F, Schwarz D, Kai L, Karbyshev M, Lohr F, Bernhard F, Dotsch V (2010a) Cell-free expression and stable isotope labelling strategies for membrane proteins. *J Biomol NMR* 46:33–43
- Sobhanifar S, Schneider B, Lohr F, Gottstein D, Ikeya T, Mlynarczyk K, Pulawski W, Ghoshdastider U, Kolinski M et al. (2010b) Structural investigation of the C-terminal catalytic fragment of presenilin 1. *Proc Natl Acad Sci U S A* 107:9644–9649
- Sobolevsky AI, Rosconi MP, Gouaux E (2009) X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature* 462:745–756
- Standfuss J, Xie G, Edwards PC, Burghammer M, Oprian DD, Schertler GF (2007) Crystal structure of a thermally stable rhodopsin mutant. *J Mol Biol* 372:1179–1188
- Standfuss J, Edwards PC, D'Antona A, Fransen M, Xie G, Oprian DD et al (2011) The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* 471:656–660
- Steyaert J, Kobilka BK (2011) Nanobody stabilization of G protein-coupled receptor conformational states. *Curr Opin Struct Biol* 21:567–572
- Stock D, Leslie AG, Walker JE (1999) Molecular architecture of the rotary motor in ATP synthase. *Science* 286:1700–1705
- Tani K, Mitsuma T, Hiroaki Y, Kamegawa A, Nishikawa K, Tanimura Y, Fujiyoshi Y (2009) Mechanism of aquaporin-4's fast and highly selective water conduction and proton exclusion. *J Mol Biol* 389:694–706
- Talbot JC, Dautant A, Polidori A, Pucci B, Cohen-Bouhacina T, Maali A, Salin B, Brèthes D, Velours J, Giraud MF (2009) Hydrogenated and fluorinated surfactants derived from Tris (hydroxymethyl)-acrylamidomethane allow the purification of a highly active yeast F1-F0 ATP-synthase with an enhanced stability. *J Bioenerg Biomembr* 41:349–360
- Tate CG (2010) Practical considerations of membrane protein instability during purification and crystallisation. *Methods Mol Biol* 601:187–203
- Tate CG, Whiteley E, Betenbaugh MJ (1999) Molecular chaperones stimulate the functional expression of the cocaine-sensitive serotonin transporter. *J Biol Chem* 274:17551–17558
- Tate CG, Haase J, Baker C, Boorsma M, Magnani F et al (2003) Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochim Biophys Acta* 1610:141–153
- Tate CG (2012) A crystal clear solution for determining G-protein-coupled receptor structures. *Trends Biochem Sci* 37:343–352

- Tao X, Avalos JL, Chen J, MacKinnon R (2009) Crystal structure of the eukaryotic strong inward-rectifier K⁺ channel Kir2.2 at 3.1 Å resolution. *Science* 326:1668–1674
- Tao X, Le A, Limapichat W, Dougherty DA, MacKinnon R (2010) A gating charge transfer center in voltage sensors. *Science* 328:67–73
- Thompson AA, Liu JJ, Chun E, Wacker D, Wu H, Cherezov V, Stevens RC (2011) GPCR stabilization using the bicelle-like architecture of mixed sterol-detergent micelles. *Methods* 55:310–317
- Thompson AA, Liu W, Chun E, Katritch V, Wu H, Vardy E, Huang XP, Trapella C, Guerrini R, Calo G, Roth BL, Cherezov V, Stevens RC (2012) Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature* 485:395–399
- Toyoshima C, Nakasako M, Nomura H, Ogawa H (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature* 405:647–655
- Traaseth NJ, Shi L, Verardi R, Mullen DG, Barany G, Veglia G (2009) Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach. *Proc Natl Acad Sci* 106:10165–10170
- Tribet C, Audebert R, Popot J-L (1996) Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci U S A* 93:15047–15050
- Ujwal R, Bowie JU (2011) Crystallizing membrane proteins using lipidic bicelles. *Methods* 55:337–341
- Unnerståle S, Måler L, Draheim RR (2011) Structural characterization of AS1-membrane interactions from a subset of HAMP domains. *Biochim Biophys Acta: Biomembr* 1808:2403–2412
- Unwin N (2005) Refined structure of the nicotinic acetylcholine receptor at 4 Å resolution. *J Mol Biol* 346:967–989
- Vaiphei ST, Tang Y, Montelione G, Inouye M (2011) The use of the condensed single protein production system for isotope-labeled outer membrane proteins, OmpA and OmpX in *E. coli*. *Mol Biotechnol* 47:205–210
- Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM (2013) Molecular signatures of G-protein-coupled receptors. *Nature* 494:185–194
- Verardi R, Shi L, Traaseth NJ, Walsh N, Veglia G (2011) Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. *Proc Natl Acad Sci U S A* 108:9101–9106
- Wacker D, Wang C, Katritch V, Han GW, Huang XP, Vardy E, McCorvy JD, Jiang Y, Chu M, Siu FY, Liu W, Xu HE, Cherezov V, Roth BL, Stevens RC (2013) Structural features for functional selectivity at serotonin receptors. *Science* 340:615–619
- Wada T, Shimono K, Kikukawa T, Hato M, Shinya N, Kim SY, Kimura-Someya T, Shirouzu M, Tamogami J, Miyauchi S et al (2011) Crystal structure of the eukaryotic light-driven proton-pumping rhodopsin, *Acetabularia* rhodopsin II, from marine alga. *J Mol Biol* 411:986–998
- Wang C, Wu H, Katritch V, Han GW, Huang XP, Liu W, Siu FY, Roth BL, Cherezov V, Stevens RC (2013) Structure of the human smoothed receptor bound to an antitumour agent. *Nature* 497:338–343
- Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454:486–491
- Warne T, Moukhametzianov R, Baker JG, Nehmé R, Edwards PC, Leslie AG, Schertler GF, Tate CG (2011) The structural basis for agonist and partial agonist action on a β(1)-adrenergic receptor. *Nature* 469:241–244
- Warne T, Edwards PC, Leslie AG, Tate CG (2012) Crystal structures of a stabilized β1-adrenoceptor bound to the biased agonists bucindolol and carvedilol. *Structure* 20:841–849
- Warschawski DE, Arnold AA, Beaugrand M, Gravel A, Chartrand E, Marcotte I (2011) Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochim Biophys Acta: Biomembr* 1808:1957–1974
- White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG, Grisshammer R (2012) Structure of the agonist-bound neurotensin receptor. *Nature* 490:508–513

- Whorton MR, MacKinnon R (2011) Crystal Structure of the Mammalian GIRK2 K⁺ Channel and Gating Regulation by G Proteins, PIP₂, and Sodium. *Cell* 147:199–208
- Whorton MR, MacKinnon R (2013) X-ray structure of the mammalian GIRK2-beta gamma G-protein complex. *Nature* 498:190–197
- Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* 330:1066–1071
- Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, Liu W, Thompson AA, Huang XP, Carroll FI, Mascarella SW, Westkaemper RB, Mosier PD, Roth BL, Cherezov V, Stevens RC (2012) Structure of the human κ -opioid receptor in complex with JDTic. *Nature* 485:327–332
- Xie G, Timasheff SN (1997) The thermodynamic mechanism of protein stabilization by trehalose. *Biophys Chem* 64:25–43
- Xu F, Stevens RC (2011) Trapping small caffeine in a large GPCR pocket. *Structure* 19:1204–1207
- Xu Y, Kong J, Kong W (2013) Improved membrane protein expression in *Lactococcus lactis* by fusion to Mistic. *Microbiology* 159:1002–1009
- Yang DW, Kay LE (1999) Improved lineshape and sensitivity in the HNCO-family of triple resonance experiments. *J Biomol NMR* 14:273–276
- Zhang C, Srinivasan Y, Arlow DH, Fung JJ, Palmer D, Zheng Y, Green HF, Pandey A, Dror RO, Shaw DE et al (2012) High-resolution crystal structure of human protease-activated receptor 1. *Nature* 492:387–392
- Zhang QH, Ma XQ, Ward A, Hong WX, Jaakola VP, Stevens RC, Finn MG, Chang G (2007) Designing facial amphiphiles for the stabilization of integral membrane proteins. *Angew Chem Int Ed* 46:7023–7025
- Zhou Y, Bowie JU (2000) Building a thermostable membrane protein. *J Biol Chem* 275:6975–6979

Chapter 2

Membrane Protein Quality Control in Cell-Free Expression Systems: Tools, Strategies and Case Studies

Davide Proverbio, Erik Henrich, Erika Orbán, Volker Dötsch
and Frank Bernhard

2.1 Introduction

Cell-free (CF) expression has emerged in the last decade as an efficient and fast approach for the production of membrane proteins (MPs) of diverse topologies and origin. Its unique design as an open accessible reaction helps to eliminate several central bottlenecks known from conventional cell-based MP expression systems. In general, problems with cell physiology, expression regulation and cell culture are reduced. On the other hand, the high diversity of CF reaction conditions requests increased time investments in controlling MP quality, fine-tuning of reaction conditions and designing sample evaluation strategies. Poor MP sample quality can be the result if this important requirement is overseen.

CF reactions can basically be operated in two flavors, the single compartment batch configuration and the two-compartment continuous exchange (CECF) configuration (Kigawa and Yokoyama 1991; Spirin et al. 1988). The batch configuration is the method of choice in throughput applications using microplate devices and analytical scale reactions (Kai et al. 2013; Savage et al. 2007; Schwarz et al. 2010). Batch reaction times are limited to few hours with consequently lower yields of protein, although a number of modifications are possible in order to considerably improve efficiencies. Higher protein yields are typically obtained with the CECF configuration containing a reaction mixture (RM) compartment containing all the high molecular weight compounds such as ribosomes, DNA template and enzymes, and a feeding mixture (FM) compartment with a certain amount of precursors such as amino acids and nucleotides. Protocols for batch and CECF configurations are highly variable and among others, expression efficiencies depend on (1) precursor

F. Bernhard (✉) · D. Proverbio · E. Henrich · E. Orbán · V. Dötsch
Centre for Biomolecular Magnetic Resonance, Institute of Biophysical Chemistry,
Goethe University Frankfurt am Main, Frankfurt am Main, Germany
e-mail: fbern@bpc.uni-frankfurt.de

concentrations, (2) energy regeneration systems, (3) RM–FM volume ratios and (4) the implementation of repeated FM exchanges.

An array of new applications, modifications, and strategies for the CF production of MP samples has been developed within the last decade. In particular, the tools for the modulation of MP quality already during translation by CF reaction condition tuning have been widely expanded. We therefore provide a current view on options and perspectives for successful MP production and we summarize diverse strategies based on CF expression technologies.

2.2 Selecting the Background: Different Extract Sources

The origin of the CF extract is the first selection to be made by approaching MP expression. In particular, within the last decade, a considerable number of new extract sources covering eukaryotic as well as prokaryotic origins have been introduced (Table 2.1). Major selection criteria before starting a CF expression approach are usually (1) the required amount of synthesized recombinant protein, (2) to provide the most favorable background for promoting protein folding, (3) to increase the likeability of posttranslational modifications, (4) general handling issues, system availability and costs.

Expression efficiencies and other characteristics of the various systems still differ significantly and best compromises have to be found. While few micrograms of recombinant protein can usually be obtained in any system, the production of preparative scale levels approaching milligram yields out of 1 ml of RM is currently only routinely possible with extracts of *Escherichia coli* or wheat germs. Frequent limiting factors for protein production efficiency in cell extracts are high concentrations of endogenous degrading enzymes, poor synchronization of ribosome activity during cell growth, or stability problems of essential enzymes. It should be noted that extracts of cells showing even high expression activities *in vivo* such as yeasts might not be very efficient in CF expression. However, protocols in particular for the efficient CECF configuration are continuously being optimized and further potential for improved protein synthesis might exist. Most systems have now been adjusted as coupled transcription/translation systems including the efficient T7 promoter for protein production and accepting plasmid or linear DNA templates (Table 2.1). The addition of translation factors or considering specific template modifications might further be necessary depending on the selected system.

A critical issue is the availability of the different CF extracts. Most systems are available as standardized commercial kits, but quality optimization and specific applications often require the set up of individual expression reactions. The preparation protocols for the various cell extracts differ significantly with sometimes even high variations in extract batch quality (Table 2.1). For eukaryotic cell extract preparations, species possible to grow in defined cell cultures might be preferred or commercial sources might be considered. The relatively fast and efficient preparation protocol is a major advantage of using *E. coli* extracts. In addition, it is best

Table 2.1 CF extract sources

CF extract source	PTM ^a	cT7 ^b	CECF ^c	Yield ^d (mg/ml)	MP ^e	Kit ^f	Protocol ^g
<i>Eukaryotes</i>							
Human HeLa S3 cells	+	+	+	0.05–0.2 ≤ 0.01	NA	+	Mikami et al. (2006, 2008); Witherell et al. (2001)
CHO	+	+	–	≤ 0.01	+	+	Arduengo et al. (2007); Craig et al. (1992)
Rabbit reticulocytes	+	+	+	≤ 2 (CECF)	+	+	Endo and Sawasaki (2005); Madin et al. (2000); Sawasaki et al. (2007); Takai et al. (2010)
Wheat germ	+	+	–	0.1–0.4 (B)	+	+	Ezure et al. (2007, 2010); Katzen and Kudlicki (2006); Stech et al. (2012)
<i>Spodoptera frugiperda</i> 9, 21	–	–	–	0.05–0.07	NA	–	Wang (2006); Wang et al. (2008)
<i>Saccharomyces cerevisiae</i>	–	+	–	≤ 0.2	NA	+	Kovtun et al. (2011)
<i>Leishmania tarentolae</i>	–	–	–	≤ 5 (CECF)	+	+	Katzen et al. (2005); Schwarz et al. (2008)
<i>Prokaryotes</i>							
<i>E. coli</i>	–	–	–	0.5–1 (B)	NA	–	Nakashima and Tamura (2004)
<i>Pseudomonas fluorescens</i> (reaction at 8–30 °C)	–	–	–	≤ 0.1	+	+	Shimizu et al. (2001)
<i>E. coli</i> –PURE (purified translation machinery)	–	–	–	≤ 0.5	NA	–	Uzawa et al. (2002); Zhou et al. (2012)
<i>Thermus thermophilus</i> (reaction at 37–65 °C)	–	–	–	≤ 0.06	NA	–	Endoh et al. (2008)
<i>Thermococcus kodakaraensis</i> (reaction at 40–80 °C)	–	–	–	≤ 0.007	NA	–	

^a Posttranslational modifications other than disulfide bridge formation reported

^b Coupled transcription/translation protocol with T7 promoter described

^c Protocols for the more efficient CECF configuration established

^d Approximate range of protein yield in 1 ml of RM, B batch configuration, CECF continuous exchange configuration

^e Expression of MPs reported; NA data not yet available

^f Commercial reaction kits available

^g Representative recent protocols for extract preparation

characterized and a large variety of compounds useful for reaction modifications is available. The vast majority of current data on MP production have thus been obtained with *E. coli* extracts.

Depending on the intended applications, the proper formation of posttranslational modifications can be a key issue for protein sample quality evaluation. Disulfide bridge formation may be triggered independently from extract origins by modulating the reducing conditions, e.g., by adding redox systems into the reaction, by supporting disulfide bridge formation with chaperones or by chemical pretreatment of extracts (Goerke and Swartz 2008, Kim and Swartz 2004, Yin and Swartz 2004). More complex modifications, such as glycosylation, lipidation, or phosphorylation, are so far only described from systems with eukaryotic extracts such as rabbit reticulocytes, insect cells, or wheat germ and at analytical scales (Table 2.1). Many modifications require supplements such as canine pancreas microsomes into the CF reaction. If modifying enzymes are provided, posttranslational modifications such as N-glycosylation appear to be possible even in extracts of *E. coli* (Guarino and DeLisa 2012). However, it might stay challenging to combine quality and homogeneity of posttranslational modifications with high-level expression purposes.

2.3 Basic Protocol Development: Improving CF Expression Efficiency

Complexity of MP production in CF systems is mainly reduced to the basic transcription/translation process. Coordination of pathways for trafficking or translocation as well as suppressing toxic effects are usually less relevant issues. Protein expression in most CF systems is controlled by the phage T7-RNA polymerase, and the corresponding regulatory promoter and terminator elements in addition to system specific enhancers have to be provided. However, other promoters could work as well. With *E. coli* extracts, derivatives of standard *Ptac* promoters recognized by the endogenous *E. coli* RNA polymerase could give even relatively high expression levels (Shin and Noireaux 2010). DNA template constructs can be generated by overlap polymerase chain reaction (PCR) strategies and added as linear DNA fragment into the CF reaction (Ahn et al. 2005; Yabuki et al. 2007). Alternatively, plasmid DNA templates based on standard vectors such as, e.g., the pET or pIVEx series can be provided. DNA templates appear to be quite stable in CF reactions and final concentrations in between 2 and 10 ng/ μ l RM are already saturating (Habersstock et al. 2012).

Initial problems with low expression efficiency are mainly associated with the translation process. Adjusting the proper Mg^{2+} ion optimum is mandatory for each new target and suboptimal conditions can have severe impacts on protein production (Rath et al. 2011; Schwarz et al. 2007). Abundance of rare codons could further reduce protein expression and induce mis-incorporation of amino acids or even the premature termination of translation. Low protein yields are even more frequently caused by the formation of unfavorable secondary structures of the mRNA

involving the 5-prime end containing the translational initiation site. Modulating the nucleotide sequence of the 5-prime coding sequence can therefore be very efficient in order to improve expression (Ahn et al. 2007; Kralicek et al. 2011). A fast approach is the tag variation screen by analyzing the effects of a small number of short sequence-optimized expression tags (Haberstock et al. 2012). The tag variation constructs are generated by overlap PCR and the resulting products can directly be used as DNA templates in CF expression screens. The construction of large fusion proteins in order to improve expression is therefore usually not necessary. Expression monitoring can initially be performed via immunodetection by using C-terminal purification tags such as a poly(His)₁₀-tag as antigen. In an ideal template design, the coding sequence is therefore modified with a C-terminal purification/detection tag, and, if necessary, with a short N-terminal expression tag (Fig. 2.1). If translation can be addressed properly with the above mentioned procedures, the protein production in CF systems is usually very efficient. In expression screens comprising MP targets of diverse sizes, topologies, and functions, high success rates could be achieved (Schwarz et al. 2010; Savage et al. 2007; Langlais et al. 2007).

Expression monitoring by taking advantage of C-terminally attached derivatives of green fluorescent protein (GFP) could be useful for CF expression protocol development and fast protein quantification (Kai et al. 2013; Müller-Lucks et al. 2012; Nozawa et al. 2011; Roos et al. 2012). For MP expression, it must be considered that the folding of wild-type or red-shifted variants of GFP is hampered in the presence of most detergents (Roos et al. 2012). More resistant is the superfolder GFP derivative most likely due to its higher tolerance for chemical denaturants and its faster folding kinetics (Roos et al. 2012; Pedelacq et al. 2006). However, the folding of superfolder GFP might not correlate with the productive folding of the N-terminal target protein as it is speculated for other GFP derivatives (Pedelacq et al. 2006). Superfolder GFP might therefore only be considered as general expression monitor while fusions with other GFP derivatives may in addition also give some preliminary evidence of the target protein folding and quality.

2.4 Folded Precipitates: P-CF Expression

Depending on the strategy and choice of supplemented additives, several basic expression modes are possible for the CF production of MPs (Fig. 2.1). The selection of the expression mode may depend on the intended application of the MP sample, but it can also have drastic consequences on the resulting MP quality (Junge et al. 2010; Lyukmanova et al. 2012). An overview on the implementation of the different CF expression modes is given in Tables 2.2–2.4 and representative case studies published during the last decade are listed.

In absence of any provided hydrophobic environment, the freshly translated MPs instantly precipitate in the RM. Successful expression in this precipitate forming (P-CF) production mode can thus even be monitored by increased turbidity of the RM during incubation. Folded structures of such P-CF-generated MP precipitates

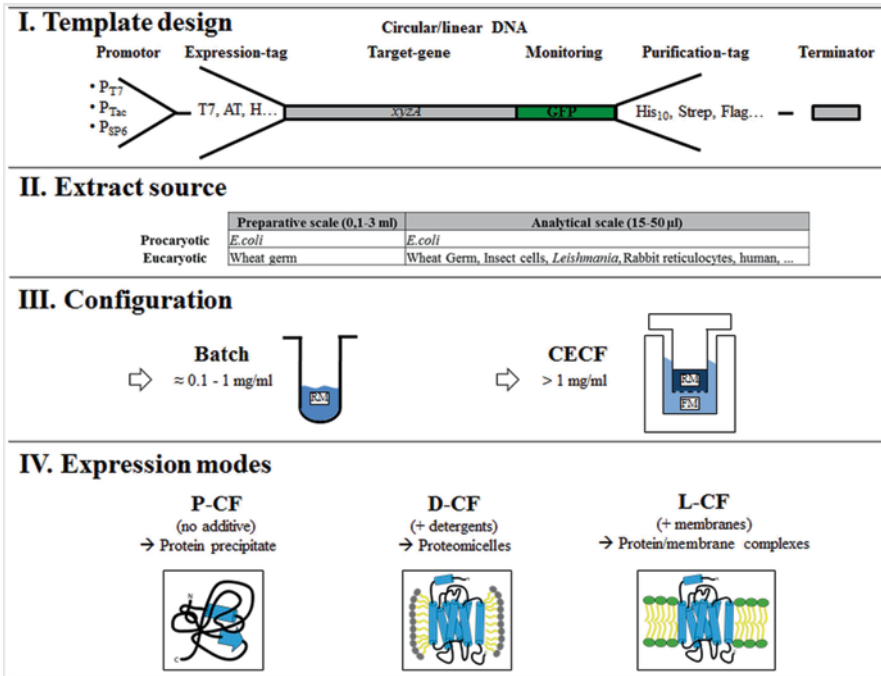


Fig. 2.1 Basic steps for the design of CF expression reactions

can be detected by solid state nuclear magnetic resonance (NMR) and resolubilized precipitates show significant structural overlaps with corresponding MP samples obtained after conventional *in vivo* production (Maslennikov et al. 2010). P-CF-expressed MPs can simply be harvested by centrifugation. The MP pellet is usually contaminated with a number of co-precipitated proteins from the extract. Washing with buffer containing mild detergents such as Brij derivatives can help to selectively reduce such contaminations. The MPs are then solubilized in buffer containing specific detergents. Best results are usually obtained with 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)] (LMPG), 1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-RAC-(1-glycerol)] (LPPG), or sodium dodecyl sulfate (SDS; Table 2.2; Klammt et al. 2004; Klammt et al. 2012; Rath et al. 2011). Milder detergents such as n-dodecyl-phosphocholine (DPC) or n-dodecyl-b-D-maltoside (DDM), detergent cocktails or mixtures of detergents and lipids could further be useful depending on the MP target (Ma et al 2011).

Critical parameters for the resulting MP quality can be (1) detergent concentration and volume of the solubilization buffer, (2) temperature of solubilization, and (3) the subsequent exchange of the primary and usually relatively harsh solubilization detergent against secondary and considerably milder detergents, e.g., upon MP immobilization during affinity chromatography. Stabilization and high recovery of ligand binding active GPCRs could be obtained by this strategy (Junge et al. 2010;

Table 2.2 Case studies of P-CF-expressed MPs

Protein ^a	Size [kDa] (TMH ^b)	Type/assay ^c	System ^d yield ^e	Solubilization ^f (%)	Reference
Bs-MraY	36 (10)	Enzyme/+	iE [3]	Triton X-100 [2], DPC [2], DHPC [2], DDM [2], LS [0.8], LMPG [0.75], SDS [2]	Ma et al. (2011)
mAqp4-M23	30 (6)	Channel/+	iE [3]	Fos12 [1], DHPC [2], Fos16 [2], LMPG [2], LPPG [1]	Kai et al. (2010)
hETB, hNPY2/5	39–51	GPCR/-	iE [3]	e.g. LMPG [2]	Schneider et al. (2010)
hMTNRIA/B					
hSSI/2, hV1BR					
hHRH1, hV2R					
rCRF					
hPSI-CTF	16 (3)	Protease/-	iE [3]	SDS [1–2]	Sobhanifar et al. (2010)
hCRFR1	47 (7)	GPCR/+	iE [3]	LMPG [2]/Nvoy	Klammt et al. (2011)
mCRFR2β	49 (7)				
KvAP-VSD	16.5 (4)	Channel/+	iE [2]	SDS [1]/DPC [0.2]	Lyukmanova et al. (2012)
<i>A. permix</i>					
hErbB3 (639–670)	5 (1)	Receptor/+	iE [3]	DPC	Mineev et al. (2011)
hErbB3 (632–675)	6 (1)	Receptor/+	iE [3]	SDS [1]	Khabibullina et al. (2010)
Ec-ArcB (1–115)	11–21	Sensor/+	iE [3]	LMPG [5]	Maslennikov et al. (2010)
Ec-QseC (1–185)	(2–4)				
Ec-KdpD (397–502)					
6 hMPs	~11 (2–3)	Unknown	iE [3]	LMPG	Klammt et al. (2012)
hLAPT4A	28 (4)	Transporter	iE [3]	Fos14	Nguyen et al. (2010)
hCX32	32 (4)	Gap junction/+			
hGLUT4	55 (12)	Transporter			
hVDAC1	36 (19)	Channel/+			
F ₁ F ₀ -ATP synthase <i>C. thermarum</i>	542 (26)	Enzyme/+	iE [1]	DDM [2]	Matthies et al. (2011)
PorA/H <i>C. glutamicum</i>	5 (7)	Channel/+	iE [3]	LDAO [1], LPPG [1], LMPG [1], LMPC [1], Triton X-100 [1], DHPC [1], DPC [1], SDS [1]	Rath et al. (2011)
134 Ec-MPs	≤112 (≤15)	–	iE [3]	SDS [0.25], LPPG [0.25], LMPC [0.25], DPC [0.1]	Schwarz et al. (2010)

Table 2.2. (continued)

Protein ^a	Size [kDa] (TMH ^b)	Type/assay ^c	System ^d yield ^e	Solubilization ^f (%)	Reference
hETA/B	49/50 (7)	GPCR/+	iE [3]	LPPG [1], LMPC [1], SDS [1], Fos16 [1], Fos12 [1]	Junge et al. (2010)
Ec-EmrE	11 (4)	Transporter/+	iE [3]	DDM [2], DPC [1], LMPG	Klammt et al. (2004)
Ec-TehA	36 (10)				
Ec-SugE	11 (4)				
rOCT1/2	55 (12)	Transporter/+	iE [2]	LMPG [1]	Keller et al. (2008, 2011)
rOAT-1	60 (12)				
MPs of diverse origin	Diverse	Diverse	iE [3]	e.g. LMPG, DPC	Schwarz et al. (2007)
Proteorhodopsin	27 (7)	H ⁺ -Pump/+	cE [2]	e.g. LMPG [0.01], LPPG [0.025]	Gourdon et al. (2008)
Bacteriorhodopsin	28(7)	H ⁺ -Pump/+	iE	SDS, refolding	Shenkarev et al. (2013)
<i>E. sibiricum</i>					
hH1R	56 (7)	GPCR/+	cE [1]	DDM [2]	Sansuk et al. (2008)
~120 MPs	10–30 (1–9)	Diverse	iE	nOG, DDM	Savage et al. (2007)
LH1- α -apoprotein	5 (1)	Light harvesting/+	cE [2]	Triton-X100 [0.5–2]	Shimada et al. (2004)
<i>R. rubrum</i>					
> 100 hMPs	8–134	Diverse	cE/cWG –		Langlais et al. (2007)

DDM n-dodecyl- β -D-maltoside, DHPC 1,2-dihexanoyl-sn-glycero-3-phosphocholine, DPC=Fos12 n-dodecylphosphocholine, Fos/4 n-tetradecylphosphocholine, Fos/6 n-hexadecylphosphocholine, LDAO lauryldimethylamine oxide, LMPG 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac(1-choline)], LMPG 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac(1-glycerol)], LPPG 1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-rac(1-glycerol)], Nvyoy NV10 polymer, SDS sodium dodecylsulfate

^a Approximate size, if documented, the origin of proteins is given in italics; *h* human; *m* murine; *r* rat; *Ec* *E. coli*, *Bs* *Bacillus subtilis*

^b TMH Proposed number of transmembrane helices

^c +: Quality analyzed by structural evaluation or functionality

^d iE Individual *E. coli* extract, cE commercial *E. coli* extract; cMG commercial wheat germ extract

^e Approximate yields per one ml RM if documented in the corresponding references. 1: ≤ 0.1 mg/ml; 2: 0.1–1 mg/ml; 3: > 1 mg/ml

^f Detergents and concentrations used for posttranslational solubilization. Exchange into secondary detergents are indicated by slash. Concentrations are given if documented

Klammt et al. 2011). Solubilization of P-CF pellets is fast and usually complete after gentle shaking for approximately 1 h. It should be noted that pellets of CF-expressed soluble proteins cannot usually be solubilized by that procedure as they are much more unstructured. Consequently, MPs having excessive soluble domains could therefore resist solubilization out of P-CF pellets.

Selecting the P-CF expression mode is the fastest approach and usually routinely employed for the first level of MP expression protocol development in order to tune protein production up to the desired yields (Junge et al. 2011). Even complex MPs such as 12 transmembrane segment containing eukaryotic ion transporters or the 10 transmembrane segment containing MraY translocase have been functionally synthesized in the P-CF mode (Keller et al. 2008; Ma et al. 2011). The P-CF mode is furthermore excellent for screening MP libraries (Langlais et al. 2007; Savage et al. 2007; Schwarz et al. 2010) and for the production of MP samples for structural analysis by NMR (Klammt et al. 2004; Maslennikov et al. 2010; Rath et al. 2011; Sobhanifar et al. 2010).

2.5 Production of Proteomicelles: D-CF Expression

CF expression systems can tolerate a considerable number of supplied hydrophobic compounds, while certain variations in between the different extract sources exist. The CF expression in the presence of detergents above their critical micellar concentration (CMC) can result into the co-translational solubilization of the expressed MPs and into the instant formation of proteomicelles (Fig. 2.1).

Extensive evaluation of detergent tolerance has been performed with *E. coli* extracts (Blesneac et al. 2012; Gourdon et al. 2008; Klammt et al. 2005; Lyukmanova et al. 2012) as well as with wheat germ extract systems (Beebe et al. 2011; Genji et al. 2010; Kaiser et al. 2008; Periasamy et al. 2013). As primary compounds of choice, long-chain polyoxyethylene-alkyl-ethers such as Brij35, Brij58, Brij78 or Brij98, and the steroid-derivative digitonin have been determined (Table 2.3). These detergents have been successfully used for the solubilization of different G protein-coupled receptors (GPCRs) as well as of prokaryotic MPs (Table 2.3). Commonly employed detergents for the extraction of MPs out of native membranes such as DPC, the alkyl-glucoside n-dodecyl- β -D-maltoside (DDM), or n-octyl- β -D-glucopyranoside (β -OG) are too harsh or only tolerated at lower concentrations. However, the tolerance can sometimes be increased if critical detergents are provided as mixed micelles together with other detergents, e.g., CHAPS together with Fos-choline derivatives (Genji et al. 2010). It might generally be advantageous to combine the provided detergent micelles with some small amounts of lipids in case the translated MPs require interaction with some lipids for stabilization (Arslan Yildiz et al. 2013; Müller-Lücks et al. 2013; Nozawa et al. 2007). For the expression of mitochondrial carrier proteins, the addition of some cardiolipin together with fluorinated surfactants or Brij35 detergent had significant beneficial effects, whereas cardiolipin had negative effects in combination with the detergent Brij58

Table 2.3 Case studies of D-CF-expressed MPs

Protein ^a	Size [kDa] (TMH ^b)	Type/assay ^c	System ^d Yield ^e	Detergent ^f	Reference
<i>Pores and Channels</i>					
Aqp3	32 (6)	Porin/+	iE [3]	Brij98 + Ec polar lipids	Müller-Lucks et al. (2013)
Cx32	32 (4)	Channel/+	cE [2]	Brij35	Nguyen et al. (2010)
VDAC	36 (1+13B)				
hVDAC1	35 (13B)	Channel/+	cE [1]	DDM, Fos12	Deniaud et al. (2010)
mAqp4	30 (6)	Porin/+	iE [3]	Brij35, Digitonin	Kai et al. (2010)
Ec-MscL	15 (2)	Channel/+	cE [3]	Triton X-100	Berrier et al. (2004); Abdine et al. (2010)
OEP24	24 (12B)	Channel/+	cE [2]	DDM	Liguori et al. (2010)
PorA/H	5	Channel/+	iE [2]	Brij72	Rath et al. (2011)
<i>C. glutamicum</i>					
hERG	25 (6)	Channel/+	iE [2]	Brij78 + soybean PC	Arslan Yildiz et al. (2013)
<i>Transporters and pumps</i>					
UCP1	30–35 (6)	Carrier/+	cE [2]	Brij35/58, DDM, digitonin, fluorinated surfactants + cardiolipin	Blesneac et al. (2012)
Ec-EmrE	12 (4)	Trans- porter/+	cE [3]	DDM	Elbaz et al. (2004)
AtPPT1, OpPPT1/2/3	30 (8)	Trans- porter/+	cE	Brij35 + Asolectin	Nozawa et al. (2007)
Bacteriorhodop- sin	28 (7)	H ⁺ -Pump/+	cE [2]	NaPol	Bazzacco et al. (2012)
Bacteriorhodop- sin	28 (7)	H ⁺ -Pump/+	WG	Chaps + Fos12, Fos14	Genji et al. (2010)
Bacteriorhodop- sin	28 (7)	H ⁺ -Pump/+	WG	Chaps, Fos12	Beebe et al. (2011)
Ec-Tsx	34 (12B)	Trans- porter/-	iE [3]	Brij78	Klammt et al. (2005)
<i>Receptors</i>					
Dopamine D2	50 (7)	GPCR/+	iE WG	–	Basu et al. (2013)
hTAAR-T4L	45 (7)	GPCR/+	iE [2]	Brij35	Wang et al. (2013)
hETA, hETB	~45 (7)	GPCR/+	iE [3]	Brij35/78	Junge et al. (2010)
Olfactory Receptors, hFPR3, hVN1R1, hVN1R5	~30 (7)	GPCR/+	cE [2]	Brij35, peptide surfactants	Corin et al. (2011)
Cytokinin Receptor CRE1/AHK4	37 (2)	Receptor/+	iE [3]	Brij58/78	Wulfetange et al. (2011)
CpxA	50 (2)	Receptor/+	cE [3]	Brij35	Miot and Betton (2011)

Table 2.3 (continued)

Protein ^a	Size [kDa] (TMH ^b)	Type/assay ^c	System ^d Yield ^e	Detergent ^f	Reference
OR17-4	36 (7)	GPCR/+	WG, cE [2]	Digitonin	Kaiser et al. (2008)
hMTNR1B, hNPY4R, rCRF, hV2R	~40 (7)	GPCR/-	iE [2]	Brij58/78	Klammt et al. (2007)
hCHRM2, hβ2AR, hNTR	~60 (7)	GPCR/+	iE [2]	Brij35, digitonin	Ishihara et al. (2005)
ARII	18 (7)	Receptor/+	cE [1]	Digitonin + PC	Wada et al. (2011)
hOR17-210, mOR103- 15, hFPR3, hTAAR5	~40 (7)	GPCR/+	iE [2]	Peptide surfactants	Wang et al. (2011)
hCRF1, CRF2β	(7)	GPCR/+	iE	Nvoy	Klammt et al. (2011)
<i>Enzymes</i>					
Bs-MraY	36 (10)	Translo- case/+	iE [3]	Brij35	Ma et al. (2011)
Bcl-2	25 (1)	Anti- apoptotic protein/+	iE [2]	Brij58	Pedersen et al. (2011)
ATP synthase	542	Multisub- unit com- plex/+	iE [1]	Brij58	Matthies et al. (2011)
CrdS	70 (7)	Enzyme/+	WG [1]	Brij58, peptide surfactants	Periasamy et al. (2013)
Bs-DesK	~40 (4–5)	Histidine kinase/+	cE [3]	Brij58, digitonin, Triton X-100	Martin et al. (2009)
<i>Diverse</i>					
>100 Ec-MPs	(<15)	Diverse	iE [3]	Brij35/58/78/98	Schwarz et al. (2010)
TM-ErbB3	5 (1)	Receptor/+	iE [3]	Brij35/58/78/98	Lyukmanova et al. (2012)
VSD-KvaP	(4)	Channel		Triton X-100, DDM	
Bacteriorhodop- sin	28 (7)	H ⁺ -Pump/+			

Brij35 polyoxyethylene-(23)-lauryl-ether, *Brij58* polyoxyethylene-(20)-cetyl-ether, *Brij72* Polyoxethylene-(2)-stearyl-ether, *Brij78* polyoxyethylene-(20)-stearyl-ether, *Brij98* polyoxyethylene-(20)-oleyl-ether, *CHAPS* 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonat, *DDM* n-dodecyl-β-D-maltoside, *Fos12* n-dodecylphosphocholine, *Fos14* n-tetradecylphosphocholine, *NaPol* Nonionic amphipols, *Nvoy* NV10 polymer, *PC* L-α-phosphatidylcholine

^a If documented, the origin of proteins is given in italics; *h* human, *m* murine, *r* rat, *Ec* *E. coli*, *Bs* *Bacillus subtilis*

^b TMH: Number of transmembrane helices or β-sheets (β)

^c +: Quality analyzed by structural evaluation or functionality

^d *iE* individual *E. coli* extracts, *cE* commercial *E. coli* extracts, *WG* wheat germ extracts

^e Approximate yields per 1 ml RM if documented in the corresponding references. 1: ≤0.1 mg/ml;

2: 0.1–1 mg/ml; 3: >1 mg/ml

^f Main detergents used for co-translational solubilization. Concentrations are given if documented

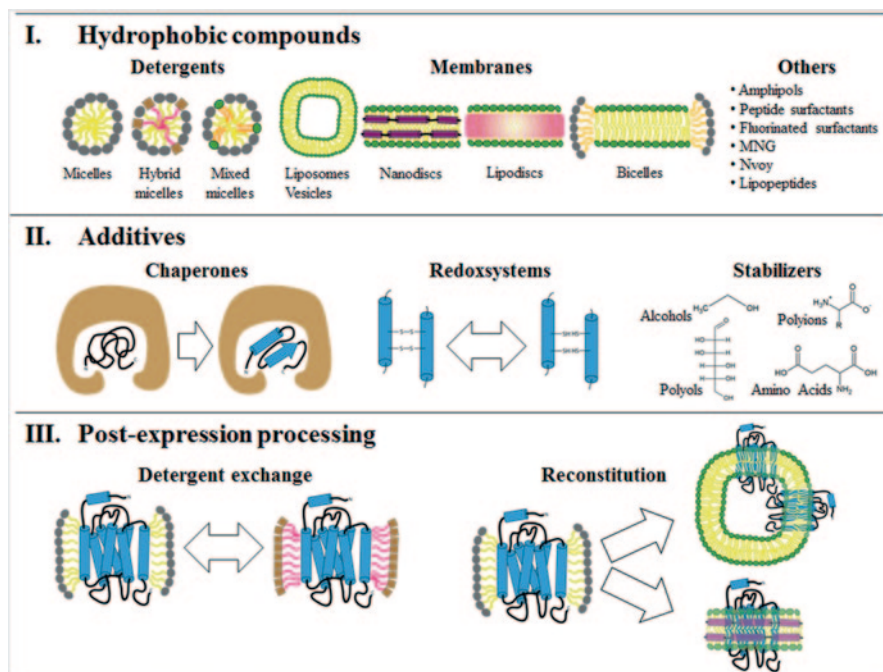


Fig. 2.2 Supplements for the co-translational and posttranslational modification of MP sample quality

(Blesneac et al. 2012). The nature or composition of the selected detergent or detergent mixture can certainly affect the efficiency of solubilization as well as the MP quality in view of folding and stability. Systematic screens for detergent type and concentration are therefore necessary in order to determine optimal conditions for each individually expressed MP (Ishihara et al. 2005; Klammt et al. 2005; Liguori et al. 2008; Martin et al. 2009; Rath et al. 2011).

The open nature of CF reactions has initiated searches for further and new hydrophobic compounds with improved properties that could substitute classical detergents in co-translational MP solubilization (Fig. 2.2). Fluorinated surfactants as well as phospholipid-like surfactants have been considered as mild hydrophobic supplements in D-CF reactions (Park et al. 2011; Blesneac et al. 2012). Amphipols and in particular the charged A8-35 derivative are not well tolerated but might be interesting as solubilizing agents for P-CF-generated MP precipitates. The polyfructose-based uncharged polymer NV10 was claimed to be beneficial in supporting the solubilization of several class B GPCRs (Klammt et al. 2011). Similar positive effects on GPCR solubilization as described for Brij detergents have been observed with peptide surfactants as D-CF supplements (Wang et al. 2011; Corin et al. 2011; Table 2.3). However, curdlan synthase was inactive if D-CF expressed in the presence of Brij58, but active if the detergent was replaced by peptide surfactants (Periasamy et al. 2013). Other compounds such as the recently described

maltose-neopentyl glycol amphiphiles might be considered in future as well (Chae et al. 2010). Although the general availability of several compounds is still somehow limited, it is evident that the variety for designing artificial hydrophobic environments in D-CF reactions is rapidly increasing.

2.6 Designing Protein/Membrane Complexes: L-CF Expression

CF extracts are almost devoid of membranes, although some residual small vesicles originating from the cell membranes might be present after S30 preparation. More complete removal of membrane fragments can be achieved by S100 (centrifugation at $100,000\times g$) extract preparation (Berrier et al. 2011). Instead of complex cell membranes, vesicles containing selected lipid compositions can be prepared *in vitro* and provided into the CF reactions. Lipids are, in contrast to detergents, much better tolerated by CF systems and mostly high final concentrations are possible (Kalmbach et al. 2007; Hovijitra et al. 2009; Roos et al. 2013; Umakoshi et al. 2009). This L-CF (lipid membrane-based) expression mode can thus facilitate the co-translational association of the expressed MPs with supplied bilayers provided as liposomes, as bicelles in combination with specific detergents or as planar membrane discs (Fig. 2.2). The L-CF mode is an excellent approach in order to evaluate lipid effects on the quality and activity of MPs (Table 2.4). Subsequent characterization of L-CF-generated MP samples can furthermore be performed in the natural context of membranes. The co-translational insertion might direct the synthesized MPs in a unidirectional inside-out orientation into the supplied membranes. This could be shown with connexins as example and thus more uniform samples can be generated if compared with conventional posttranslational reconstitution approaches (Moritani et al. 2010).

It should be realized that by selecting the L-CF expression mode, translocation problems of the expressed MP can become an issue again. For many MPs, complex translocation machineries are essential for their proper membrane insertion *in vivo* (Shao and Hedge 2011). The knowledge of translocation mechanisms of MPs in artificial L-CF systems is still at the very beginning. However, it is already evident that the dependency of MP insertion on translocation systems might not be as strict as *in vivo*. The membrane insertions of the CF-expressed channel MscL and of the MtlA permease were independent of the YidC insertase and of the SecYEG complex, respectively (Berrier et al. 2011; Nishiyama et al. 2006). Characteristics that could facilitate membrane insertion might be that (1) the provided membranes are empty, (2) the membrane concentration in CF reaction can be high, and (3) lipid compositions of the supplied membranes can specifically be modulated, e.g., increase of anionic lipids can improve MP insertion efficiencies (Roos et al. 2012). It is generally advisable to perform a lipid screening with individual MPs in order to determine the appropriate composition of supplied membranes in view of lipid charge, length, and flexibility. If translocation systems appear to be mandatory,

Table 2.4 Case studies of L-CF-expressed MPs

Protein ^a	Size [kDa] (TMH ^b)	Type/assay ^c	System ^d yield ^e	Lipid ^f	Reference
<i>Liposomes/Vesicles</i>					
VDAC	31 (19 ⁸)	Channel/+	cE	Thylakoid vesicles	Liguori et al. (2008)
Ec-AqpZ	24 (6)	Channel/+	iE [2]	DOPC, POPG/E	Hovijitra et al. (2009)
OEP24	24 (7 ⁸)	Channel/+	cE [2]	DOPC/PE/DMPA/cholesterol, Soybean asolectin	Liguori et al. (2010)
<i>P. sativum</i>					
Ec-MscL	15 (2)	Channel/+	cE	Soybean asolectin	Berrier et al. (2011)
Ec-OmpA	37 (8 ⁸)	Pore	PURE [1]	<i>E. coli</i> INV	Kuruma et al. (2005)
PulD	68	Pore/+	cE [1]	Soybean asolectin	Guilvout et al. (2008)
<i>K. oxytoca</i>					
α -Hemolysin	33	Pore/+	iE	POPC	Chalmeau et al. (2011)
<i>S. aureus</i>					
Connexin-43	43 (4)	Pore/+	PURE [1]	DOPC/G	Moritani et al. (2010)
Ec-MtlA	70 (6)	Permease	PURE [1]	<i>E. coli</i> INV	Kuruma et al. (2005)
MtlA,	637 (6)	Permease	iE [2]	<i>E. coli</i> INV	Wuu and Swartz (2008)
Ec-TetA	41 (12)	Transporter/+			
Ec-MtlA	70 (6)	Permease	iE	PL + DAG	Nishiyama et al. (2006)
BR	262 (7)	H ⁺ -Pump/+	iE [1]	DOPC	Kalmbach et al. (2007)
Ec-FtsQ	31	Enzyme	PURE [1]	<i>E. coli</i> INV	Kuruma et al. (2005)
Δ gp91	63 (6)	Enzyme/+	cE	MS: Spinach thylacoids	Marques et al. (2007)
hSCD1	40 (4)	Enzyme/+	cWG [2]	MS: Soybean tissue	Goren and Fox (2008)
TbSL1-4	~30 (6)	Enzyme/+	cWG [2]	Soybean lecithin	Sevova et al. (2010)
<i>T. brucei</i>					
Ec-MraY	40 (10)	Enzyme/+	iE [3]	Soybean PC, PL	Ma et al. (2011)
Bs-MraY					
Cyt B5	17 (1)	Enzyme	cWG	DOPC/egg PC	Nomura et al. (2008)
F ₀ subunit of ATP synthase	26	Enzyme	cE	Soybean PC	Kuruma et al. (2010)
3 L-Pf3	5 (1)	Coat protein	cE	<i>E. coli</i> INV + DAG	Kawashima et al. (2008)

Table 2.4 (continued)

Protein ^a	Size [kDa] (TMH ^b)	Type/assay ^c	System ^d yield ^e	Lipid ^f	Reference
hERG,	24 (6)	Channel/+	cE [1]	PC/DMPE	Yildiz et al. (2012)
Ec-Cyt-b03 ubiquinol oxid.	35	Enzyme/+			
CrdS	75 (7)	Enzyme/+	iWG [2]	DOPG/E, <i>E. coli</i> lipids	Periasamy et al. (2013)
<i>Agrobacterium</i>					
Proteorhodopsin	27 (7)	H ⁺ -Pump/+	iE [2]	DMPC/PA	Mörs et al. (2013)
Bacteriorhodopsin	28 (7)	H ⁺ -Pump/+	cE [2]	DMPC	Cappuccio et al. (2008)
Bs-DesK	40 (4–5)	Histidine kinase/+	cE [3]	DOPC	Martin et al. (2009)
40 mammalian MPs	32–130 (1–14)	Diverse/+	cWG [1]	Soybean asolectin	Nozawa et al. (2011)
Connexin-43	43 (4)	Pore/+	RRL	Egg PC	Kaneda et al. (2009)
<i>Lipid + Detergent</i>					
AtPPT1	~30	Transporter/+	cWG	Soybean asolectin + Brij35	Nozawa et al. (2007)
Bacteriorhodopsin	28 (7)	H ⁺ -Pump/+	iE [2]	Egg PC + CHAPS/Cholesterol/ Digitonin	Shimono et al. (2009)
ATP synthase (a + c subunits)	23 + 11	Complex/+	cE [1]	DHPC, DMPC	Uhlemann et al. (2012)
<i>Nanodiscs/lipoprotein complexes</i>					
Ec-EmrE	12 (4)	Transporter/+	cE [2]	apoA1 (DMPC)	Katzen et al. (2008)
Bacteriorhodopsin	28 (7)	H ⁺ -Pump/+	cWG		
32 diverse MPs			RRL		
hTM-ErbB3	5 (1)	Receptor	iE [3]	MSP1 (DMPC/G/POPC, DOPG)	Lyukmanova et al. (2012)
VSD-KvaP	28 (7)	Channel			
Bacteriorhodopsin		H ⁺ -Pump/+			
<i>E. sibiricum</i>					
hβ2-AR-T4L	63 (7)	GPCR/+	cE	apoA1 (DMPC)	Yang et al. (2011)
hETB, hETA	–	GPCR/+	iE [2]	MSP1E3 (DMPC/G)	Proverbio et al. (2013)
npSR11 (<i>N. pharaonis</i>)	26 (6)	Receptor	iE	apoA1 (DMPC)	Isaksson et al. (2012)
pR-AND4	32 (7)	H ⁺ -Pump/+			

Table 2.4 (continued)

Protein ^a	Size [kDa] (TMH ^b)	Type/assay ^c	System ^d yield ^e	Lipid ^f	Reference
Ec-MraY	40 (10)	Enzyme/+	iE [3]	MSPIE3 (DMPC/G, POPC/G, DOPC/G/E, TL, PL)	Roos et al. 2012
Proteorhodopsin	27 (7)	H ⁺ -Pump/+			
Proteorhodopsin	27 (7)	H ⁺ -Pump/+	iE [2]	MSPI (DMPC, POPC)	Mörs et al. (2013)
Bacteriorhodopsin	28 (7)	H ⁺ -Pump/+	cE	–	Cappuccio et al. (2009)
<i>H. halobium</i>					
ADRB2, DRD1, NK1R	(7)	GPCR/+	cE	Δ49A1 (DMPC)	Gao et al. (2012)

MS microsomes, *IMV* inner membrane vesicles, *OMV* outer membrane vesicles, *INV* inverted vesicles, *Brij35* polyoxyethylene-(23)-lauryl-ether, *CHAPS* 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate, *DAG* diacylglycerol, *DHPC* 1,2-diheptanoyl-sn-glycero-3-phosphocholine, *DMPA* 1,2-ditetradecanoyl-sn-glycero-3-phosphate, *DMPC* 1,2-dimyristoyl-sn-glycero-3-phosphocholine, *DMPE* 1,2-ditetradecanoyl-sn-glycero-3-phosphoethanolamine, *DMPG* 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol), *DOPA* 1,2-dioleoyl-sn-glycero-3-phosphate, *DOPC* 1,2-dioleoyl-sn-glycero-3-phosphocholine, *DOPE* 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine, *DOPG* 1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol), *PC* L- α -phosphatidylcholine, *PE* L- α -phosphatidylethanolamine, *PL* polar lipid extract, *POPC* 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine, *POPE* 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine, *POPG* 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol), *TL* total lipid extract

^a If documented, the origin of proteins is given in italics; *h* human, *m* murine, *r* rat, *Ec E. coli*, *Bs Bacillus subtilis*

^b Size and number of transmembrane helices or β -sheets (B), if documented

^c +: Quality analyzed by structural evaluation or functionality

^d iE individual *E. coli* extracts, cE commercial *E. coli* extracts, iWG individual wheat germ extracts, cWG commercial wheat germ extracts, PURE protein synthesis using recombinant elements, RRL rabbit reticulocyte extract

^e Approximate yields per 1 ml RM if documented in the corresponding references. 1: ≤ 0.1 mg/ml; 2: 0.1–1 mg/ml; 3: > 1 mg/ml

^f Lipids or lipid mixtures used for co-translational reconstitution

vesicles isolated out of complex cell membranes and containing natural translocation machineries could be considered as supplements as well (Stech et al. 2012).

L-CF expression in the presence of nanodiscs appears to be in particular promising in order to obtain soluble and functionally active MP/membrane complexes (Cappuccio et al. 2009; Lyukmanova et al. 2012; Proverbio et al. 2013; Roos et al. 2012). Specific advantages might be that nanodisc membranes are accessible from both sides and it could be speculated that inhomogeneities of the membrane/membrane scaffold protein interface may provide additional entry sides for MP integration. MP/nanodisc complexes are highly soluble and can be used for a variety of applications such as surface plasmon resonance (SPR) measurements (Proverbio et al. 2013). Nanodiscs provide furthermore a membrane compartment that is stable in size and topology. In contrast, liposomes supplied to CF reactions show excessive fusion and increase in size, resulting into their almost quantitative precipitation during the reaction (Barrier et al. 2011; Roos et al. 2013). An interesting modification by using the CECF configuration is the initial supply of a lipid/detergent mixture in the RM, whereas the FM is devoid of any detergent. Freshly translated MPs can therefore first physically associate with the provided lipomicelles and become increasingly trapped into membranes that slowly formed by the continuously decreased detergent concentration in the RM. This method was successfully applied for bacteriorhodopsin with combinations of steroid detergents and phosphatidylcholine lipids (Shimono et al. 2009). In a similar approach, the plant solute transporter AtPPT1 was functionally reconstituted by using a mixture of Brij35 and soybean asolectin lipids (Nozawa et al. 2007).

Besides hydrophobic environments, a variety of further additives could be beneficial for the production of high quality MP samples (Fig. 2.2). Chemical chaperones such as sugars, alcohols, or polyions are tolerated by CF systems and synergies of several compounds could be determined in correlated concentration screens (Kai et al. 2013). Such additives might be in particular beneficial for supporting the folding of larger soluble domains of MPs.

2.7 Handling the Toolbox of CF Expression: Strategies for Protein Quality Optimization

During the last decade when CF expression as a new platform for MP production was emerging, the three basic modes P-CF, D-CF, and L-CF have been employed in more or less comparable frequencies. The variety of supplements useful for MP quality optimization is rapidly expanding and defining specific conditions for the production of sufficient MP quantities that are homogenous, functionally folded, and stable is generally the key issue in CF expression protocols. The systematic screening of (1) expression modes, (2) type and concentration of hydrophobic compounds, (3) additives such as stabilizer or chaperones and (4) post-expression processing procedures generate an array of MP samples that have to be analyzed (Fig. 2.2). A strategic plan for MP quality control is therefore crucial and the first evaluation

of solubility, sample homogeneity, protein integrity, stability, or specific activity should be implemented as soon as possible, at best already in the crude reactions. The above mentioned GFP fusion approach can, e.g., provide already a first short list of compounds and compound combinations giving an efficient solubilization of the MP. Enzymatic reactions or binding of labeled ligands might be considered as well in specific cases (Gourdon et al. 2008; Kalmbach et al. 2007). Solubility and integrity can be assayed with immunoblotting by taking advantage of terminal tags. Rare occurring MP fragmentation generated by premature translational termination or proteolytic degradation could be addressed by using synthetic genes and by screening protease inhibitors. In the case of some GPCRs, ligand-binding assays by SPR measurement can be performed with crude RMs containing the expressed MP (Proverbio et al. 2013). Feedback from first quality evaluations should then be considered for re-optimization, fine-tuning of compound concentrations, and for analyzing cocktails of beneficial compounds for synergistic effects. In many cases, tremendous variations in the quality of MP samples produced at different CF conditions have been observed. A striking example is the *MraY* translocase, a membrane-embedded enzyme responsible for lipid-I precursor formation in the bacterial cell wall biosynthesis pathway. The *Bacillus subtilis* *MraY* can be functionally synthesized in a large variety of CF conditions implementing detergents and lipids. In contrast, the *E. coli* *MraY* enzyme was only functional if L-CF synthesized in the presence of nanodiscs containing anionic lipids (Roos et al. 2012; Ma et al. 2011).

If a short list of few promising reaction conditions has been determined, more time-consuming assays using purified samples such as size exclusion chromatography, multi-angle light scattering, or circular dichroism spectroscopy can be implemented in order to analyze homogeneity, folding, and oligomeric states of the MP. Functionality of MPs is often more difficult to analyze and currently available case studies are compiled in Tables 2.2–2.4. As some examples, ligand binding of GPCRs and transporters were shown by radioassays (Gao et al. 2012; Ishihara et al. 2005; Sansuk et al. 2008; Yang et al. 2011), fluorescence anisotropy measurement (Junge et al. 2010), SPR (Kaiser et al. 2008; Proverbio et al. 2013), or thermophoresis (Corin et al. 2011). Functional samples of channels and transporters after co-translationally or posttranslationally reconstitution into lipid bilayers were obtained from *MscL* (Berrier et al. 2004, 2011), *EmrE* (Elbaz et al. 2004), *TetA* (Wuu and Swartz 2008), *PorA* and *PorH* (Rath et al. 2011), eukaryotic organic ion transporters (Keller et al. 2008, 2011), as well as from aquaporins (Hovijitra et al. 2009; Kai et al. 2010; Müller-Lucks et al. 2013).

2.8 Perspectives for Structural Approaches

The possibility to produce pure and concentrated samples of even very difficult MPs in a short time by CF expression is certainly of major interest for structural studies. Efficient incorporation of selenomethionine is of value for X-ray crystallography. In addition, MPs could already co-translationally be stabilized by inhibitors or other ligands. X-ray structures of CF-expressed MPs are still limited to the multidrug transporter *EmrE* (Chen et al. 2007) and *Acetabularia* proteorhodopsin

(Wada et al. 2011) synthesized in a mixture of detergents and lipids. Crystallization was furthermore successful with the CF-expressed human voltage-gated anion channel VDAC1 (Nguyen et al. 2010; Deniaud et al. 2010). Despite some success, the still relatively low number of X-ray structures derived from CF-expressed MPs indicates that reaction conditions obviously may often not have been optimal in order to obtain crystallization grade MP samples. Systematic case studies are needed in order to identify the key parameters to be considered by choosing CF expression approaches. Sample homogeneity might be optimized by (1) intensive screening of reaction modes and hydrophobic supplements, (2) providing detergents in cocktails with some stabilizing lipids, (3) adjusting redox conditions for proper disulfide bridge formation, and (4) fine-tuning of solubilization conditions of P-CF samples.

The efficient and cost-effective labeling opportunities of CF expression are excellent prerequisites for structural approaches by NMR spectroscopy and a variety of sophisticated labeling tools and schemes have been developed and are already established standards (Klammt et al. 2012; Reckel et al. 2008; Ozawa et al. 2006). Nevertheless, liquid-state NMR of MPs is still a challenging task. MPs in micellar as well as in lipid environment are prone to signal broadening due to the large size and slow rotational tumbling. Furthermore, amino acids in α -helical structures tend to display narrow-range chemical shifts, resulting in severe peak overlaps. Numerous liquid-state NMR structures of P-CF (Klammt et al. 2012; Maslennikov et al. 2010; Sobhanifar et al. 2010) or D-CF (Reckel et al. 2008) expressed MPs have been reported (see also chapter by L. Catoire and D. Warschawski in this volume). For solid-state NMR, P-CF precipitates, samples posttranslationally reconstituted into liposomes, or L-CF samples by additions of liposomes or nanodiscs can be used (see also chapter by L. Catoire and D. Warschawski in this volume). Examples of successfully analyzed MPs are the mechanosensitive channel MscL (Abdine et al. 2010) and the multidrug transporter EmrE (Lehner et al. 2008). In particular, L-CF expression in the presence of nanodiscs can become attractive as shown in initial studies with photoorhodopsin (Mörs et al. 2013).

2.9 Conclusions

Besides exploring new approaches in particular for drug screening or single-molecule approaches, some major current challenges are the scale-up of CF reactions to industrial dimensions, the streamlined determination of expression conditions for crystallization grade MP samples, and the production of larger assemblies and MP complexes. Initial milestones have already been achieved and give promising perspectives. The manufacturing of multigram to kilogram scales in 100 L or even higher CF reaction volumes appears to become feasible (Zawada et al. 2011). Providing complex and more elaborated hydrophobic expression environments might be a direction for producing homogeneous samples suitable for crystallization (Wada et al. 2011). Tuning expression and template design might be a prerequisite for successful synthesis and assembly of MP complexes as recently been shown for the 542-kDa ATP synthase complex (Matthies et al. 2011).

Acknowledgments This work was supported by the Collaborative Research Center (SFB) 807 of the German Research Foundation (DFG) and by the German Ministry of Education and Science (BMBF). Erika Orbán was supported by the Alexander von Humboldt Foundation.

References

- Abdine A, Verhoeven MA, Park K-H, Ghazi A, Guittet E, Berrier C, Van Heijenoort C, Warschawski DE (2010) Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J Magn Reson* 204:155–159
- Ahn J-H, Chu H-S, Kim T-W, Oh I-S, Choi C-Y, Hahn G-H, Park C-G, Kim D-M (2005) Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that of plasmid-based reactions. *Biochem Biophys Res Commun* 338:1346–1352
- Ahn J-H, Hwang M-Y, Lee K-H, Choi C-Y, Kim D-M (2007) Use of signal sequences as an in situ removable sequence element to stimulate protein synthesis in cell-free extracts. *Nucleic Acids Res* 35:e21
- Arduengo M, Schenborn E, Hurst R (2007) The role of cell-free rabbit reticulocyte expression systems in functional proteomics. In: Kudlicki WA, Katzen F, Bennett RP (eds) *Cell-free protein expression*, Landes Bioscience, Austin, pp 1–18
- Arslan Yildiz A, Kang C, Sinner E-K (2013) Biomimetic membrane platform containing hERG potassium channel and its application to drug screening. *Analyst* 138:2007–2012
- Basu D, Castellano JM, Thomas N, Mishra RK (2013) Cell-free protein synthesis and purification of human dopamine D2 receptor long isoform. *Biotechnol Prog* 29:601–608
- Bazzacco P, Billon-Denis E, Sharma KS et al (2012) Nonionic homopolymeric amphipols: application to membrane protein folding, cell-free synthesis, and solution nuclear magnetic resonance. *Biochemistry* 51:1416–1430
- Beebe ET, Makino S-I, Nozawa A, Matsubara Y, Frederick RO, Primm JG, Goren MA, Fox BG (2011) Robotic large-scale application of wheat cell-free translation to structural studies including membrane proteins. *N Biotechnol* 28:239–249
- Berrier C, Park K-H, Abes S, Bibonne A, Betton J-M, Ghazi A (2004) Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry* 43:12585–12591
- Berrier C, Guilvout I, Bayan N, Park K-H, Mesneau A, Chami M, Pugsley AP, Ghazi A (2011) Coupled cell-free synthesis and lipid vesicle insertion of a functional oligomeric channel MscL. MscL does not need the insertase YidC for insertion in vitro. *Biochim Biophys Acta* 1808:41–46
- Blesneac I, Ravaud S, Juillan-Binard C, Barret L-A, Zoonens M, Polidori A, Miroux B, Pucci B, Pebay-Peyroula E (2012) Production of UCP1 a membrane protein from the inner mitochondrial membrane using the cell free expression system in the presence of a fluorinated surfactant. *Biochim Biophys Acta* 1818:798–805
- Cappuccio JA, Blanchette CD, Sulchek TA et al (2008) Cell-free co-expression of functional membrane proteins and apolipoprotein, forming soluble nanolipoprotein particles. *Mol Cell Proteomics* 7:2246–2253
- Cappuccio JA, Hinz AK, Kuhn EA et al (2009) Cell-free expression for nanolipoprotein particles: building a high-throughput membrane protein solubility platform. *Methods Mol Biol* 498:273–296
- Chae PS, Rasmussen SGF, Rana RR et al (2010) Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* 7:1003–1008
- Chalmeau J, Monina N, Shin J, Vieu C, Noireaux V (2011) α -Hemolysin pore formation into a supported phospholipid bilayer using cell-free expression. *Biochim Biophys Acta* 1808:271–278

- Chen Y-J, Pornillos O, Lieu S, Ma C, Chen AP, Chang G (2007) X-ray structure of EmrE supports dual topology model. *Proc Natl Acad Sci U S A* 104:18999–19004
- Corin K, Baaske P, Ravel DB et al (2011) Designer lipid-like peptides: a class of detergents for studying functional olfactory receptors using commercial cell-free systems. *PLoS One* 6:e25067. doi:10.1371/journal.pone.0025067
- Craig D, Howell MT, Gibbs CL, Hunt T, Jackson RJ (1992) Plasmid cDNA-directed protein synthesis in a coupled eukaryotic in vitro transcription-translation system. *Nucleic Acids Res* 20:4987–4995
- Deniaud A, Liguori L, Blesneac I, Lenormand JL, Pebay-Peyroula E (2010) Crystallization of the membrane protein hVDAC1 produced in cell-free system. *Biochim Biophys Acta* 1798:1540–1546
- Elbaz Y, Steiner-Mordoch S, Danieli T, Schuldiner S (2004) In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc Natl Acad Sci U S A* 101:1519–1524
- Endo Y, Sawasaki T (2005) Advances in genome-wide protein expression using the wheat germ cell-free system. *Methods Mol Biol* 310:145–167
- Endoh T, Kanai T, Imanaka T (2008) Effective approaches for the production of heterologous proteins using the *Thermococcus kodakaraensis*-based translation system. *J Biotechnol* 133:177–182
- Ezure T, Suzuki T, Shikata M, Ito M, Ando E, Nishimura O, Tsunasawa S (2007) Expression of proteins containing disulfide bonds in an insect cell-free system and confirmation of their arrangements by MALDI-TOF MS. *Proteomics* 7:4424–4434
- Ezure T, Suzuki T, Shikata M, Ito M, Ando E (2010) A cell-free protein synthesis system from insect cells. *Methods Mol Biol* 607:31–42
- Gao T, Petlova J, He W, Huser T, Kudlick W, Voss J, Coleman MA (2012) Characterization of de novo synthesized GPCRs supported in nanolipoprotein discs. *PLoS One* 7:e44911
- Genji T, Nozawa A, Tozawa Y (2010) Efficient production and purification of functional bacteriorhodopsin with a wheat-germ cell-free system and a combination of Fos-choline and CHAPS detergents. *Biochem Biophys Res Commun* 400:638–642
- Goerke AR, Swartz JR (2008) Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechnol Bioeng* 99:351–367
- Goren MA, Fox BG (2008) Wheat germ cell-free translation, purification, and assembly of a functional human stearyl-CoA desaturase complex. *Protein Expr Purif* 62:171–178
- Gourdon P, Alfredsson A, Pedersen A et al (2008) Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump. *Protein Expr Purif* 58:103–113
- Guarino C, DeLisa MP (2012) A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins. *Glycobiology* 22:596–601
- Guilvout I, Chami M, Berrier C, Ghazi A, Engel A, Pugsley AP, Bayan N (2008) In vitro multimerization and membrane insertion of bacterial outer membrane secretin PulD. *J Mol Biol* 382:13–23
- Haberstock S, Roos C, Hoevels Y, Dötsch V, Schnapp G, Pautsch A, Bernhard F (2012) A systematic approach to increase the efficiency of membrane protein production in cell-free expression systems. *Protein Expr Purif* 82:308–316
- Hovijitra NT, Wu JJ, Peaker B, Swartz JR (2009) Cell-free synthesis of functional aquaporin Z in synthetic liposomes. *Biotechnol Bioeng* 104:40–49
- Isaksson L, Enberg J, Neutze R, Göran Karlsson B, Pedersen A (2012) Expression screening of membrane proteins with cell-free protein synthesis. *Protein Expr Purif* 82:218–225
- Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M, Yokoyama S (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr Purif* 41:27–37
- Junge F, Luh LM, Proverbio D, Schäfer B, Abele R, Beyermann M, Dötsch V, Bernhard F (2010) Modulation of G-protein coupled receptor sample quality by modified cell-free expression protocols: a case study of the human endothelin A receptor. *J Struct Biol* 172:94–106

- Junge F, Haberstock S, Roos C, Stefer S, Proverbio D, Dötsch V, Bernhard F (2011) Advances in cell-free protein synthesis for the functional and structural analysis of membrane proteins. *N Biotechnol* 28:262–271
- Kai L, Kaldenhoff R, Lian J, Zhu X, Dötsch V, Bernhard F, Cen P, Xu Z (2010) Preparative scale production of functional mouse aquaporin 4 using different cell-free expression modes. *PLoS One* 5:e12972
- Kai L, Dötsch V, Kaldenhoff R, Bernhard F (2013) Artificial environments for the co-translational stabilization of cell-free expressed proteins. *PLoS One* 8:e56637
- Kaiser L, Graveland-Bikker J, Steuerwald D, Vanberghem M, Herlihy K, Zhang S (2008) Efficient cell-free production of olfactory receptors: detergent optimization, structure, and ligand binding analyses. *Proc Natl Acad Sci U S A* 105:15726–15731
- Kalmbach R, Chizhov I, Schumacher MC, Friedrich T, Bamberg E, Engelhard M (2007) Functional cell-free synthesis of a seven helix membrane protein: in situ insertion of bacteriorhodopsin into liposomes. *J Mol Biol* 371:639–648
- Kaneda M, Nomura SM, Ichinose S, Kondo S, Nakahama K, Akiyoshi K, Morita I (2009) Direct formation of proteo-liposomes by in vitro synthesis and cellular cytosolic delivery with connexin-expressing liposomes. *Biomaterials* 30:3971–3977
- Katzen F, Kudlicki W (2006) Efficient generation of insect-based cell-free translation extracts active in glycosylation and signal sequence processing. *J Biotechnol* 125:194–197
- Katzen F, Chang G, Kudlicki W (2005) The past, present and future of cell-free protein synthesis. *Trends Biotechnol* 23:150–156
- Katzen F, Fletcher JE, Yang J-P et al (2008) Insertion of membrane proteins into discoidal membranes using a cell-free protein expression approach. *J Proteome Res* 7:3535–3542
- Kawashima Y, Miyazaki E, Müller M, Tokuda H, Nishiyama K (2008) Diacylglycerol specifically blocks spontaneous integration of membrane proteins and allows detection of a factor-assisted integration. *J Biol Chem* 283:24489–24496
- Keller T, Schwarz D, Bernhard F, Dötsch V, Hunte C, Gorboulev V, Koepsell H (2008) Cell free expression and functional reconstitution of eukaryotic drug transporters. *Biochemistry* 47:4552–4564
- Keller T, Egenberger B, Gorboulev V et al (2011) The large extracellular loop of organic cation transporter 1 influences substrate affinity and is pivotal for oligomerization. *J Biol Chem* 286:37874–37886
- Khabibullina NF, Liukmanova EN, Kopeina GS, Shenkarev ZO, Arsen'ev AS, Dolgikh DA, Kirpichnikov MP (2010) [The development and optimization of coupled cell-free expression system for production of the transmembrane domain of the receptor tyrosine kinase ErbB3]. *Bioorg Khim* 36:654–660
- Kigawa T, Yokoyama S (1991) A continuous cell-free protein synthesis system for coupled transcription-translation. *J Biochem* 110:166–168
- Kim D-M, Swartz JR (2004) Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of *Escherichia coli*. *Biotechnol Bioeng* 85:122–129
- Klammt C, Löh F, Schäfer B, Haase W, Dötsch V, Rüterjans H, Glaubitz C, Bernhard F (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* 271:568–580
- Klammt C, Schwarz D, Fendler K, Haase W, Dötsch V, Bernhard F (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J* 272:6024–6038
- Klammt C, Schwarz D, Eifler N, Engel A, Piehler J, Haase W, Hahn S, Dötsch V, Bernhard F (2007) Reprint of “cell-free production of G protein-coupled receptors for functional and structural studies” [*J Struct Biol* 158 (2007) 482–493]. *J Struct Biol* 159:194–205
- Klammt C, Perrin MH, Maslennikov I, Renault L, Krupa M, Kwiatkowski W, Stahlberg H, Vale W, Choe S (2011) Polymer-based cell-free expression of ligand-binding family B G-protein coupled receptors without detergents. *Protein Sci* 20:1030–1041
- Klammt C, Maslennikov I, Bayrhuber M et al (2012) Facile backbone structure determination of human membrane proteins by NMR spectroscopy. *Nat Methods* 9:834–839

- Kovtun O, Mureev S, Jung W, Kubala MH, Johnston W, Alexandrov K (2011) Leishmania cell-free protein expression system. *Methods* 55:58–64
- Kralicek AV, Radjainia M, Mohamad Ali NAB, Carraher C, Newcomb RD, Mitra AK (2011) A PCR-directed cell-free approach to optimize protein expression using diverse fusion tags. *Protein Expr Purif* 80:117–124
- Kuruma Y, Nishiyama K-I, Shimizu Y, Müller M, Ueda T (2005) Development of a minimal cell-free translation system for the synthesis of presecretory and integral membrane proteins. *Biotechnol Prog* 21:1243–1251
- Kuruma Y, Suzuki T, Ueda T (2010) Production of multi-subunit complexes on liposome through an *E. coli* cell-free expression system. *Methods Mol Biol* 607:161–171
- Langlais C, Guillaume B, Wermke N, Scheuermann T, Ebert L, LaBaer J, Korn B (2007) A systematic approach for testing expression of human full-length proteins in cell-free expression systems. *BMC Biotechnol* 7:64
- Lehner I, Basting D, Meyer B, Haase W, Manolikas T, Kaiser C, Karas M, Glaubitc C (2008) The key residue for substrate transport (Glu14) in the EmrE dimer is asymmetric. *J Biol Chem* 283:3281–3288
- Liguori L, Marques B, Lenormand J-L (2008) A bacterial cell-free expression system to produce membrane proteins and proteoliposomes: from cDNA to functional assay. *Curr Protoc Protein Sci*, Chapter 5: Unit 5.22
- Liguori L, Blesneac I, Madern D, Vivaudou M, Lenormand J-L (2010) Single-step production of functional OEP24 proteoliposomes. *Protein Expr Purif* 69:106–111
- Lyukmanova EN, Shenkarev ZO, Khabibullina NF et al (2012) Lipid-protein nanodiscs for cell-free production of integral membrane proteins in a soluble and folded state: comparison with detergent micelles, bicelles and liposomes. *Biochim Biophys Acta* 1818:349–358
- Ma Y, Münch D, Schneider T, Sahl H-G, Bouhss A, Ghoshdastider U, Wang J, Dötsch V, Wang X, Bernhard F (2011) Preparative scale cell-free production and quality optimization of *MraY* homologues in different expression modes. *J Biol Chem* 286:38844–38853
- Madin K, Sawasaki T, Ogasawara T, Endo Y (2000) A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc Natl Acad Sci U S A* 97:559–564
- Marques B, Liguori L, Pacllet M-H, Villegas-Méndez A, Rothe R, Morel F, Lenormand J-L (2007) Liposome-mediated cellular delivery of active gp91(phox). *PLoS One* 2:e856
- Martín M, Albanesi D, Alzari PM, De Mendoza D (2009) Functional in vitro assembly of the integral membrane bacterial thermosensor DesK. *Protein Expr Purif* 66:39–45
- Maslennikov I, Klammt C, Hwang E, Kefala G, Okamura M, Esquivies L, Mörs K, Glaubitc C, Kwiatkowski W, Jeon YH, Choe S (2010) Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis. *Proc Natl Acad Sci U S A* 107:10902–10907
- Matthies D, Haberstock S, Joos F, Dötsch V, Vonck J, Bernhard F, Meier T (2011) Cell-free expression and assembly of ATP synthase. *J Mol Biol* 413:593–603
- Mikami S, Masutani M, Sonenberg N, Yokoyama S, Imataka H (2006) An efficient mammalian cell-free translation system supplemented with translation factors. *Protein Expr Purif* 46:348–357
- Mikami S, Kobayashi T, Masutani M, Yokoyama S, Imataka H (2008) A human cell-derived in vitro coupled transcription/translation system optimized for production of recombinant proteins. *Protein Expr Purif* 62:190–198
- Miniev KS, Khabibullina NF, Lyukmanova EN, Dolgikh DA, Kirpichnikov MP, Arseniev AS (2011) Spatial structure and dimer—monomer equilibrium of the ErbB3 transmembrane domain in DPC micelles. *Biochim Biophys Acta* 1808:2081–2088
- Miot M, Betton J-M (2011) Reconstitution of the Cpx signaling system from cell-free synthesized proteins. *N Biotechnol* 28:277–281. doi:10.1016/j.nbt.2010.06.012
- Moritani Y, Nomura SM, Morita I, Akiyoshi K (2010) Direct integration of cell-free-synthesized connexin-43 into liposomes and hemichannel formation. *FEBS J* 277:3343–3352

- Mörs K, Roos C, Scholz F, Wachtveitl J, Dötsch V, Bernhard F, Glaubitc C (2013) Modified lipid and protein dynamics in nanodiscs. *Biochim Biophys Acta* 1828:1222–1229
- Müller-Lucks A, Bock S, Wu B, Beitz E (2012) Fluorescent in situ folding control for rapid optimization of cell-free membrane protein synthesis. *PLoS One* 7:e42186
- Müller-Lucks A, Gena P, Frascaria D, Altamura N, Svelto M, Beitz E, Calamita G (2013) Preparative scale production and functional reconstitution of a human aquaglyceroporin (AQP3) using a cell free expression system. *N Biotechnol*. doi:10.1016/j.nbt.2013.03.007
- Nakashima N, Tamura T (2004) Cell-free protein synthesis using cell extract of *Pseudomonas fluorescens* and CspA promoter. *Biochem Biophys Res Commun* 319:671–676
- Nguyen TA, Lieu SS, Chang G (2010) An *Escherichia coli*-based cell-free system for large-scale production of functional mammalian membrane proteins suitable for X-ray crystallography. *J Mol Microbiol Biotechnol* 18:85–91
- Nishiyama K, Ikegami A, Moser M, Schiltz E, Tokuda H, Müller M (2006) A derivative of lipid A is involved in signal recognition particle/SecYEG-dependent and -independent membrane integrations. *J Biol Chem* 281:35667–35676
- Nomura S-IM, Kondoh S, Asayama W, Asada A, Nishikawa S, Akiyoshi K (2008) Direct preparation of giant proteo-liposomes by in vitro membrane protein synthesis. *J Biotechnol* 133:190–195
- Nozawa A, Nanamiya H, Miyata T, Linka N, Endo Y, Weber APM, Tozawa Y (2007) A cell-free translation and proteoliposome reconstitution system for functional analysis of plant solute transporters. *Plant Cell Physiol* 48:1815–1820
- Nozawa A, Ogasawara T, Matsunaga S, Iwasaki T, Sawasaki T, Endo Y (2011) Production and partial purification of membrane proteins using a liposome-supplemented wheat cell-free translation system. *BMC Biotechnol* 11:35
- Ozawa K, Wu PSC, Dixon NE, Otting G (2006) N-Labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes. *FEBS J* 273:4154–4159
- Park K-H, Billon-Denis E, Dahmane T, Lebaupain F, Pucci B, Breyton C, Zito F (2011) In the cauldron of cell-free synthesis of membrane proteins: playing with new surfactants. *N Biotechnol* 28:255–261
- Pédelaçq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* 24:79–88
- Pedersen A, Wallgren M, Karlsson BG, Gröbner G (2011) Expression and purification of full-length anti-apoptotic Bcl-2 using cell-free protein synthesis. *Protein Expr Purif* 77:220–223
- Periasamy A, Shadiac N, Amalraj A, Garajová S, Nagarajan Y, Waters S, Mertens HDT, Hrmova M (2013) Cell-free protein synthesis of membrane (1,3)- β -d-glucan (curdlan) synthase: co-translational insertion in liposomes and reconstitution in nanodiscs. *Biochim Biophys Acta* 1828:743–757
- Proverbio D, Roos C, Beyermann M, Orban E, Dötsch V, Bernhard F (2013) Functional properties of cell-free expressed human endothelin A and endothelin B receptors in artificial membrane environments. *Biochim Biophys Acta*. doi:10.1016/j.bbamem.2013.05.031
- Rath P, Demange P, Saurel O, Tropis M, Daffé M, Dötsch V, Ghazi A, Bernhard F, Milon A (2011) Functional expression of the PorAH channel from *Corynebacterium glutamicum* in cell-free expression systems: implications for the role of the naturally occurring mycolic acid modification. *J Biol Chem* 286:32525–32532
- Reckel S, Sobhanifar S, Schneider B, Junge F, Schwarz D, Durst F, Löhr F, Güntert P, Bernhard F, Dötsch V (2008) Transmembrane segment enhanced labeling as a tool for the backbone assignment of alpha-helical membrane proteins. *Proc Natl Acad Sci U S A* 105:8262–8267
- Roos C, Zocher M, Müller D et al (2012) Characterization of co-translationally formed nanodisc complexes with small multidrug transporters, proteorhodopsin and with the *E. coli* MraY translocase. *Biochim Biophys Acta* 1818:3098–3106
- Roos C, Kai L, Proverbio D, Ghoshdastider U, Filipek S, Dötsch V, Bernhard F (2013) Co-translational association of cell-free expressed membrane proteins with supplied lipid bilayers. *Mol Membr Biol* 30:75–89

- Sansuk K, Balog CIA, Van der Does AM, Booth R, De Grip WJ, Deelder AM, Bakker RA, Leurs R, Hensbergen PJ (2008) GPCR proteomics: mass spectrometric and functional analysis of histamine H1 receptor after baculovirus-driven and in vitro cell free expression. *J Proteome Res* 7:621–629
- Savage DF, Anderson CL, Robles-Colmenares Y, Newby ZE, Stroud RM (2007) Cell-free complements in vivo expression of the *E. coli* membrane proteome. *Protein Sci* 16:966–976
- Sawasaki T, Morishita R, Gouda MD, Endo Y (2007) Methods for high-throughput materialization of genetic information based on wheat germ cell-free expression system. *Methods Mol Biol* 375:95–106
- Schneider B, Junge F, Shirokov VA, Durst F, Schwarz D, Dötsch V, Bernhard F (2010) Membrane protein expression in cell-free systems. *Methods Mol Biol* 601:165–186
- Schwarz D, Junge F, Durst F, Frölich N, Schneider B, Reckel S, Sobhanifar S, Dötsch V, Bernhard F (2007) Preparative scale expression of membrane proteins in *Escherichia coli*-based continuous exchange cell-free systems. *Nat Protoc* 2:2945–2957
- Schwarz D, Dötsch V, Bernhard F (2008) Production of membrane proteins using cell-free expression systems. *Proteomics* 8:3933–3946
- Schwarz D, Daley D, Beckhaus T, Dötsch V, Bernhard F (2010) Cell-free expression profiling of *E. coli* inner membrane proteins. *Proteomics* 10:1762–1779
- Sevova ES, Goren MA, Schwartz KJ, Hsu F-F, Turk J, Fox BG, Bangs JD (2010) Cell-free synthesis and functional characterization of sphingolipid synthases from parasitic trypanosomatid protozoa. *J Biol Chem* 285:20580–20587
- Shao S, Hegde RS (2011) Membrane protein insertion at the endoplasmic reticulum. *Annu Rev Cell Dev Biol* 27:25–56
- Shenkarev ZO, Lyukmanova EN, Butenko IO, Petrovskaya LE, Paramonov AS, Shulepko MA, Nekrasova OV, Kirpichnikov MP, Arseniev AS (2013) Lipid-protein nanodiscs promote in vitro folding of transmembrane domains of multi-helical and multimeric membrane proteins. *Biochim Biophys Acta* 1828:776–784
- Shimada Y, Wang Z-Y, Mochizuki Y, Kobayashi M, Nozawa T (2004) Functional expression and characterization of a bacterial light-harvesting membrane protein in *Escherichia coli* and cell-free synthesis systems. *Biosci Biotechnol Biochem* 68:1942–1948
- Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. *Nat Biotechnol* 19:751–755
- Shimono K, Goto M, Kikukawa T, Miyauchi S, Shirouzu M, Kamo N, Yokoyama S (2009) Production of functional bacteriorhodopsin by an *Escherichia coli* cell-free protein synthesis system supplemented with steroid detergent and lipid. *Protein Sci* 18:2160–2171
- Shin J, Noireaux V (2010) Efficient cell-free expression with the endogenous *E. Coli* RNA polymerase and sigma factor 70. *J Biol Eng* 4:8
- Sobhanifar S, Reckel S, Junge F, Schwarz D, Kai L, Karbyshev M, Löhr F, Bernhard F, Dötsch V (2010) Cell-free expression and stable isotope labelling strategies for membrane proteins. *J Biomol NMR* 46:33–43
- Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* 242:1162–1164
- Stech M, Merk H, Schenk JA, Stöcklein WFM, Wüstenhagen DA, Micheel B, Duschl C, Bier FF, Kubick S (2012) Production of functional antibody fragments in a vesicle-based eukaryotic cell-free translation system. *J Biotechnol* 164:220–231
- Takai K, Sawasaki T, Endo Y (2010) The wheat-germ cell-free expression system. *Curr Pharm Biotechnol* 11:272–278
- Uhlemann E-ME, Pierson HE, Fillingame RH, Dmitriev OY (2012) Cell-free synthesis of membrane subunits of ATP synthase in phospholipid bicelles: NMR shows subunit a fold similar to the protein in the cell membrane. *Protein Sci* 21:279–288
- Umakoshi H, Suga K, Bui HT, Nishida M, Shimanouchi T, Kuboi R (2009) Charged liposome affects the translation and folding steps of in vitro expression of green fluorescent protein. *J Biosci Bioeng* 108:450–454

- Uzawa T, Yamagishi A, Oshima T (2002) Polypeptide synthesis directed by DNA as a messenger in cell-free polypeptide synthesis by extreme thermophiles, *Thermus thermophilus* HB27 and *Sulfolobus tokodaii* strain 7. *J Biochem* 131:849–853
- Wada T, Shimono K, Kikukawa T et al (2011) Crystal structure of the eukaryotic light-driven proton-pumping rhodopsin, acetabularia rhodopsin II, from marine alga. *J Mol Biol* 411:986–998
- Wang Z (2006) Controlled expression of recombinant genes and preparation of cell-free extracts in yeast. *Methods Mol Biol* 313:317–331
- Wang X, Liu J, Zheng Y, Li J, Wang H, Zhou Y, Qi M, Yu H, Tang W, Zhao WM (2008) An optimized yeast cell-free system: sufficient for translation of human papillomavirus 58 L1 mRNA and assembly of virus-like particles. *J Biosci Bioeng* 106:8–15
- Wang X, Corin K, Baaske P, Wienken CJ, Jerabek-Willemsen M, Duhre S, Braun D, Zhang S (2011) Peptide surfactants for cell-free production of functional G protein-coupled receptors. *Proc Natl Acad Sci U S A* 108:9049–9054
- Wang X, Cui Y, Wang J (2013) Efficient expression and immunoaffinity purification of human trace amine-associated receptor 5 from *E. coli* cell-free system. *Protein Pept Lett* 20:473–480
- Witherell G (2001) In vitro translation using HeLa extract. *Curr Protoc Cell Biol. Unit* 11:8
- Wulfetange K, Saenger W, Schmülling T, Heyl A (2011) *E. coli*-based cell-free expression, purification and characterization of the membrane-bound ligand-binding CHASE-TM domain of the cytokinin receptor CRE1/AHK4 of *Arabidopsis thaliana*. *Mol Biotechnol* 47:211–219
- Wuu JJ, Swartz JR (2008) High yield cell-free production of integral membrane proteins without refolding or detergents. *Biochim Biophys Acta* 1778:1237–1250
- Yabuki T, Motoda Y, Hanada K, Nunokawa E, Saito M, Seki E, Inoue M, Kigawa T, Yokoyama S (2007) A robust two-step PCR method of template DNA production for high-throughput cell-free protein synthesis. *J Struct Funct Genomics* 8:173–191
- Yang J-P, Cirico T, Katzen F, Peterson TC, Kudlicki W (2011) Cell-free synthesis of a functional G protein-coupled receptor complexed with nanometer scale bilayer discs. *BMC Biotechnol* 11:57
- Yildiz AA, Knoll W, Gennis RB, Sinner E-K (2012) Cell-free synthesis of cytochrome bo(3) ubiquinol oxidase in artificial membranes. *Anal Biochem* 423:39–45
- Yin G, Swartz JR (2004) Enhancing multiple disulfide bonded protein folding in a cell-free system. *Biotechnol Bioeng* 86:188–195
- Zawada JF, Yin G, Steiner AR, Yang J, Naresh A, Roy SM, Gold DS, Heinsohn HG, Murray CJ (2011) Microscale to manufacturing scale-up of cell-free cytokine production—a new approach for shortening protein production development timelines. *Biotechnol Bioeng* 108:1570–1578
- Zhou Y, Asahara H, Gaucher EA, Chong S (2012) Reconstitution of translation from *Thermus thermophilus* reveals a minimal set of components sufficient for protein synthesis at high temperatures and functional conservation of modern and ancient translation components. *Nucleic Acids Res* 40:7932–7945

Chapter 3

Bacterial Expression and Stabilization of GPCRs

Jean-Louis Banères

3.1 Introduction

Among integral membrane proteins (IMPs), G protein-coupled receptors (GPCRs) constitute the largest family with more than 800 different receptors identified (Lagerstrom and Schiöth 2008). GPCRs are involved in most essential cellular processes and are the targets of about 30% of current pharmaceutical drugs (Lagerstrom and Schiöth 2008). It is therefore crucial to gain knowledge of their structure and their conformational dynamics in order to understand their functions and/or dysfunctions, as well as to design more selective and specific drugs. However, although the field has witnessed spectacular progresses during the past years, getting detailed structural information on GPCRs is still a long and winding road.

Except for rhodopsin (Hofmann et al. 2009), the low abundance of GPCRs precludes their purification in biochemically relevant amounts from natural sources. Overexpression is thus a prerequisite for structure/function analyses. However, overexpressing GPCRs, in particular in their unmodified state, may be problematic. Many different expression systems have been tested so far (see Chaps. 1, 2, 4–6 from this issue). One of the most popular for structural analyses actually is the insect cell system where most of the receptors crystallized have been produced, although in a modified state (Cooke et al. 2013). Yeast cells have also been used to express the histamine receptor whose crystal structure has been solved (Shiroishi et al. 2011). Other expression systems such as cell-free synthesis (Proverbio et al. 2013; see also Chap. 2) and mammalian cells (Andrell and Tate 2013) are explored to produce GPCRs.

Among all these expression systems, *Escherichia coli* still holds promises. This is due to its simplicity, scalability, and homogeneity of the recombinant protein obtained. Moreover, this system is essentially unique when it comes to specific

J.-L. Banères (✉)

Institute of Biomolecules Max Mousseron, UMR-CNRS 5247, Faculty of Pharmacy,
Université Montpellier 1 and 2, Montpellier, France
e-mail: Jean-louis.baneres@univ-montp1.fr

labeling of the expressed protein. This is particularly true when labeling with isotopes for nuclear magnetic resonance (NMR) studies is considered. Indeed, complex labeling protocols, homogeneous ^{15}N and/or ^{13}C labeling, and perdeuteration are routinely carried out in bacterial expression systems only (Verardi et al. 2012). This recently allowed the structure of the unmodified CXCR1 chemokine receptor to be solved in its lipid environment using solid-state NMR (Park et al. 2012b, see also Chap. 12 from Catoire et al. in this volume). In this case, the labeled receptor was expressed in *E. coli* inclusion bodies (IBs) and subsequently folded back to its native state in a lipid medium (Park et al. 2012a).

Once efficient expression has been achieved, purifying sufficient amounts of functional and stable protein still remains a challenge. GPCRs, as most IMPs, have to be handled in aqueous solutions in complex with surfactants, usually detergents. Because detergents tend to be inactivating, identifying a detergent or a lipid/detergent mixture that ensures protein homogeneity, functionality, and stability is often a limiting step. Different approaches have been developed that consist in either increasing the stability of the receptor through specific mutations or through the use of new generations of detergents, surfactant, and membrane-mimicking environments (Popot 2010, see also Chaps. 7 and 8 from M. Zoonens and from G. Durand, respectively, in this volume).

Here, we review the most recent advances in the production of GPCRs in *E. coli*, with a special emphasis on the strategy that is based on expression in IBs followed by solubilization and in vitro folding of the receptor. We then take a brief glimpse of the methods used to stabilize purified receptors in solution for subsequent structure/function analyses.

3.2 Functional Expression of GPCRs in *E. Coli* Inner Membranes

Since GPCRs are IMPs, targeting recombinant receptors to the inner membrane of the bacterium is considered as the most obvious strategy. Different approaches have been developed so far to optimize membrane expression that consist in screening for a fusion partner to efficiently address the receptor to the membrane, manipulating the bacterial strains to avoid toxicity, or manipulating the receptor sequence to increase its expression level and/or stability in inner membranes.

In most cases, efficient insertion into the bacterial inner membrane is achieved only by fusing the GPCR to a protein partner. During the first time of GPCR expression in *E. coli*, β -galactosidase was used to produce the β_2 -adrenergic receptor (Marullo et al. 1988). Since then, other fusion partners have been used. These include, for instance, a bacterial membrane protein, OmpF (Wiktor et al. 2013). Mystic, an IMP expression enhancer, also allowed an enhancement of the expression of the cannabinoid CB2 receptor at the bacterial membrane in combination with TarCF, a C-terminal fragment of the bacterial chemosensory transducer Tar

(Chowdhury et al. 2012). In the same way, fusion of the green fluorescent protein to the Cterminus of the cannabinoid CB1 and bradykinin B2 receptors led to efficient membrane expression (Skretas and Georgiou 2009). The most extensively described and used fusion system, however, is maltose-binding protein (MBP) alone or combined to thioredoxine A (TRX). Such fusions have been shown to be particularly well adapted for high-level expression of the neurotensin NTS1 (Tucker and Grisshammer 1996), the adenosine A_{2A} (Weiss and Grisshammer 2002), or the cannabinoid CB2 (Yeliseev et al. 2005) receptors in the *E. coli* inner membrane.

Another way to optimize expression of functional GPCRs in *E. coli* consists in manipulating the bacterial strain to suppress toxicity and/or enhance expression. We mention this approach only briefly since it is detailed in Chap. 4 of this volume. The best-illustrated example is that of the C41 and C43 strains (Miroux and Walker 1996) where, besides several MPs, some GPCRs have been produced (e.g., chemokine receptors; Ren et al. 2009). In the same way, some bacterial genes have been identified such as those encoding the ribonucleotide phosphatase NagD, a fragment of the predicted lipoprotein NlpD, and the three-gene cluster *ptsN-yhbJ-npr*, which globally enhance the production of properly folded GPCRs in *E. coli* (Skretas et al. 2012).

Finally, a most promising strategy to get functional receptors may consist in modifying the sequence of the receptor gene itself to optimize its stability at the bacterial membrane. An exciting evolutionary approach has been recently reported to identify structural features in a receptor that are important for biosynthesis and stability in the membrane of *E. coli* (Schlinkmann et al. 2012a; Schlinkmann et al. 2012b; Schlinkmann and Plückthun 2013). Using an in vitro recombination technology, Plückthun and coworkers obtained evolved variants of the neurotensin NTS1 receptor with both maximal functional expression and stability in the short-chain detergents used for crystallization. Less extensive mutagenesis strategies, essentially based on the replacement of the cysteine residues of the receptor, have also been reported that, in some cases, may enhance expression, although in a much more limited manner (Wiktor et al. 2013).

Expression of the receptor in the membrane of the bacteria then implies that the protein is subsequently solubilized from the membrane, usually with detergents, to achieve purification and biochemical and biophysical characterizations. Although trivial at first look, this step may be as problematic as achieving high-level expression or in vitro refolding; indeed, finding a detergent or a detergent combination that preserves the function of the expressed receptor can sometimes be an exhausting, time-consuming task. This step, however, is not detailed here since the problems encountered are common to the general process of GPCR stabilization in solution that is addressed further. They are also addressed in Chaps. 1, 7, and 8 of the present book.

3.3 Expression of GPCRs in *E. Coli* Inclusion Bodies and *In Vitro* Refolding

Expression of heterologous proteins in *E. coli* is frequently associated with incorrect folding and accumulation of the recombinant protein in cytoplasmic aggregates, IBs. Although commonly judged as a failure, getting GPCRs in IBs may present some advantages. IBs are not toxic to the cell so that high-level expression can be achieved, and they are resistant to proteolytic degradation. Accumulating receptors in IBs implies, however, that the receptors expressed have to be subsequently folded back to their native state; although considerable progresses have been achieved in the field, *in vitro* refolding still represents an arduous challenge and the bottleneck of this production strategy.

3.3.1 Accumulating GPCRs in IBs

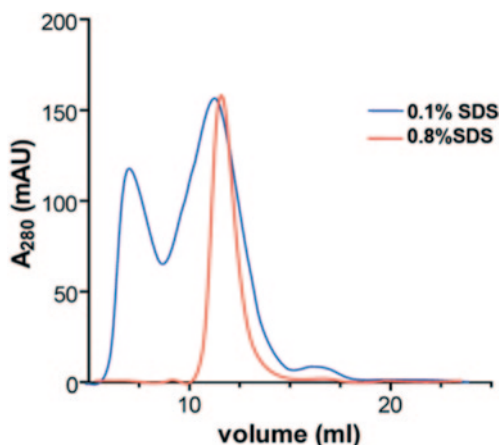
As is the case for functional expression in *E. coli* membranes, a fusion partner is generally required for efficient accumulation of GPCRs in IBs. An example of such is glutathione S-transferase (GST) that was first used to express olfactory receptor OR5 (Kiefer et al. 1996) and then several other receptors (Kiefer et al. 2000), including the chemokine receptor whose three-dimensional structure has been recently solved using solid-state NMR (Park et al. 2012a, b). In the case of the serotonin 5-HT4(a) receptor, the receptor was fused to the bacterial ketosteroid isomerase (Baneres et al. 2005). A high-throughput effort aimed at evaluating the efficacy of various fusion partners to target GPCRs to IBs identified GST and TRX as the most efficient ones, although some GPCRs were overexpressed without any protein tag (Michalke et al. 2009).

In the search for a fusion partner that could be used to systematically express any GPCR in IBs, we stumbled over a fragment of the extracellular domain of $\alpha 5$ integrin ($\alpha 5I$; Arcemisbehere et al. 2010). When fused to the N-terminus of the receptor, this fragment allowed many different rhodopsin-like GPCRs to be expressed in their full-length, unmodified state, regardless of their size and sequence. Importantly, in this case, high-level expression does not require any optimization of the receptor coding sequence, of the culture conditions, or of the extraction/purification procedures. This $\alpha 5I$ fusion strategy may therefore represent an important breakthrough for expression strategies involving accumulation of the GPCR in IBs by providing a systematic way to accumulate receptors in IBs.

3.3.1.1 Solubilizing GPCRs from IBs

Following expression, IBs are solubilized under denaturing conditions and the fusion partner removed. Although usually considered as a trivial step, solubilization of the receptor is critical for its subsequent refolding. While water-soluble proteins

Fig. 3.1 Influence of the aggregation state of the solubilized protein on the folding ratio. Size-exclusion (S200HR; 16 × 300 mm) profile of the purified leukotriene B₄ receptor BLT1R solubilized from inclusion bodies at different SDS concentrations. The folding yield obtained from the protein in 0.1% SDS amounted 5–6%, whereas that of the receptor in 0.8% SDS was ca. 30% under the same refolding conditions



are usually denatured with chaotropic compounds such as guanidium chloride or urea, IMPs such as GPCRs require harsh detergents. The rat olfactory receptor OR5 was solubilized from IBs in *N*-lauroyl sarcosine (Kiefer et al. 1996) while all other GPCRs were best solubilized with sodium dodecyl sulfate (SDS). Importantly, although one could imagine that solubilization is a straightforward step that simply consists in adding the harsh detergent, the solubilization conditions, and in particular the detergent concentration, need to be carefully adjusted. Indeed, an efficient refolding requires the protein not to be aggregated, and this is highly dependent on different parameters, in particular SDS concentration (see Fig. 3.1).

While it is generally considered that the solubilized state used as starting point in refolding protocols is an unfolded one, it appears that, in SDS, IMPs retain a significant amount of secondary structure. For instance, for the leukotriene B₄ BLT1 receptor, the amount in predicted α -helical structure is ca. 40%, and closely related to that measured after *in vitro* refolding (ca. 55%; Fig. 3.2). This is the case also for the μ -opioid receptor (Muller et al. 2008) and other IMPs such as bacteriorhodopsin (Huang et al. 1981) or the drug transporter *emrE* (Miller et al. 2009). However, although it likely contains some or even most parts of the native secondary structure, the SDS-solubilized protein is totally unable to bind its ligands (inset to Fig. 3.2), suggesting that the crucial intramolecular contacts required for maintaining/promoting the native state of the receptor are not achieved. The SDS-unfolded state after IB's solubilization should thus be considered as a nonnative, pre-folded state rather than a totally unfolded one.

3.3.2 *In Vitro* Folding of GPCRs

GPCR *in vitro* folding consists in exchanging the detergent used for solubilizing the receptor from IBs, usually SDS, with a milder surfactant. Under these conditions, regions that have a propensity to fold may do so, allowing native-like interactions

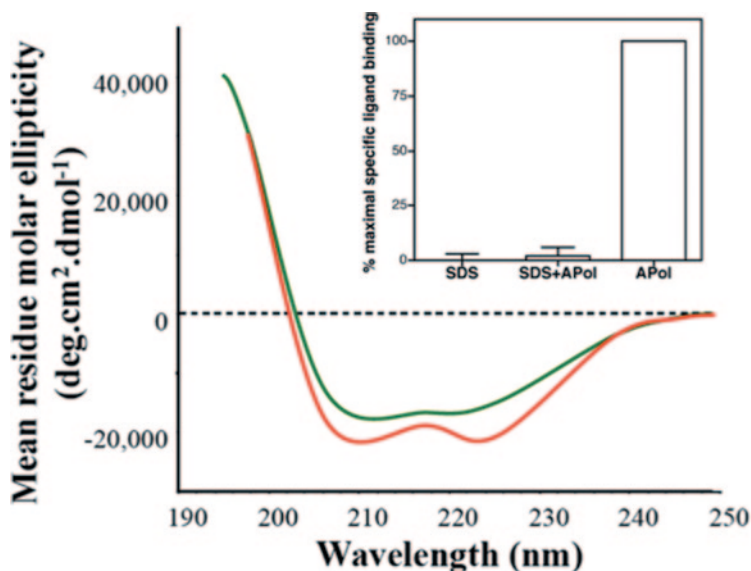


Fig. 3.2 Secondary structure and ligand-binding properties of the leukotriene B₄ receptor BLT1R before and after refolding. Far-UV circular dichroism spectra of BLT1R after solubilization in 0.8% SDS (*green trace*) and after refolding using LDAO as a detergent (*red trace*). *Inset*: percentage of ligand-competent receptor in the SDS-solubilized state (*SDS*), in the presence of both SDS and the amphipol used for refolding (*SDS+APol*) and after folding in amphipol (*APol*). Data are normalized to the amount of active receptor after refolding

between folded segments to form. Finding favorable folding conditions implies identifying a surfactant that will favor intramolecular native interactions versus nonnative intermolecular ones. Indeed, the former will favor folding back to the native state whereas the latter will lead to aggregation. Various environments that could stabilize the native fold of GPCRs have been reported so far, including detergents, lipid vesicles, and surfactants such as amphipols.

3.3.2.1 Folding Methods

Several methods have been used to efficiently exchange the detergents. These include matrix-assisted exchange, dilution, and SDS precipitation. Matrix-assisted methods are based on the exchange of the detergent while the receptor is immobilized on a chromatographic column, usually through an engineered Histag. Their main advantage is a better control of the intermolecular interactions that promote aggregation during refolding; several GPCRs have been folded using this method, including the olfactory OR5 receptor (Kiefer et al. 1996), the leukotriene B₄ BLT1 receptor (Baneres et al. 2003), and the serotonin 5-HT₄(a) receptor (Baneres et al. 2005). Dilution of the SDS-solubilized protein into a buffer containing the

Table 3.1 Example of some GPCRs refolded in detergents

Receptor	Detergent	SDS removal method	Folding yield (%)	Reference
BLT1R	LDAO	Matrix assisted	~20–30	(Baneres et al. 2003)
BLT2R	DPC/HDM	Matrix assisted	4–5	(Arcemisbehere et al. 2010)
OR5	Digitonin	Matrix assisted	80	(Kiefer et al. 1996)
PTH1R	DDM, cymal6	Dilution (in the presence of MCD)	50	(Michalke et al. 2010)
CB1R	DDM, cymal6	Dilution (in the presence of MCD)	50	(Michalke et al. 2010)
GLP1R	Brij78	Dilution (in the presence of MCD)	50–80	(Schroder-Tittmann et al. 2010)
NPY2R	DDM/CHAPS	Dilution (in the presence of MCD)	25–30	(Bosse et al. 2011)

CHAPS 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate, *Cymal6* 6-cyclohexyl-1-hexyl- β -D-maltoside, *DDM* n-dodecyl- β -D-maltopyranoside, *DPC* n-dodecylphosphocholine, *HDM* n-hexadecyl- β -D-maltopyranoside, *LDAO* n-dodecyl-N, N-dimethylamine-N-oxide

surfactant supposed to stabilize the native fold of the receptor, usually in the presence of methyl- β -D-cyclodextrin, has also been used to fold back to their native state several GPCRs such as the cannabinoid CB1 receptor (Michalke et al. 2010), the parathyroid hormone receptor PTHR1 (Michalke et al. 2010), the neuropeptide Y receptor NPY2R (Schmidt et al. 2009), and the glucagon-like peptide-1 receptor GLP1R (Schroder-Tittmann et al. 2010). Finally, while initially developed to refold bacteriorhodopsin (Popot et al. 1987), the removal of SDS through precipitation of its potassium salt allowed an efficient renaturation of several GPCRs into different environments (Dahmane et al. 2009; Park et al. 2012a). Besides these methods, other approaches have been tested. A particularly original one is that developed by Zaccai and coworkers (Zaccai et al. 2007) who used microfluidic channels to refold correctly an IMP, bacteriorhodopsin (Zaccai et al. 2007).

In all cases, the crucial parameter appears to be the kinetics of SDS-to-surfactant exchange that probably needs to be fast enough for an efficient folding of the receptor. The rationale behind is that a fast exchange likely leaves the protein little chances to explore misfolding or aggregation opportunities offered by partially denaturing environments. This probably explains why slower methods such as dialysis or adsorption of the detergent onto BioBeads or cyclodextrins appear to be rather inefficient.

3.3.2.2 Folding in Detergents

As shown in Table 3.1, efficient refolding of GPCRs in detergents has been reported for several receptors with yields ranging from a few percent (BLT2; Arcemisbehere et al. 2010) to about 80% (OR5; Kiefer et al. 1996). Most of the percentages of functional recovery lie in the 20–40% range, however. One of the major problems

Table 3.2 Example of some GPCRs refolded in lipids

Receptor	Lipid composition	SDS removal method	Folding yield (%)	Reference
CXCR1	DMPC (proteoliposome)	Precipitation	–	(Park et al. 2012a)
5-HT4(a)R	DMPC/CHAPS (bicelle)	Matrix assisted	20–25	(Baneres et al. 2005)
NPY2R	DMPC/CHAPS (bicelle)	Dilution	65	(Schmidt et al. 2009)

DMPC 1,2-dimyristoyl-sn-glycero-3-phosphocholine, *CHAPS* 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate

in refolding GPCRs in classical detergents is that finding favorable folding conditions implies identifying a detergent that will favor native intramolecular interactions, and very different detergent compositions have been reported so far. Inferring a general rule is difficult so that one has to screen for many different detergents and/or detergent mixtures to find the most appropriate one. Identifying the optimal conditions is therefore time and material consuming, even with miniaturized assays. Moreover, as stated above, detergents are in most cases denaturing for IMPs. This is the case when solubilizing an IMP from a membrane fraction, but also when *in vitro* refolding is considered. As a consequence, refolding in detergents usually implies an additional step that consists in replacing the detergent used during the refolding procedure with a more stabilizing environment (see Sect. 3.4).

3.3.2.3 Folding in Lipids

The efficiency of lipid environments to fold GPCRs back to their native state after solubilization from IBs has been described for a handful of receptors (Table 3.2). For instance, the serotonin 5-HT4(a) receptor (Baneres et al. 2005) and the neuropeptide Y2 receptor (Schmidt et al. 2009) were refolded in 1,2-dimyristoyl-sn-glycero-3-phosphocholine/3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (DMPC/CHAPS) bicelles. Bicelles are bilayer discs composed of mixtures of long-chain and short-chain phospholipids. Another example of refolding into a lipid environment comes from the pioneering work of H. Kiefer with the olfactory receptor OR5 that was refolded in 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine/1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPC/POPG) lipid vesicles (Kiefer et al. 1996). However, in this case, the refolding protocol included a preliminary N-lauroylsarcosine-to-digtonin exchange step. The receptor in digtonin was already able to bind its lial ligand, indicating that folding back to the native state probably occurred before insertion into the lipid vesicle already. Finally, successful folding of GPCRs by direct transfer to lipid vesicles is limited to a very few examples such as the chemokine receptor CXCR1 that was inserted in bilayers directly after SDS precipitation (Park et al. 2012a).

3.3.2.4 Folding in Amphipols

Amphipols (APols) and their applications, including *in vitro* folding, are described in detail in Chap. 8. Briefly, they are short amphipathic polymers with a high density of hydrophobic chains and highly hydrophilic groups that are able to keep individual IMPs soluble under the form of small, stable complexes (Popot 2010). Although initially developed to maintain the native fold of membrane proteins in solution with better efficacy than classical detergents (Tribet et al. 1996), APols appear as an efficient medium for folding IMPs also. This was demonstrated using model proteins such as *OmpA*, *FomA*, and bacteriorhodopsin (Pocanschi et al. 2006; Dahmane et al. 2013; Pocanschi et al. 2013) and further extended to GPCRs recovered from IBs (Bazzacco et al. 2012; Dahmane et al. 2009). An example for such a folding efficiency is illustrated in the inset to Fig. 3.2. Of importance, conditions initially established to refold bacteriorhodopsin were applied essentially without any major change to several GPCRs to achieve folding yields between 30 and 50% (Bazzacco et al. 2012; Damian et al. 2012; Dahmane et al. 2009), suggesting that this method of refolding could be of general use, in contrast to that based on detergents. If this is the case, as our data suggest, it will represent an important breakthrough for *in vitro* studies of purified GPCRs.

Of importance, as observed for most APol-trapped membrane proteins, GPCRs folded in amphipols are significantly more stable than those kept in lipid/detergent mixtures while the pharmacological properties of the receptor are maintained (Dahmane et al. 2009). Moreover, APols can be readily exchanged with detergents without significant unfolding of the protein (Zoonens et al. 2007) so that the APol-refolded receptor can be either used as such in biophysical and biochemical studies (Catoire et al. 2010) or, if necessary, as a starting state for subsequent insertion into other membrane-mimicking environments. This procedure has been recently used to reconstitute monomers and heterodimers of the ghrelin receptor into lipid nanodiscs (Damian et al. 2012; Mary et al. 2012; Mary et al. 2013).

3.3.2.5 Miscellaneous Considerations

In all cases, efficient refolding is dependent on several parameters that have to be considered. For instance, in strategies that involve a refolding in detergents or APols, the absence of lipids when solubilizing proteins from IBs is to be taken into account. As shown for several receptors, not only folding yields but also the stability of the folded protein are significantly increased in the presence of lipids (Dahmane et al. 2009, 2013). One possibility is that they do so by binding to sites that form when the transmembrane surface achieves its native state. Thereby, they would contribute to driving folding toward the latter. Scrambling of disulfide bridges can also be a problem in some cases. This problem is nevertheless alleviated by careful control of the redox potential during *in vitro* folding and/or by specific mutagenesis of the receptor cysteines (Witte et al. 2013).

To quantify the proportion of native receptor is also fundamental to evaluate the success of the folding strategy. In this context, classical strategies used when *in vitro* folding of globular proteins is considered may be inefficient to evaluate the efficacy of GPCR renaturation. For instance, estimation of the amount of secondary structure of the folded protein with circular dichroism (CD) is useless since, as stated above, the SDS-solubilized protein already displays a significant amount of α -helices. In the same way, Trp fluorescence cannot be used in the absence of a reference spectrum for the receptor in its native fold. Most of the folding experiments therefore rely on the availability of a sensitive, down-scalable, and robust functional assay, generally based on ligand-binding measurements. Finally, since the refolding yields are never quantitative, one has to be able to isolate and concentrate the native fraction into a homogeneous sample. This requires having at hand an efficient functional purification method. In most cases, it consists in an affinity-chromatography step on an immobilized ligand (Baneres et al. 2005; Bazzacco et al. 2012; Bosse et al. 2011).

3.4 In Vitro Stabilization of the Purified GPCRs

Detergents are often denaturing for IMPs (Popot 2010). This is true for purified GPCRs solubilized from membrane fractions of bacteria or other organisms or after refolding in detergents. For these reasons, different approaches have been developed to increase the functional stability of GPCRs in solution. Two divergent strategies have been used that rely on the modification either of the environment of the receptor or of the GPCR sequence. The latter consists in introducing random stabilizing mutations within the receptor sequence so that the receptor displays an increased stability, in particular in short-chain detergents (Scott et al. 2013; Tate and Schertler 2009). Such an approach has been used to crystallize several GPCRs such as the β_1 -adrenergic receptor (Warne et al. 2008), the adenosine A_{2A} receptor (Lebon et al. 2011), or the neurotensin NTS1 receptor (White et al. 2012). Whether stabilization of the GPCRs through specific mutations could favor receptor *in vitro* folding in strategies based on the accumulation of the protein in *E. coli* IBs is still an open question. To our knowledge, this has not been tested yet. These mutations could favor receptor refolding by stabilizing the native arrangement during the folding pathway or, on the contrary, prevent the receptor from exploring its conformational landscape and thus limit its ability to find its native fold.

Modification of the receptor sequence can introduce some bias, however, in some structure/function analyses. An alternative approach to the mutagenesis-based strategy therefore consists in using an unmodified receptor reconstituted in a membrane-mimicking environment that stabilizes the native fold of the protein better than any classical detergent. Many different original detergents/surfactants have been developed during the past years and some of them used with purified GPCRs. We detail here only three of them, the maltose-neopentyl glucose series (MNGs), amphipols, and nanodiscs. Other compounds, such as bicatenary detergents, fluorinated

surfactants, or lipid-like peptides have been developed to stabilize IMPs in solution (see Chaps. 7, 8, and 9 in the present book) but their application to GPCRs remains anecdotic (Damian et al. 2007; Corin et al. 2011; Zhao et al. 2006).

MNGs are original amphiphiles of recent development that are characterized by a central quaternary carbon atom derived from neopentyl glycol and hydrophilic groups derived from maltose (Chae et al. 2010). They appear to be particularly useful for stabilizing GPCRs during extraction from the membrane and subsequent manipulation of the purified receptor. In particular, they allowed stabilization of several purified receptors and receptor-G protein complexes during crystallization assays. So far, however, no improvement of the folding efficiency of GPCRs from IBs has been observed with MNGs compared to other classical detergents.

Amphipols have been used to stabilize the native fold of GPCRs in solution independently of their folding properties detailed above. For instance, they allowed the stabilization and the subsequent structural and functional characterization of the purified vasopressin V2 receptor extracted from insect cell membrane (Rahmeh et al. 2012). Of importance, as stated above, when refolding GPCRs in amphipols, one combines the advantages of high-yield folding ratios with the fact that the protein after renaturation is directly inserted in its stabilizing environment (Dahmane et al. 2009).

The last way used to stabilize GPCRs in solution better than classical detergents is to provide it with an environment mimicking that encountered in native conditions, i.e., a lipid bilayer. Besides lipid vesicles that have been used for several purified receptors but are not always adapted to biophysical applications, two of these kinds of lipid-containing structures have been used to stabilize GPCRs in solution, bicelles, and nanodiscs. As stated above, bicelles are bilayer discs composed of mixtures of long-chain and short-chain phospholipids. They are popular for many structural applications, in particular for NMR (Chap. 12); as such, they recently allowed the structure of the unmodified CXCR1 in its lipid environment to be obtained (Park et al. 2012b). Bicelles have also been used to stabilize different GPCRs such as opsin (McKibbin et al. 2007), the serotonin receptor (Baneres et al. 2005), or the NOP receptor (Thompson et al. 2011). Nanodiscs are also a discoidal lipid structure where the lipid bilayer is stabilized, in this case, by two encircling helical lipoproteins (Bayburt and Sligar 2010). These supramolecular structures make some sort of “lipidic cassettes” where the MP is soluble and stable. Lipid nanodiscs have been used to stabilize in solution a handful of GPCRs from natural sources (rhodopsin; Bayburt et al. 2007), or expressed in heterologous systems such as sf9 insect cells (β_2 -adrenergic receptor, Whorton et al. 2007; μ -opioid receptor, Kuszak et al. 2009, metabotropic glutamate receptor, El Moustaine et al. 2012), or mammalian cells (parathyroid hormone 1 receptor; Mitra et al. 2013). In the context of the present chapter dedicated to bacterial expression of GPCRs, lipid nanodiscs have been used to stabilize in solution receptors obtained both from *E. coli* inner membranes after solubilization with detergents (Inagaki et al. 2012) and from IBs after APol-assisted in vitro folding (Damian et al. 2012).

3.5 Conclusion

Many different expression systems have been developed to produce GPCRs for structural studies. Among them, one of the most popular actually is that based on the use of insect cells that has been used for most of the receptors crystallized so far. However, alternate systems such as yeast and mammal cells have also made the production of receptors possible. Among all the systems, bacterial ones have also their usefulness, in particular for specific approaches such as labeling with isotopes or nonnatural amino acids. The interest of such applications is illustrated by the recent NMR-based structure of the unmodified chemokine receptor CXCR1 obtained from bacterial IBs after *in vitro* refolding in a lipid-like environment. Many progresses have been made to improve the success rate of expression in *E. coli*. One can mention original mutagenesis strategies, the development of efficient fusion partners for high-level expression, and diversified protocols for successful *in vitro* folding. All this, combined to the development of original systems aimed at stabilizing purified receptors in solution, makes *E. coli* still a host of choice for expressing GPCRs and subsequent structural studies.

References

- Andrell J, Tate CG (2013) Overexpression of membrane proteins in mammalian cells for structural studies. *Mol Membr Biol* 30(1):52–63
- Arcemisbehere L, Sen T, Boudier L, Balestre MN, Gaibelet G, Detouillon E, Orcel H, Mendre C, Rahmeh R, Granier S, Vives C, Fieschi F, Damian M, Durroux T, Baneres JL, Mouillac B (2010) Leukotriene BLT2 receptor monomers activate the G(i2) GTP-binding protein more efficiently than dimers. *J Biol Chem* 285(9):6337–6347
- Baneres JL, Martin A, Hullot P, Girard JP, Rossi JC, Parello J (2003) Structure-based analysis of GPCR function: conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant BLT1. *J Mol Biol* 329(4):801–814
- Baneres JL, Mesnier D, Martin A, Joubert L, Dumuis A, Bockaert J (2005) Molecular characterization of a purified 5-HT4 receptor: a structural basis for drug efficacy. *J Biol Chem* 280(21):20253–20260
- Bayburt TH, Sligar SG (2010) Membrane protein assembly into Nanodiscs. *FEBS Lett* 584(9):1721–1727
- Bayburt TH, Leitz AJ, Xie G, Oprian DD, Sligar SG (2007) Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. *J Biol Chem* 282(20):14875–14881
- Bazzacco P, Billon-Denis E, Sharma KS, Catoire LJ, Mary S, Le Bon C, Point E, Baneres JL, Durand G, Zito F, Pucci B, Popot JL (2012) Nonionic homopolymeric amphipols: application to membrane protein folding, cell-free synthesis, and solution nuclear magnetic resonance. *Biochemistry* 51(7):1416–1430
- Bosse M, Thomas L, Hassert R, Beck-Sickinger AG, Huster D, Schmidt P (2011) Assessment of a fully active class A G protein-coupled receptor isolated from *in vitro* folding. *Biochemistry* 50(45):9817–9825
- Catoire LJ, Damian M, Giusti F, Martin A, van Heijenoort C, Popot JL, Guittet E, Baneres JL (2010) Structure of a GPCR ligand in its receptor-bound state: leukotriene B4 adopts a highly constrained conformation when associated to human BLT2. *J Am Chem Soc* 132(26):9049–9057

- Chae PS, Rasmussen SG, Rana RR, Gotfryd K, Chandra R, Goren MA, Kruse AC, Nurva S, Loland CJ, Pierre Y, Drew D, Popot JL, Picot D, Fox BG, Guan L, Gether U, Byrne B, Kobilka B, Gellman SH (2010) Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* 7(12):1003–1008
- Chowdhury A, Feng R, Tong Q, Zhang Y, Xie XQ (2012) Mystic and TarCF as fusion protein partners for functional expression of the cannabinoid receptor 2 in *Escherichia coli*. *Protein Expr Purif* 83(2):128–134
- Cooke RM, Koglin M, Errey JC, Marshall FH (2013) Preparation of purified GPCRs for structural studies. *Biochem Soc Trans* 41(1):185–190
- Corin K, Baaske P, Ravel DB, Song J, Brown E, Wang X, Wienken CJ, Jerabek-Willemsen M, Duhr S, Luo Y, Braun D, Zhang S (2011) Designer lipid-like peptides: a class of detergents for studying functional olfactory receptors using commercial cell-free systems. *PLoS One* 6(11):e25067
- Dahmane T, Damian M, Mary S, Popot JL, Baneres JL (2009) Amphipol-assisted in vitro folding of G protein-coupled receptors. *Biochemistry* 48(27):6516–6521
- Dahmane T, Rappaport F, Popot JL (2013) Amphipol-assisted folding of bacteriorhodopsin in the presence or absence of lipids: functional consequences. *Eur Biophys J* 42(2–3):85–101
- Damian M, Perino S, Polidori A, Martin A, Serre L, Pucci B, Baneres JL (2007) New tensio-active molecules stabilize a human G protein-coupled receptor in solution. *FEBS Lett* 581(10):1944–1950
- Damian M, Marie J, Leyris JP, Fehrentz JA, Verdie P, Martinez J, Baneres JL, Mary S (2012) High constitutive activity is an intrinsic feature of ghrelin receptor protein: a study with a functional monomeric GHS-R1a receptor reconstituted in lipid discs. *J Biol Chem* 287(6):3630–3641
- El Moustaine D, Granier S, Doumazane E, Scholler P, Rahmeh R, Bron P, Mouillac B, Baneres JL, Rondard P, Pin JP (2012) Distinct roles of metabotropic glutamate receptor dimerization in agonist activation and G-protein coupling. *Proc Natl Acad Sci U S A* 109(40):16342–16347
- Hofmann KP, Scheerer P, Hildebrand PW, Choe HW, Park JH, Heck M, Ernst OP (2009) A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem Sci* 34(11):540–552
- Huang KS, Bayley H, Liao MJ, London E, Khorana HG (1981) Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. *J Biol Chem* 256(8):3802–3809
- Inagak S, Ghirlando R, White JF, Gvozdenovic-Jeremic J, Northup JK, Grisshammer R (2012) Modulation of the interaction between neurotensin receptor NTS1 and Gq protein by lipid. *J Mol Biol* 417(1–2):95–111
- Kiefer H, Krieger J, Olszewski JD, Von Heijne G, Prestwich GD, Breer H (1996) Expression of an olfactory receptor in *Escherichia coli*: purification, reconstitution, and ligand binding. *Biochemistry* 35(50):16077–16084
- Kiefer H, Vogel R, Maier K (2000) Bacterial expression of G-protein-coupled receptors: prediction of expression levels from sequence. *Receptors Channels* 7(2):109–119
- Kuszak AJ, Pitchiaya S, Anand JP, Mosberg HL, Walter NG, Sunahara RK (2009) Purification and functional reconstitution of monomeric mu-opioid receptors: allosteric modulation of agonist binding by Gi2. *J Biol Chem* 284(39):26732–26741
- Lagerstrom MC, Schioth HB (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov* 7(4):339–357
- Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AG, Tate CG (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* 474(7352):521–525
- Marullo S, Delavier-Klutchko C, Eshdat Y, Strosberg AD, Emorine L (1988) Human beta 2-adrenergic receptors expressed in *Escherichia coli* membranes retain their pharmacological properties. *Proc Natl Acad Sci U S A* 85(20):7551–7555
- Mary S, Damian M, Louet M, Floquet N, Fehrentz JA, Marie J, Martinez J, Baneres JL (2012) Ligands and signaling proteins govern the conformational landscape explored by a G protein-coupled receptor. *Proc Natl Acad Sci U S A* 109(21):8304–8309

- Mary S, Fehrentz JA, Damian M, Gaibelet G, Orcel H, Verdie P, Mouillac B, Martinez J, Marie J, Baneres JL (2013) Heterodimerization with Its splice variant blocks the ghrelin receptor 1a in a non-signaling conformation: a study with a purified heterodimer assembled into lipid discs. *J Biol Chem* 288(34):24656–24665
- McKibbin C, Farmer NA, Jeans C, Reeves PJ, Khorana HG, Wallace BA, Edwards PC, Villa C, Booth PJ (2007) Opsin stability and folding: modulation by phospholipid bicelles. *J Mol Biol* 374(5):1319–1332
- Michalke K, Graviere ME, Huyghe C, Vincentelli R, Wagner R, Pattus F, Schroeder K, Oschmann J, Rudolph R, Cambillau C, Desmyter A (2009) Mammalian G-protein-coupled receptor expression in *Escherichia coli*: I. High-throughput large-scale production as inclusion bodies. *Anal Biochem* 386(2):147–155
- Michalke K, Huyghe C, Lichiere J, Graviere ME, Siponen M, Sciara G, Lepaul I, Wagner R, Magg C, Rudolph R, Cambillau C, Desmyter A (2010) Mammalian G protein-coupled receptor expression in *Escherichia coli*: II. Refolding and biophysical characterization of mouse cannabinoid receptor 1 and human parathyroid hormone receptor 1. *Anal Biochem* 401(1):74–80
- Miller D, Charalambous K, Rotem D, Schuldiner S, Curnow P, Booth PJ (2009) In vitro unfolding and refolding of the small multidrug transporter EmrE. *J Mol Biol* 393(4):815–832
- Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260(3):289–298
- Mitra N, Liu Y, Liu J, Serebryany E, Mooney V, DeVree BT, Sunahara RK, Yan EC (2013) Calcium-dependent ligand binding and G-protein signaling of family B GPCR parathyroid hormone 1 receptor purified in nanodiscs. *ACS Chem Biol* 8(3):617–625
- Muller I, Sarramegna V, Renault M, Lafaquiere V, Sebai S, Milon A, Talmont F (2008) The full-length mu-opioid receptor: a conformational study by circular dichroism in trifluoroethanol and membrane-mimetic environments. *J Membr Biol* 223(1):49–57
- Park SH, Casagrande F, Chu M, Maier K, Kiefer H, Opella SJ (2012a) Optimization of purification and refolding of the human chemokine receptor CXCR1 improves the stability of proteoliposomes for structure determination. *Biochim Biophys Acta* 1818(3):584–591
- Park SH, Das BB, Casagrande F, Tian Y, Nothnagel HJ, Chu M, Kiefer H, Maier K, De Angelis AA, Marassi FM, Opella SJ (2012b) Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature* 491(7426):779–783
- Pocanschi CL, Dahmane T, Gohon Y, Rappaport F, Apell HJ, Kleinschmidt JH, Popot JL (2006) Amphipathic polymers: tools to fold integral membrane proteins to their active form. *Biochemistry* 45(47):13954–13961
- Pocanschi CL, Popot JL, Kleinschmidt JH (2013) Folding and stability of outer membrane protein A (OmpA) from *Escherichia coli* in an amphipathic polymer, amphipol A8-35. *Eur Biophys J* 42(2–3):103–118
- Popot JL (2010) Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annu Rev Biochem* 79:737–775
- Popot JL, Gerchman SE, Engelman DM (1987) Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process. *J Mol Biol* 198(4):655–676
- Proverbio D, Roos C, Beyersmann M, Orban E, Dotsch V, Bernhard F (2013) Functional properties of cell-free expressed human endothelin A and endothelin B receptors in artificial membrane environments. *Biochim Biophys Acta* 1828(9):2182–2192
- Rahmeh R, Damian M, Cottet M, Orcel H, Mendre C, Durroux T, Sharma KS, Durand G, Pucci B, Trinquet E, Zwier JM, Deupi X, Bron P, Baneres JL, Mouillac B, Granier S (2012) Structural insights into biased G protein-coupled receptor signaling revealed by fluorescence spectroscopy. *Proc Natl Acad Sci U S A* 109(17):6733–6738
- Ren H, Yu D, Ge B, Cook B, Xu Z, Zhang S (2009) High-level production, solubilization and purification of synthetic human GPCR chemokine receptors CCR5, CCR3, CXCR4 and CX3CR1. *PLoS One* 4(2):e4509

- Schlinkmann KM, Hillenbrand M, Rittner A, Kunz M, Strohner R, Pluckthun A (2012a) Maximizing detergent stability and functional expression of a GPCR by exhaustive recombination and evolution. *J Mol Biol* 422(3):414–428
- Schlinkmann KM, Honegger A, Tureci E, Robison KE, Lipovsek D, Pluckthun A (2012b) Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc Natl Acad Sci U S A* 109(25):9810–9815
- Schlinkmann KM, Pluckthun A (2013) Directed evolution of G-protein-coupled receptors for high functional expression and detergent stability. *Methods Enzymol* 520:67–97
- Schmidt P, Lindner D, Montag C, Berndt S, Beck-Sickingler AG, Rudolph R, Huster D (2009) Prokaryotic expression, in vitro folding, and molecular pharmacological characterization of the neuropeptide Y receptor type 2. *Biotechnol Prog* 25(6):1732–1739
- Schroder-Tittmann K, Bosse-Doenecke E, Reedt-Runge S, Ihling C, Sinz A, Tittmann K, Rudolph R (2010) Recombinant expression, in vitro refolding, and biophysical characterization of the human glucagon-like peptide-1 receptor. *Biochemistry* 49(36):7956–7965
- Scott DJ, Kummer L, Tremmel D, Pluckthun A (2013) Stabilizing membrane proteins through protein engineering. *Curr Opin Chem Biol* 17(3):427–435
- Shiroishi M, Kobayashi T, Ogasawara S, Tsujimoto H, Ikeda-Suno C, Iwata S, Shimamura T (2011) Production of the stable human histamine H(1) receptor in *Pichia pastoris* for structural determination. *Methods* 55(4):281–286
- Skretas G, Georgiou G (2009) Genetic analysis of G protein-coupled receptor expression in *Escherichia coli*: inhibitory role of DnaJ on the membrane integration of the human central cannabinoid receptor. *Biotechnol Bioeng* 102(2):357–367
- Skretas G, Makino T, Varadarajan N, Pogson M, Georgiou G (2012) Multi-copy genes that enhance the yield of mammalian G protein-coupled receptors in *Escherichia coli*. *Metab Eng* 14(5):591–602
- Tate CG, Schertler GF (2009) Engineering G protein-coupled receptors to facilitate their structure determination. *Curr Opin Struct Biol* 19(4):386–395
- Thompson AA, Liu JJ, Chun E, Wacker D, Wu H, Cherezov V, Stevens RC (2011) GPCR stabilization using the bicelle-like architecture of mixed sterol-detergent micelles. *Methods* 55(4):310–317
- Tribet C, Audebert R, Popot JL (1996) Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci U S A* 93(26):15047–15050
- Tucker J, Grisshammer R (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem J* 317(Pt 3):891–899
- Verardi R, Traaseth NJ, Masterson LR, Vostrikov VV, Veglia G (2012) Isotope labeling for solution and solid-state NMR spectroscopy of membrane proteins. *Adv Exp Med Biol* 992:35–62
- Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454(7203):486–491
- Weiss HM, Grisshammer R (2002) Purification and characterization of the human adenosine A(2a) receptor functionally expressed in *Escherichia coli*. *Eur J Biochem/FEBS* 269(1):82–92
- White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG, Grisshammer R (2012) Structure of the agonist-bound neurotensin receptor. *Nature* 490(7421):508–513
- Whorton MR, Bokoch MP, Rasmussen SG, Huang B, Zare RN, Kobilka B, Sunahara RK (2007) A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. *Proc Natl Acad Sci U S A* 104(18):7682–7687
- Wiktor M, Morin S, Sass HJ, Keibel F, Grzesiek S (2013) Biophysical and structural investigation of bacterially expressed and engineered CCR5, a G protein-coupled receptor. *J Biomol NMR* 55(1):79–95
- Witte K, Kaiser A, Schmidt P, Splith V, Thomas L, Berndt S, Huster D, Beck-Sickingler AG (2013) Oxidative in vitro folding of a cysteine deficient variant of the G protein-coupled neuropeptide Y receptor type 2 improves stability at high concentration. *Biol Chem* 394(8):1045–1056

- Yeliseev AA, Wong KK, Soubias O, Gawrisch K (2005) Expression of human peripheral cannabinoid receptor for structural studies. *Protein Sci* 14(10):2638–2653
- Zaccai NR, Yunus K, Matthews SM, Fisher AC, Falconer RJ (2007) Refolding of a membrane protein in a microfluidics reactor. *Eur Biophys J* 36(6):581–588
- Zhao X, Nagai Y, Reeves PJ, Kiley P, Khorana HG, Zhang S (2006) Designer short peptide surfactants stabilize G protein-coupled receptor bovine rhodopsin. *Proc Natl Acad Sci U S A* 103(47):17707–17712
- Zoonens M, Giusti F, Zito F, Popot JL (2007) Dynamics of membrane protein/amphipol association studied by Forster resonance energy transfer: implications for in vitro studies of amphipol-stabilized membrane proteins. *Biochemistry* 46(36):10392–10404

Chapter 4

Membrane Protein Production in *Escherichia coli*: Overview and Protocols

Georges Hattab, Annabelle Y. T. Suisse, Oana Ilioaia, Marina Casiraghi,
Manuela Dezi, Xavier L. Warnet, Dror E. Warschawski, Karine Moncoq,
Manuela Zoonens and Bruno Miroux

4.1 Introduction

Production of biological molecules is a challenge for the next decade in the field of medicinal chemistry. After heterologous production, the biological molecule must be active, well defined homogenous and the cost of its production should remain low. An interesting example is given by the relative success of therapeutic antibodies. Twenty monoclonal therapeutic antibodies are presently on the market (Oldham and Dillman 2008). All of them are produced with the hybridoma technology, which significantly increases the social cost of treating corresponding diseases and prevents the worldwide distribution of these drugs. Smaller-sized antibody peptides, named nanobodies, are being produced in bacteria to circumvent the cost of the hybridoma technology. Although *Escherichia coli* is probably the most versatile and the cheapest host for protein production, several obstacles remain: inclusion bodies formation, LPS contamination, incomplete synthesis, degradation by proteases, and the lack of post-translational modifications.

Georges Hattab and Annabelle Y. T. Suisse are equal first authors.

B. Miroux (✉) · G. Hattab · A. Y. T. Suisse · O. Ilioaia · M. Casiraghi · X. L. Warnet ·
D. E. Warschawski · K. Moncoq · M. Zoonens
Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099,
Institute of Physico-Chemical Biology, Université Paris Diderot, Paris, France
Phone: 33 1 58 41 52 25
e-mail: Bruno.Miroux@ibpc.fr

M. Dezi
Laboratory of Crystallography and NMR Biology, UMR-CNRS 8015, Université Paris
Descartes, Paris, France

In the case of membrane proteins, the situation is even more complex because they are difficult not only to produce but also to keep, in an active state, in solution. In medicinal chemistry, the need for large-scale production of membrane proteins is increasing. For instance, producing the major outer membrane protein (MOMP) of *Chlamydia trachomatis* is a major issue for establishing a robust vaccine against this pathogenic bacterium. Although the protein can easily be produced in bacteria and refolded in several detergents, only the native protein can be used to generate protective antibodies. Indeed, its quaternary structure must be preserved to generate an efficient B-cell response. Despite recent progress in maintaining the MOMP quaternary structure in solution (Tifrea et al. 2011), large-scale production of the protein is still a challenge. Bioproduction is a challenge not only for producing biological drugs or drug targets but also for the development of new drugs. Membrane proteins represent up to 50% of human drug targets (Overington et al. 2006). Several milligrams to grams of proteins are required to screen and validate drugs, which is a major limitation in pharmaceutical research.

Beyond its impact in medicinal chemistry and in the pharmaceutical industry, bioproduction is also a bottleneck for biologists and biophysicists. For instance, there are 424 unique membrane protein structures in the Protein Data Bank (PDB; see <http://blanco.biomol.uci.edu/mpstruc/>), which corresponds to only 2% of the total number of solved protein structures. Despite the exponential growth of membrane protein structures, they are still 20 years behind soluble protein, in terms of number of structures solved. Over the past decades, a tremendous effort has been invested in developing alternative expression systems and new surfactants (see Zoonens et al., Chap. 7 of this book for review; Chae et al. 2010; Popot et al. 2011) to purify and refold membrane proteins in an active state (Catoire et al. 2010). However, it becomes clear that determining the atomic structure of membrane proteins isolated in detergent might not answer fundamental biological questions. Membrane proteins may also need to be studied in native-like lipid membranes, which is even more challenging (Abdine et al. 2010; Park et al. 2012).

There is a need to develop robust expression systems for producing membrane proteins in native membranes. Although mammalian cell-based expression systems have been very successful for crystallization of G protein-coupled receptors (GPCR; Tate 2012), microorganisms, mainly bacteria and yeast, are still subject to intense studies and technological developments. For instance, Le Maire and colleagues have obtained in 2005 the first X-ray structure of a eukaryotic membrane protein after overexpression in the yeast *Saccharomyces cerevisiae*, by fusing the rabbit sarco/endoplasmic reticulum Ca^{2+} -ATPase (SERCA ATPase) with a biotin acceptor domain peptide (Jidenko et al. 2005). In parallel, 16 membrane protein structures have been obtained using the *Pichia pastoris* yeast expression system (for review, see Alkhalifioui et al. 2011), including two GPCRs. In bacteria, the lactobacillus expression system is highly promising because it has the main advantage of avoiding inclusion bodies formation (see Frelet-Barrand et al., Chap. 5 of this book and Frelet-Barrand et al. 2010 for review). However, the yield of membrane protein production remains low and, to our knowledge, this expression system has not generated any membrane protein structure.

The bacteria *E. coli* today is still the most widely used host for protein overexpression. Most prokaryotic membrane protein structures found in the PDB have been obtained after production of the corresponding protein in *E. coli*. Extending the production of membrane proteins in *E. coli* to eukaryotic sequences is facing two major problems: the formation of inclusion bodies (see the review from Banères, Chap. 3 of this book) and the toxicity associated with the induction of the target gene expression, which frequently results in cell death. This chapter will focus on the second aspect because overcoming the toxicity of expression has proven to be extremely useful and productive. A good example is given by bacterial mutants isolated using the T7 RNA polymerase-based expression system (see below for a full description). In this expression system, induction of the expression of the target gene by addition of the inducer Isopropyl- β -D-1-thiogalactopyranoside (IPTG) kills the cells, usually the BL21(DE3) bacterial host. This phenotype was used to screen for spontaneous mutants on IPTG-containing plates. Starting with the expression of the mitochondrial oxoglutarate carrier protein (OGCP) in the BL21(DE3) bacterial host, a first mutant was isolated, named C41(DE3), in which OGCP protein levels were strongly increased despite a tenfold reduction of corresponding mRNA levels (Miroux and Walker 1996). A second round of selection was conducted expressing *uncF*, which encodes AtpF, the *E. coli* b-subunit of the F1Fo ATP synthase, in C41(DE3) bacterial host. A second mutant C43(DE3) was isolated.

Overproduction of AtpF in its adapted bacterial host C43(DE3) resulted in the development of a large network of internal membranes. The bacterial host C43(DE3) reacted to the overproduction of a membrane protein by synthesizing lipids and by converting phosphatidyl glycerol into cardiolipids at the stationary phase (Weiner et al. 1984; Arechaga et al. 2000). Whereas de novo lipid synthesis may serve to maintain the lipid/protein ratio constant, the function of the increased cardiolipid content is unclear. Although the mutation in the C43(DE3) genome remains unknown, a delay in the transcription of the *uncF* gene (60 min) was observed, allowing membrane synthesis and proper folding of the b-subunit. Indeed, although AtpF forms inclusion bodies in C41(DE3) cells, it is readily inserted and folded in the membranes of C43(DE3) (Arechaga et al. 2000). Thus, slowing down the expression of *uncF* improved coupling between transcription, translation folding-insertion processes and consequently the storage of the b-subunit into internal proliferating membranes (Miroux and Walker 1996).

Membrane proliferation upon overexpression of a membrane protein has been observed before in *E. coli* (Weiner et al. 1984; von Meyenburg et al. 1984; Wilkison et al. 1986; Arechaga et al. 2000; Eriksson et al. 2009) and in the yeast (Wright et al. 1988). For instance, overproduction of the enzyme 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase in the yeast *Saccharomyces cerevisiae* resulted in the formation of paired membranes closely associated with the nuclear envelope called “Karmellae” (Wright et al. 1988). Proliferation of endoplasmic reticulum structures has also been observed upon the regulated overexpression of the P-type H(+) ATPase (Supply et al. 1993). However, in the case of AtpF, the stronger intensity of membrane proliferation opens a way to the study of AtpF in situ in its native membrane environment (see Chap. 12 from Catoire et al. of this book). Co-expression of AtpF with other membrane proteins of interest is also a promising avenue (Zoonens and Miroux 2010).

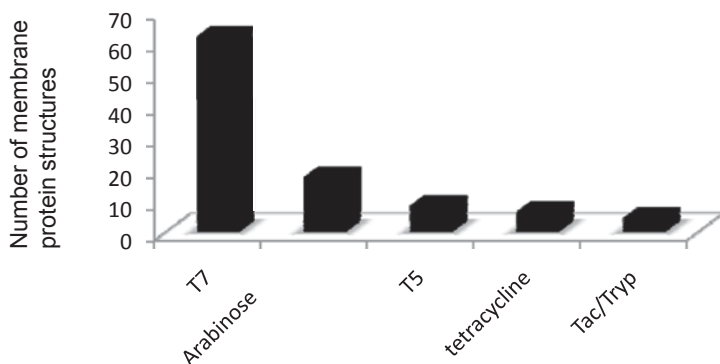


Fig. 4.1 Distribution of bacterial promoter usage in structural biology of membrane proteins (adapted from Hattab et al. 2014). A Hundred and fifty one unique membrane protein structures were extracted from the Protein Data Bank (Warschawski 2013; White 2013) on the basis of heterologous production of the protein in *Escherichia coli* (homologous production in *Escherichia coli* was excluded). The chart shows the number of solved membrane protein structures for each promoter used to produce the corresponding protein.

In order to assess the impact of these mutant hosts on structural biology of membrane proteins, we have conducted, 20 years after their discovery, a large-scale analysis of membrane protein structure databases (<http://www.drorlist.com/nmr/MPNMR.html> and <http://blanco.biomol.uci.edu/mpstruc/>; Hattab et al. 2014). Figure 4.1, adapted from Hattab et al. (2014), summarizes the PDB search and shows that the T7 RNA polymerase-based expression system (Novagen) accounts for more than 60% of membrane protein structures obtained after heterologous production in *E. coli*. The arabinose promoter-based expression system (Guzman et al. 1995) comes second, followed by T5 (Quiagen) and tetracycline promoter-based expression systems (IBA). Within the T7 expression system, the bacterial mutant hosts C41(DE3) and C43(DE3), commercially available from Lucigen, have been used for 50% of solved prokaryotic integral membrane protein structures so far. In this chapter, we will therefore focus on the T7 expression system, which, thanks to its multiple levels of regulation, has been the most successful in structural biology of membrane proteins.

4.2 Overview of the T7-Based Expression System

4.2.1 Regulation Levels of the T7 Expression System

In its most usual configuration, the T7 RNA polymerase gene is inserted in the genome of a lambda DE3 bacteriophage that is maintained into the lysogenic *E. coli* BL21(DE3) host. The T7 RNA polymerase gene is under the control of a *lacUV5*

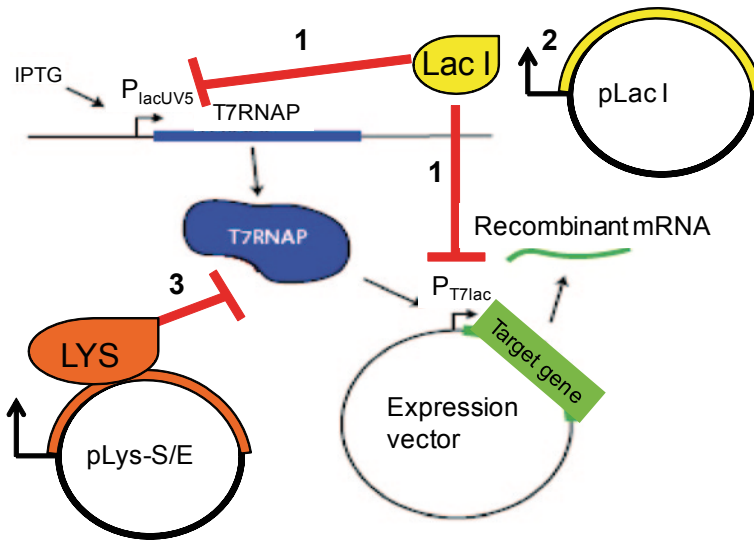


Fig. 4.2 Levels of regulation of the T7-based expression system. The amount of active T7 RNA polymerase is controlled in several ways: 1. repressing the *lacUV5* and *T7lac* promoters using the *lacI* repressor which binds to the *lacO* operating sequence; 2. expression of *lacI* gene from the expression plasmid or from a companion plasmid; and 3. expression of lysozyme, which inhibits the T7 RNA polymerase, from a companion plasmid.

promoter, which is weakly sensitive to cAMP regulation, associated with a *lacO*-regulating sequence (Fig. 4.2). The DE3 insert also contains *lacI* gene, which product represses the *lacUV5* promoter upon binding to the *lacO* sequence. The second level of regulation is the expression vector itself. In its simplest version, the vector only contains the T7 promoter (pRSET, Invitrogen; pMW7/pHIS; Way et al. 1990; Orriss et al. 1996), but many of the pET derivatives (Novagen) also contain a *T7lac* promoter that is fully repressed by the LacI repressor. In addition, the *lacI* gene is often expressed separately in a companion expression plasmid to ensure a multi-copy expression of the LacI repressor. A third level of regulation relies on the expression of lysozyme, either constitutively expressed or inducible by rhamnose (Wagner et al. 2008). Lysozyme specifically inhibits the T7 RNA polymerase, thus further attenuating the expression system. Two other parameters also influence the final strength and stability of the system: the plasmid copy number and antibiotic resistance genes. Multiple other versions of this expression system are still under development. For instance, in the BL21AI bacterial host (Invitrogen), the arabinose promoter replaces the *lacUV5* promoter. Expression of the T7 RNA polymerase can be titrated using increasing concentrations of arabinose. In the *Lemo* bacterial host, expression of the lysozyme is under the control of rhamnose promoter, which indirectly titrates the amount of active RNA polymerase via the expression of lysozyme (Wagner et al. 2008).

Table 4.1 Toxicity of T7 expression vectors

Plasmid name	Tag	Size of colonies on 2 × TY plate		
		BL21(DE3)	BL21(DE3)	C41(DE3)
		–IPTG	+IPTG	+IPTG
pMW7 ^a	None	Normal	None	Normal
pHis17 ^b	C-ter (6*) His	Normal	None	Normal
pET17b ^c	N-ter T7	Normal	Very small	Normal
pET29a ^d	N-ter S	Normal	None	Normal
pGEMEX-1	N-ter T7gene10	Small	None	Small

^aHigh copy number T7 plasmid from (Way et al. 1990)

^bDerivative of pMW7(Orriss et al. 1996)

^cpET series vector are low copy number derivatives of pBR322

^dContains T7/*lac*, *lacI* and Kan genes

4.2.2 Toxicity Associated with Membrane Protein Expression

Very early in the construction of the T7-based expression system, Studier and Mofatt noticed that the size of bacterial colonies on plate was dependent on the genomic insertion site of the lambda DE3 (Studier et al. 1990). Actually, the BL21(DE3) host was selected on its ability to form normal-sized colonies in the presence of expression plasmids but only in the absence of IPTG inducer. In the presence of IPTG, most expression plasmids, even without an inserted coding sequence behind their T7 promoter (“empty plasmid”), prevent the formation of colonies on plate. Table 4.1 gives an overview of different types of plasmids toxicities in the BL21(DE3) host. Very high copy number plasmids (>200 copies) that do not contain a *lacO* sequence, such as pMW7 or pHis vectors, do not allow colony formation on an IPTG-containing plate, even when they are empty. Low copy number plasmids (50 copies), i.e. deriving from pBR322, are slightly less toxic to BL21(DE3), showing that the expression plasmid copy number is an important parameter (Table 4.1, see pET17b phenotype). However, the addition of a coding sequence, even a small tag sequence such as S-Tag, completely prevents the growth on an IPTG plate. In contrast, none of the empty expression plasmids are toxic to the bacterial host C41(DE3), in which the production of T7 RNA polymerase is ten times decreased (Wagner et al. 2008). This suggests that a first level of toxicity occurs at the transcriptional level, when the T7 RNA polymerase is produced in excess. This basic level of toxicity does not necessarily compromise the successful expression of a target protein. Actually, in some cases where the target protein is produced at high levels, it can be advantageous to stop bacterial growth while expressing the target protein, in order to increase its concentration per cell and therefore to facilitate its purification. In isotope-labelling experiments, it can also be useful to specifically label the expressed protein so that, after purification, the remaining contaminants will be invisible on nuclear magnetic resonance (NMR) spectra. For this reason, we provide in the protocol section of this chapter some pragmatic tips to express proteins in toxic conditions.

1 Overload of the target mRNA



2 Unfolding stress and loss of periplasm integrity

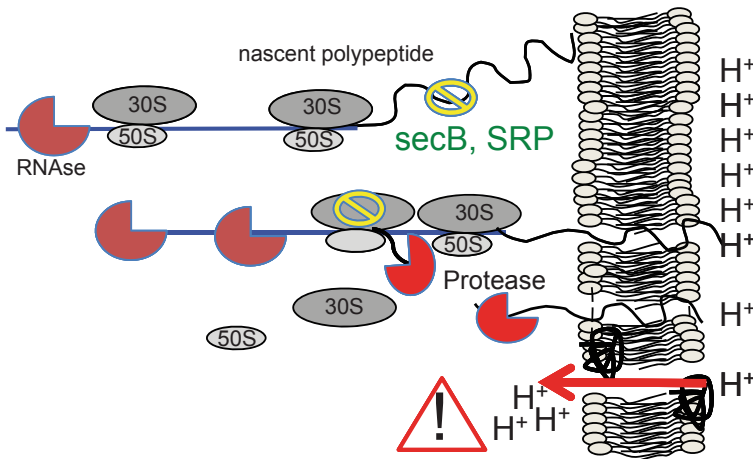


Fig. 4.3 Origins of toxicity during overexpression of membrane proteins in *Escherichia coli*. Overproduction of the target mRNA is toxic to the cell because it overloads the translation machinery at the expense of the cell's intrinsic protein synthesis (Dong et al. 1995). A second level of toxicity is linked to the folding and insertion of the newly synthesized protein. Co-translational folding of secondary structures is non-optimal and can lead to either inclusion bodies formation or mistargeting of the protein. Mistargeting occurs due to the lack of membrane-targeting signalling sequences and failure of chaperones to recognize the foreign protein sequence. A side effect of mistargeting the protein is the destabilization of the membranes upon aggregation of the proteins. Local disruption of the membrane triggers proton leak and loss of energy homeostasis.

A second level of toxicity occurs when the target protein, for instance a membrane protein, needs and therefore recruits and overloads the *E. coli* folding or insertion machinery (Wagner et al. 2007). In the best-case scenario, the chaperones recognize the foreign membrane protein but cannot synchronize its insertion into *E. coli* membranes because the T7 RNA polymerase is working too fast (Fig. 4.3). Consequently, an increasing fraction of the target membrane protein is partially inserted and folded in *E. coli* membranes, which in turn compromises ion gradient homeostasis and ultimately adenosine triphosphate (ATP) synthesis (Fig. 4.3).

Several strategies have been developed to overcome the toxicity associated with membrane protein expression: (1) adjusting the time course of expression of the target membrane protein by selection of bacterial mutants (Miroux and Walker

1996; see protocol in Sect. 4.3.2), (2) optimizing growth conditions (see protocol in Sect. 4.3.3 and Sevastyanovich et al. 2010, for review), (3) co-expressing bacterial chaperones (Chen et al. 2003), (4) inserting signal-targeting sequences to help the recognition of the foreign membrane protein by the *E. coli* machinery (for maltose-binding protein (MBP) fusion, see Miroux et al. 1993; Bocquet et al. 2008; Nury et al. 2011 as examples), (5) preventing misfolding into bacterial membranes by facilitating inclusion bodies formation (see Chap. 3 from JL Banère in this book and Mouillac and Banères 2010, for review), (6) introducing mutations into the target membrane protein to enhance its thermostability and/or its folding in vivo (Sarkar et al. 2008) and (7) cell-free expression of the target membrane protein using bacterial extracts (Rogé and Betton 2005; Miot and Betton 2011).

4.3 Protocol Section

4.3.1 *Choosing the Appropriate Strategy and Host/Vector Combination*

In a previous study (Hattab et al. 2014), we have conducted a systematic analysis of expression protocols in bacteria, based on membrane protein structures solved after heterologous expression of the protein in *E. coli*. Table 4.2 lists genotypes of all bacterial hosts that are used in structural biology of membrane proteins. Figure 4.1 summarizes one of the major outcomes of this study: T7 and arabinose-based promoters account for 80% of membrane protein structures. Therefore, we recommend running both expression systems in parallel to maximize your chances of getting your target membrane protein in sufficient amounts. The arabinose promoter-based expression system is well defined in terms of vector/bacterial host combination (Guzman et al. 1995). However, we have found ten membrane protein structures in the PDB that were solved after overproduction of the protein in the C43(DE3) bacterial host transformed with a pBAD arabinose-inducible vector (Hattab et al. 2014). This is unusual and requires further investigation. In this chapter, we focus on the T7 promoter-based expression system because C41(DE3) and C43(DE3) bacterial hosts account for 50% of heterologous integral membrane protein structures (Hattab et al. 2014).

A large survey on these bacterial host users revealed that high copy number vectors harbouring a wild-type T7 promoter, like the pRSET (Invitrogen), pMW7/pHis (Way et al. 1990; Orriss et al. 1996) or pPR/pPSG (IBA) vectors, are most frequently associated with C41(DE3) and C43(DE3) bacterial hosts. Avoid pET vectors bearing a pBR322 origin of replication and most importantly those carrying T7lac promoter and *lacI* gene. If you need to down-regulate your expression system, use, instead, the bacterial hosts BL21AI (Invitrogen) or C41(DE3) and C43(DE3) (Lucigen). In these hosts, the amount of T7 RNA polymerase is decreased or can be titrated with the inducer. Avoid companion plasmids that express lysozyme (pLyS/E) to inhibit the T7 RNA polymerase activity (Moffatt and Studier 1987) because they

Table 4.2 Genotypes of *Escherichia coli* hosts used for structural determination of membrane proteins

<i>T7-based expression hosts</i>	<i>Genotype</i>
BL21λ(DE3)	F- <i>ompT hsdS</i> (rB- mB-) <i>dcm gal</i> λ(DE3 [<i>lacI lacUV5-T7 gene 1 ind1 sam7 nin5</i>])
C41λ(DE3)	BL21λ(DE3 [<i>lacI lac-T7 gene 1 ind1 sam7 nin5</i>])
C43λ(DE3) ^a	C41λ(DE3) derivative
BL21λ(DE3) pLysS	BL21λ(DE3) pLysS (CamR)
BL21λ(DE3) CodonPlus ^a	BL21 <i>dcm</i> + TetR λ(DE3) <i>endA Hte</i> [<i>argU proL</i> CamR]
BL21 Star λ(DE3)	BL21 <i>rne131</i> λ(DE3)
BL21 Rosetta λ(DE3) pLysS	BL21 λ(DE3) pLysSRARE (CamR)
BL21λ(DE3) Tuner	BL21 <i>lacY1</i> λ(DE3)
BL21(AI)	BL21 <i>lon araB::T7RNAP-tetA</i>
<i>Other expression hosts</i>	<i>Genotype</i>
BL21Rosetta	BL21 RARE (CamR)
BL21-Gold	BL21 <i>dcm</i> + TetR <i>endA Hte</i>
BL21-T1 ^R competent	<i>fhuA2</i> [<i>lon ompT gal</i> [<i>dcm</i>] Δ <i>hsdS</i>
Origami B	BL21 <i>lacY1 aphC gor522::Tn10 trxB</i> (KanR TetR)
B834	F- <i>ompT hsdSB</i> (rB- mB-) <i>gal dcm met</i>
BLR	F- <i>ompT hsdSB</i> (rB- mB-) <i>gal dcm</i> Δ(<i>srl-recA</i>)306::Tn10 (TetR)
DH10B	F- <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74</i>
TOP10	<i>nupG recA1 araD139</i> Δ(<i>ara-leu</i>)7697 <i>galE15 galK16 rpsL endA1</i> λ-
	DH10B <i>rpsL</i> (StrR)
KRX	[F', <i>traD36, ΔompP proA</i> + B + <i>lacIq</i> Δ(<i>lacZ</i>)M15] Δ <i>ompT endA1 recA1 gyrA96</i> (Nalr) <i>thi-1 hsdR17</i> (rk-mk +) e14- (McrA-) <i>relA1 supE44</i> Δ(<i>lac-proAB</i>) Δ(<i>rhaBAD</i>)::T7 RNA polymerase
XL10-Gold	F' [<i>proAB lacIqZ</i> ΔM15 Tn10(TetR Amy CmR)] <i>recA1 endA1 glnV44 thi-1 gyrA96 relA1 lac Hte</i> Δ(<i>mcrA</i>)183 Δ(<i>mcrCB-hsdSMR-mrr</i>)173 TetR
XL1-Blue	F' [<i>proAB, lacIq Z</i> ΔM15 Tn10(TetR)] <i>recA1 endA1 gyrA96 thi-1 relA1 supE44 hsdR17</i> (rk-mk +) l-
DH5α	F- <i>ø80dlacZ</i> ΔM15 Δ(<i>lacZYA-argF</i>)U169 <i>deoR recA1 endA1 hsdR17</i> (rk-mk +) <i>phoA supE44</i> λ- <i>thi-1 gyrA96 relA1</i>
SG13009	NaI[s] Str[s] Rif[s] Thi[-] lac[-] Ara[+] Gal[+] Mtl[-] F[-] RecA[+] Uvr[+] Lon[+]
LS6164	Δ <i>fadR</i> Δ <i>fadL</i>
MC4100	F- [<i>araD139</i>]B/r Δ(<i>argF-lac</i>)169* &lambda- <i>e14-flhD5301</i> Δ(<i>fruK-yeiR</i>)725 (<i>fruA25</i>) <i>relA1 rpsL150</i> (strR) <i>rbsR22</i> Δ(<i>fimB-fimE</i>)632(::IS1) <i>deoC1</i>
SCM6	NS (Patented)
MC1061	F- Δ(<i>ara-leu</i>)7697 [<i>araD139</i>]B/r Δ(<i>codB-lacI</i>)3 <i>galK16 galE15</i> λ- e14- <i>mcrA0 relA1 rpsL150</i> (strR) <i>spoT1 mcrB1 hsdR2</i> (r- m+)
JM83	<i>rpsL ara</i> Δ(<i>lac-proAB</i>) Φ80 <i>dlacZ</i> ΔM15
Other	PA(Δ <i>oprH</i>)

^aAlso used in the arabinose expression system

require the addition of a second antibiotic (chloramphenicol), which quite substantially decreases cell growth. Moreover, excess of lysozyme impairs cell growth.

Once you have chosen your T7 vector, you need to decide whether to make a fusion to either direct your target gene to the *E. coli* membrane (for MBP fusion, see Bocquet et al. 2008 and Nury et al. 2011 as examples) or form inclusion bodies (for $\alpha 5$ integrin fusion, see Mouillac and Banères 2010, for review). If your protein is of prokaryotic origin, avoid fusion protein constructs or use a green fluorescent protein (GFP) fusion to monitor the yield and aggregation state of your protein on crude extract before any purification step (Drew et al. 2006). GFP fusions also offer the great advantage either to directly assess the production of your protein (Sarkar et al. 2008) or to select new bacterial hosts (Walker and Miroux 1999; Alfasi et al. 2011). If you wish to express an eukaryotic protein, be aware that there are almost no solved integral eukaryotic membrane protein structure after production in *E. coli*. There is one noticeable exception where the author succeeded in refolding and transferring directly the CXCR receptor into liposomes and solved the structure of the receptor by solid-state NMR analysis (Park et al. 2012). Thus, refolding of inclusion bodies from integral eukaryotic membrane proteins is an emerging promising avenue (see Chap. 3 from Banères and Chap. 12 from Catoire et al. of this book and references herein; Catoire et al. 2010; Banères et al. 2011; Park et al. 2012).

4.3.2 Selection of Bacterial Mutant Hosts

Transformation Transform your expression plasmid into BL21(DE3), which is the best host to start with because its induction power is maximal and easy to down-regulate. Prepare 2×Tryptone Yeast (TY) plates with antibiotic and IPTG. Two concentrations of IPTG may be used, i.e., 0.4 and 1 mM (Hattab et al. 2014). Use calcium chloride transformation with 10 ng of plasmid. After 1-h incubation at 37°C of the 1-ml transformation culture, plate 100 µl on 2×TY plate with antibiotic and 100 µl on 2×TY plates with antibiotic and both 0.4 and 1-mM IPTG concentrations. If you do not get any colony on any 2×TY plates even in the absence of IPTG, then switch to an electroporation protocol. In all cases, check that you do not have any colony on any IPTG plate. If you have the same number of colonies in the presence and in the absence of IPTG, the expression of your protein is not toxic or is partially toxic but you cannot select mutants. If you get hundreds of colonies in the absence of IPTG but very few in the presence of IPTG, some mutants may appear at high frequency. To make sure that you do not carry any contamination, repeat the experiment from a single colony culture after transformation of your bacterial host with a freshly prepared plasmid.

Culture and Mutant Isolation You should have 50 ready-to-use plates, supplemented with IPTG and antibiotic. Make sure the plates are not wet (incubate them for 16 h at 37°C). Prewarm five 250-ml flasks containing 50 ml 2×TY medium with antibiotic and inoculate each flask with one bacterial colony to perform five independent selection experiments the same day. Measure the optical density at 600 nm

every 30 min starting 2 h after inoculation. Meanwhile, in sterile conditions, label 40 clean and autoclaved Eppendorf tubes and add 900 μl of sterile water in each tube to perform 1/10 serial dilutions of each culture. Water is preferable to $2\times\text{TY}$ to avoid external contamination. Once the culture has reached 0.4–0.6 $\text{OD}_{600\text{ nm}}$, induce the expression of the target gene by adding IPTG at 1 mM final concentration. One hour after induction, transfer 1 ml of the culture into a new clean and sterile Eppendorf tube and gently spin it down for 2 min at 300 g. Discard the supernatant (secreted β -lactamase often gives false positive colonies when the culture is plated without dilution) and resuspend the pellet in 1 ml sterile water. Perform serial 1/10 dilutions until 10^{-4} and immediately plate 100 μl of all dilutions on IPTG and antibiotic-containing plates. Repeat the experiment 2 h after induction. The purpose of diluting the culture is that it is critical to have less than 200 colonies on a plate so that individual colonies can easily be sub-cloned and isolated. Given that the number of mutant hosts appearing on the plate is difficult to predict, it is safer to have extended dilutions. The frequency of appearance of mutant hosts varies from $1/10^{-4}$ to $1/10^{-6}$ (Miroux and Walker 1996). A 1/100 dilution is often the best compromise.

After an overnight incubation at 37°C (or at a lower temperature for thermo-sensitive mutants), estimate the number of colonies of different sizes. Typically, you should see a majority of large colonies, which, in most of cases, have lost the ability to express the target gene. Small colonies arise at a frequency of 1–20%. If you do not see any obvious difference in sizes between colonies, there are two plausible explanations: (1) The cells need to grow for a longer period of time; leave the plates for 8 additional hours at 37°C to reveal mutant clones of smaller sizes. (2) The cells that have lost the expression of the target gene divide faster, rapidly overgrow the culture and outcompete bacterial mutants that form small colonies. In this case, repeat the selection experiment and plate the culture shortly after induction (20–30 min) to avoid “dilution” of small colonies on the plates.

Figure 4.4 provides examples of selection experiments with the GFP as a reporter gene. Panel a shows the typical size difference between mutant hosts. Panel b shows the same plate under UV exposure. Almost all the small colonies are green and therefore express high amounts of GFP. Large colonies exhibit no or weak fluorescence. Panel c shows a selection experiment where all colonies are small. Among them, some exhibit very high fluorescence intensity. Panel d shows that, in this experiment, medium colonies are fluorescent while the very small ones are not. If you do not have GFP co-expressed or in fusion with your target membrane protein, then incubate 50 2-ml $2\times\text{TY-Amp}$ tubes, each containing one small colony (ten small colonies per selection experiment), and make an over-day culture. When the culture is turbid (2–3 h after inoculation), add IPTG (1 mM final concentration) and induce synthesis of your target protein overnight. The next morning, run 10 μl of the overnight culture on a sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and check the expression of your membrane protein either by immuno-detection with a specific antibody against your protein (or against a tag) or simply by staining the gel with Coomassie blue.

Once you have isolated interesting mutant hosts, you have to cure the strain from the plasmid and check if the mutation is within the bacterial or the plas-

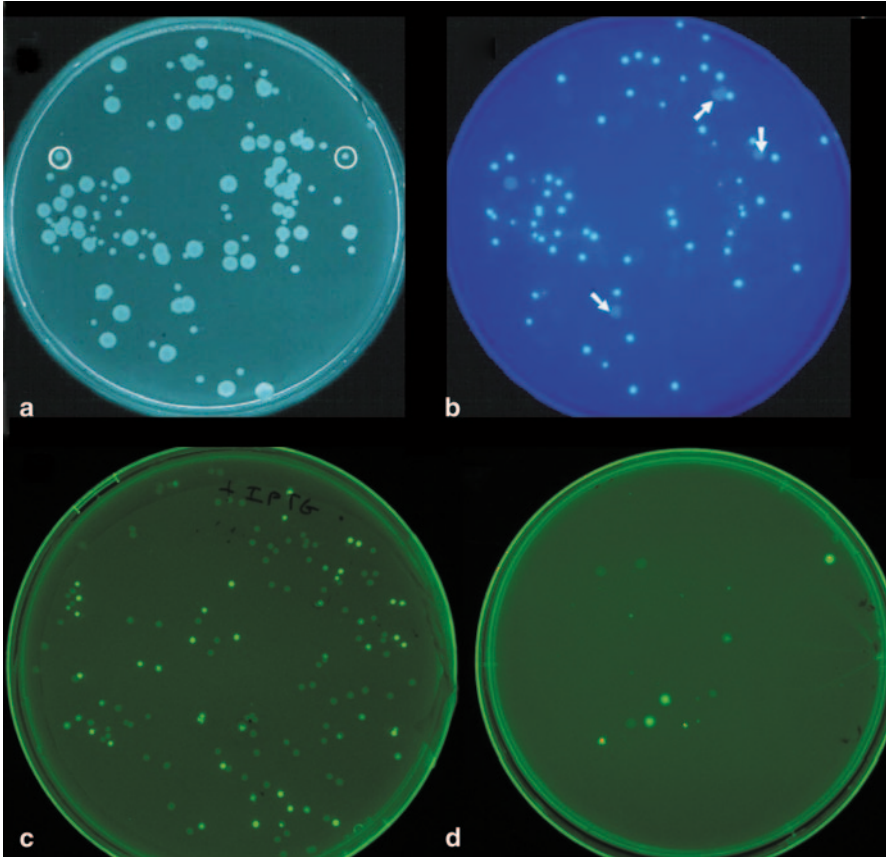


Fig. 4.4 Selection of bacterial hosts using GFP as gene reporter. Isolation of bacterial mutant hosts was performed according to the protocol in Sect. 4.3.2 and to Miroux and Walker (1996) and Walker and Miroux (1999). Briefly, pMW7-GFP-Xa expression plasmid was transformed into BL21(DE3) host (**a**, **b** and **c**) or into C41(DE3) host (**d**) and a single colony was inoculated in 50-ml $2 \times$ TY medium. At $OD_{600\text{ nm}}=0.4$, cells were diluted in water and 100 μ l of the 1/10 dilutions were plated on IPTG-containing plates. The plates are illuminated under normal light (**a**) or UV light (**b**, **c** and **d**).

mid DNA (Fig. 4.5). To do so, prepare a miniprep of plasmids from each clone, transform them into the BL21(DE3) reference host and check colony formation on IPTG-containing plates (left panel). If you obtain colonies, then the mutation is within the plasmid; if not, then the bacterial host carries the mutation. To cure the strain from the plasmid, the easiest method is to wait for spontaneous loss of the plasmid (right panel). Inoculate a 50-ml $2 \times$ TY culture without antibiotic with one single colony and maintain the culture for a week by transferring every day 100 μ l of the culture into a new Erlenmeyer containing 50 ml fresh medium. Every day make serial dilutions of the new culture until 10^{-8} and plate 100 μ l of the 10^{-6} , 10^{-7} and 10^{-8} dilutions on IPTG-containing $2 \times$ TY plates without antibiotic. Since your

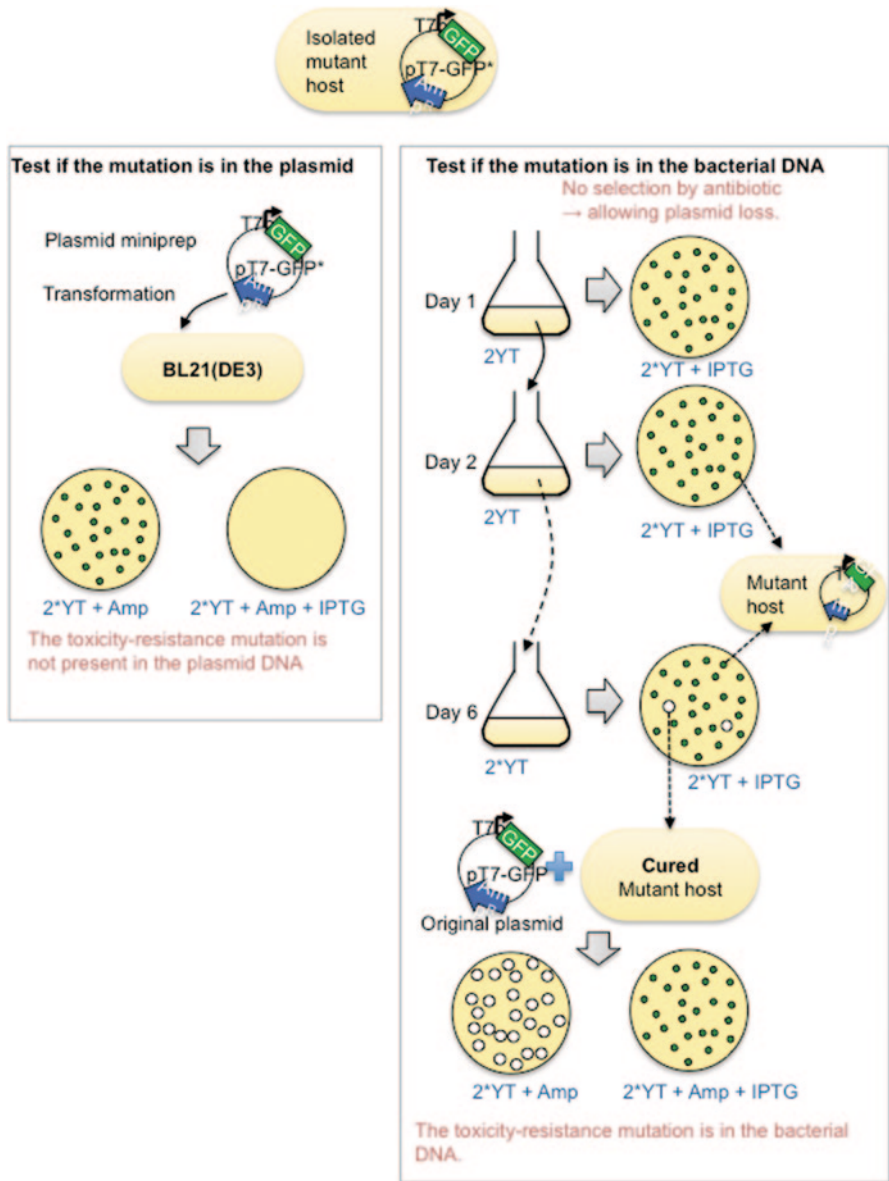


Fig. 4.5 Localization of the mutation in the isolated bacterial mutant host. This step is performed according to the protocol in Sect. 4.3.2, Miroux and Walker (1996) and Walker and Miroux (1999). Basically, your isolated mutant strain has to be cured from the expression plasmid (here, $pT7-GFP^*$) to check that the mutation is present in its genomic DNA (and not in the plasmid DNA). *Left panel:* the plasmid $pT7-GFP^*$ is rescued from the mutant strain and transformed into the original BL21(DE3) host. The transformation is plated onto $2 \times$ YT plates with antibiotic, with and without IPTG. If no colonies are formed in the presence of IPTG, this means that the expression of the target gene from this plasmid is still toxic to BL21(DE3) and, therefore, that the mutation that removed the toxicity is absent from the plasmid. *Right panel:* in parallel, the mutant strain is cured

mutant forms small colonies on these plates, cells that have lost the plasmid over the 1-week-time culture period should form large colonies (Fig. 4.5). Isolate two of those colonies and check that they are antibiotic sensitive. Prepare competent cells from these isolated mutant hosts, transform your reference expression plasmid and plate half of the competent cells on a $2 \times$ TY plate with antibiotic and the other half on a $2 \times$ TY plate supplemented with both antibiotic and IPTG. You should see the same number of colonies on both plates, the IPTG-containing plates carrying only small ones.

4.3.3 *Tuning Growth Conditions*

This guideline is adapted from previous reviews (Shaw and Miroux 2003; Zoonens and Miroux 2010) and enriched with rules from a large-scale bibliographic analysis of the T7-based expression system that we have recently conducted (Hattab et al. 2014). The protocol is divided into two parts, depending on the toxicity of the expression system. For simplicity, we only refer to the T7-based expression system but most advices that are given below can be applied to expression systems other than T7 based.

4.3.3.1 **Expression of the Target Gene is Toxic**

Despite the toxicity of the target gene, it is possible to optimise the expression level of the corresponding protein by adjusting growth conditions.

Plasmid Stability Transform your expression plasmid on a $2 \times$ TY plate with antibiotic and prepare five individual 2-ml precultures from independent colonies. After overnight growth, make serial 1/10 dilutions and plate 100 μ l of 10^{-6} , 10^{-7} and 10^{-8} dilutions on $2 \times$ TY plates with and without antibiotic. If you get the same number of colonies with or without the presence of antibiotic, then the plasmid is stable and you can proceed with larger cultures. If the number of colonies is increased on $2 \times$ TY plate without antibiotic, then it is unsafe to prepare a large culture from a preculture.

Large-Scale Experiment Start from freshly transformed bacterial cells. Some authors do not plate cells after heat shock but use the whole transformation medium as a preculture (Rogé and Betton 2005). By doing this, they avoid strong

from the plasmid by dropping the selective pressure by antibiotic. Serial cultures are performed for a week, during which each is plated on $2 \times$ YT plates with IPTG, after serial dilutions. Mutants that have lost the plasmid will form large colonies that are no longer GFP positive. These cured mutants are then transformed again with the original expression plasmid pT7-GFP and plated on $2 \times$ YT plates, with and without IPTG. If small colonies are retrieved on both plates, this will mean that the strain contains a mutation in its DNA that allows it to overcome the toxicity associated with the expression of the target gene.

Table 4.3 Optimisation of growth conditions in the IPTG inducible T7 expression system

Size of colonies on IPTG plate ^a	Inoculation	Induction	IPTG concentration	Temperature after induction (°C)
No colony	No preculture ^b	No induction ^c OD _{600 nm} = 1	None 10 µM ^d , 0.1 µM	30 or below
Small (> 10% reduction)	Preculture ^c	OD _{600 nm} < 0.6	0.4 or 1 mM	37 or 25
Minor reduction (< 10%)	Preculture	OD _{600 nm} < 0.4	1 mM	37 or 25

^aTry 1 mM and 0.4 mM of IPTG and check phenotype on plates at 37°C and room temperature

^bIf you need a preculture to grow large volumes or to inoculate a fermenter then check plasmid stability

^cSee (Walse et al. 2008; Fairman et al. 2012)

^dSee (Alfasi et al. 2011) for a complete description of the procedure

^ePre-warm the medium and use the preculture at 1/100 dilution; when the plasmid is stable antibiotic is no longer required in the large culture

variability in the target protein expression level from one colony to the other. Prewarm 500 ml of 2 × TY medium in a 2.5-L Erlenmeyer. This is a critical step if you wish to perform the experiment over a day. An alternate option is to incubate the flasks overnight in a 37°C incubator and to add antibiotic the next morning just before use. Inoculate the warm medium with one single colony and follow the optical density at 600 nm. The culture should reach an optical density of 0.6 in less than 5 h; if not, then the basal level of expression of your target gene is sufficient to severely impair cell growth. In addition to the classical protocol of induction (1-mM IPTG at OD_{600 nm} = 0.6), there are two alternative protocols worth trying (Table 4.3): (1) Do not add IPTG; let the culture grow overnight at 30 or 37°C. This protocol works well when your high copy number plasmid is not regulated (no T7lac promoter or *lacI* gene) and in combination with the regular BL21(DE3) bacterial host without a companion plasmid (pLysS/E). We have found two membrane protein structures where the authors specifically mentioned this protocol (Walse et al. 2008; Fairman et al. 2012). (2) Add IPTG at the beginning of the stationary phase either in trace amount (10 µM) following the “improved protocol” from Alfasi and colleagues (Alfasi et al. 2011) or at a high concentration (1 mM).

4.3.3.2 Expression of the Target Gene is Non-toxic or Moderately Toxic

In this configuration, the induction of the expression of the target gene slows down cell growth but does not compromise colony formation on IPTG plate. The expression plasmid is usually highly stable and, in most cases, you can use a preculture to inoculate large flasks. This is also an ideal configuration for using bioreactor, as you do not have to worry about plasmid loss at high cell density. In several occasions, we have observed that antibiotic use is not required anymore in the culture, provided that you have added antibiotic to the preculture (Shaw and Miroux 2003).

The induction protocol must be adjusted depending on the size of the colonies you get on IPTG plate. If the size reduction is marginal (<10%), it is not necessarily a good sign because it could simply mean that the production of your target membrane protein is very low. To maximize your chances of having high level of expression of the target gene, add 1 mM IPTG at the early stage of the exponential phase ($\leq 0.4 \text{ OD}_{600 \text{ nm}}$). If the size of the colonies is decreased by 10% or more, then add IPTG at $\text{OD}_{600 \text{ nm}}=0.6$ and test the two concentrations that are most frequently used (Hattab et al. 2014): 0.4 and 1 mM (Table 4.3).

4.4 Frequently Asked Questions

1. What are the main differences between C41(DE3) and BL21(DE3) bacterial hosts?

C41(DE3) is a derivative of BL21(DE3), isolated on an IPTG plate upon the expression of oxoglutarate mitochondrial carrier, expressed as inclusion bodies. We initially observed that the level of the oxoglutarate mRNA was ten times decreased in this host, 3 h after the addition of IPTG (Miroux and Walker 1996). The group of de Gier has recently shown that expression of the T7 RNA polymerase is strongly decreased in both C41(DE3) and C43(DE3) hosts, thus explaining the reduction in target mRNA levels. The mutation in C41(DE3) is most likely the replacement of the *lacUV5* promoter located upstream of the T7 RNA polymerase by the genomic wild-type copy of the *lac* promoter (Wagner et al. 2008).

2. Does the C43(DE3) bacterial host produce constitutively internal membranes?

No. Internal membrane proliferation occurs upon overexpression of the b-subunit of the *E. coli* F_1F_o ATP synthase, which was used to isolate the mutant host from C41(DE3). Membrane proliferation has been observed by electron microscopy on bacteria cross section 3 h after induction by IPTG at 37 or 25 °C in $2 \times \text{TY}$ medium. The best pictures were taken after an overnight induction at 25 °C (Arechaga et al. 2000).

3. What is the purpose of decreasing culture temperature upon induction by IPTG?

The main advantage of decreasing the temperature is to slow down the activity of the transcription/translation machinery and consequently cell growth. At 20 °C, *E. coli* does not initiate translation, which helps reducing translational stress. For soluble proteins, it has been shown to increase target protein solubility. It also helps the insertion of MBP fusion proteins into the bacterial membrane. Another reason to reduce culture temperature is to avoid overgrowth of the culture by cells that have lost the expression plasmid. Therefore, decreasing the temperature is highly recommended when the expression of the target membrane protein is toxic. Using C41(DE3) at 25–20 °C is often optimal while overexpression of the target gene below 37 °C in C43(DE3) is unpredictable and gene dependent.

4. Does removing the toxicity by selection of bacterial mutant hosts always increase the yield of expression of the protein?

No, unfortunately. A good example is given by the a-subunit of the *E. coli* F₁F₀ ATP synthase. The *uncB* gene, encoding the a-subunit, is regulated by RNA degradation and, consequently, its expression under the T7 promoter does not increase the amount of *uncB* mRNA or the a-subunit peptide (Arechaga et al. 2003). If your target mRNA is not overexpressed, then try breaking mRNA secondary structures by using silent mutations. A complete synthetic gene might help although internal mRNA degradation sites are difficult to predict. Protein degradation is also frequent but can be overcome by making fusion proteins.

5. Is supplementing tRNA for rare codons useful?

We have found that 10% of membrane protein structures have been solved following expression of the protein in the BL21(DE3)CodonPlus bacterial host. On the one hand, it is not negligible and certainly worth trying. In addition, the group of von Heijne has recently demonstrated that codon optimization is critical in the N-terminal sequence of the protein to ensure a proper initiation of translation (Nørholm et al. 2013). It has also been suggested that rare codons are useful for co-translational folding of the nascent polypeptide (Pechmann and Frydman 2013).

6. Are culture media important for protein expression in C41(DE3)/C43(DE3)?

The bacterial mutant hosts C41(DE3) and C43(DE3) support all classical media (minimal medium, Luria Bertani, LB, 2×TY, terrific broth, TB). However, as a general rule, we have observed that toxicity of expression plasmids is increased in LB medium and decreased in 2×TY medium. As expected, levels of expression of the target gene are increased in 2×TY or TB medium.

Acknowledgments This work was supported by the Agence National de La Recherche (ANR MIT-2M, 2010 BLAN1518), the Centre National de la Recherche Scientifique, and by the “Initiative d’Excellence” programme from the French State (Grant “DYNAMO”, ANR-11-LABEX-0011-01).

References

- Abdine A, Verhoeven MA, Park K-H et al (2010) Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J Magn Reson* 204:155–159. doi:10.1016/j.jmr.2010.02.003
- Alfasi S, Sevastyanovich Y, Zaffaroni L et al (2011) Use of GFP fusions for the isolation of *Escherichia coli* strains for improved production of different target recombinant proteins. *J Biotechnol* 156:11–21. doi:10.1016/j.jbiotec.2011.08.016
- Alkhalfioui F, Logez C, Bornert O, Wagner R (2011) In: Robinson AS (ed) Production of membrane proteins: strategies for expression and isolation. Wiley-VCH, Weinheim, pp 75–108
- Arechaga I, Miroux B, Karrasch S et al (2000) Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F₁F₀ ATP synthase. *FEBS Lett* 482:215–219

- Arechaga I, Miroux B, Runswick MJ, Walker JE (2003) Over-expression of *Escherichia coli* F₁F₀-ATPase subunit a is inhibited by instability of the *uncB* gene transcript. *FEBS Lett* 547:97–100
- Banères J-L, Popot J-L, Mouillac B (2011) New advances in production and functional folding of G-protein-coupled receptors. *Trends Biotechnol* 29:314–322. doi:10.1016/j.tibtech.2011.03.002
- Bocquet N, Nury H, Baaden M et al (2008) X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* 457:111–114. doi:10.1038/nature07462
- Catoire LJ, Damian M, Giusti F et al (2010) Structure of a GPCR ligand in its receptor-bound state: leukotriene B4 adopts a highly constrained conformation when associated to human BLT2. *J Am Chem Soc* 132:9049–9057. doi:10.1021/ja101868c
- Chae PS, Rasmussen SGF, Rana RR et al (2010) Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* 7:1003–1008. doi:10.1038/nmeth.1526
- Chen Y, Song J, Sui S, Wang D-N (2003) DnaK and DnaJ facilitated the folding process and reduced inclusion body formation of magnesium transporter CorA overexpressed in *Escherichia coli*. *Protein Expr Purif* 32:221–231. doi:10.1016/S1046-5928(03)00233-X
- Dong H, Nilsson L, Kurland CG (1995) Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J Bacteriol* 177:1497–1504
- Drew D, Lerch M, Kunji E et al (2006) Optimization of membrane protein overexpression and purification using GFP fusions. *Nat Methods* 3:303–313. doi:10.1038/nmeth0406-303
- Eriksson HM, Wessman P, Ge C et al (2009) Massive formation of intracellular membrane vesicles in *Escherichia coli* by a monotopic membrane-bound lipid glycosyltransferase. *J Biol Chem* 284:33904–33914. doi:10.1074/jbc.M109.021618
- Fairman JW, Dautin N, Wojtowicz D et al (2012) Crystal structures of the outer membrane domain of intimin and invasins from enterohemorrhagic *E. coli* and enteropathogenic *Y. pseudotuberculosis*. *Structure* 20:1233–1243. doi:10.1016/j.str.2012.04.011
- Frelet-Barrand A, Boutigny S, Kunji ERS, Rolland N (2010) Membrane protein expression in *Lactococcus lactis*. *Methods Mol Biol* 601:67–85. doi:10.1007/978-1-60761-344-2_5
- Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177:4121–4130
- Hattab G, Moncoq K, Warschawski DE, Miroux B (2014) *Escherichia coli* as host for membrane protein structure determination: a global analysis. *Biophys J* 106(2, Suppl 1):46a
- Jidenko M, Nielsen RC, Sørensen TL-M et al (2005) Crystallization of a mammalian membrane protein overexpressed in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 102:11687–11691. doi:10.1073/pnas.0503986102
- Miot M, Betton J-M (2011) Reconstitution of the Cpx signaling system from cell-free synthesized proteins. *New Biotechnol* 28:277–281. doi:10.1016/j.nbt.2010.06.012
- Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260:289–298. doi:10.1006/jmbi.1996.0399
- Miroux B, Frossard V, Raimbault S et al (1993) The topology of the brown adipose tissue mitochondrial uncoupling protein determined with antibodies against its antigenic sites revealed by a library of fusion proteins. *EMBO J* 12:3739–3745
- Moffatt BA, Studier FW (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* 49:221–227
- Mouillac B, Banères J-L (2010) Mammalian membrane receptors expression as inclusion bodies in *Escherichia coli*. *Methods Mol Biol* 601:39–48. doi:10.1007/978-1-60761-344-2_3
- Nørholm MHH, Toddo S, Virkki MTI et al (2013) Improved production of membrane proteins in *Escherichia coli* by selective codon substitutions. *FEBS Lett* 587:2352–2358. doi:10.1016/j.febslet.2013.05.063
- Nury H, Renterghem CV, Weng Y et al (2011) X-ray structures of general anaesthetics bound to a pentameric ligand-gated ion channel. *Nature* 469:428–431. doi:10.1038/nature09647
- Oldham RK, Dillman RO (2008) Monoclonal antibodies in cancer therapy: 25 years of progress. *J Clin Oncol* 26:1774–1777. doi:10.1200/JCO.2007.15.7438

- Orriss GL, Runswick MJ, Collinson IR et al (1996) The delta- and epsilon-subunits of bovine F1-ATPase interact to form a heterodimeric subcomplex. *Biochem J* 314(Pt 2):695–700
- Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov* 5:993–996. doi:10.1038/nrd2199
- Park SH, Das BB, Casagrande F et al (2012) Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature* 491:779–783. doi:10.1038/nature11580
- Pechmann S, Frydman J (2013) Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat Struct Mol Biol* 20:237–243. doi:10.1038/nsmb.2466
- Popot J-L, Althoff T, Bagnard D et al (2011) Amphipols from A to Z. *Annu Rev Biophys* 40:379–408. doi:10.1146/annurev-biophys-042910-155219
- Rogé J, Betton J-M (2005) Use of pIVEX plasmids for protein overproduction in *Escherichia coli*. *Microb Cell Fact* 4:18. doi:10.1186/1475-2859-4-18
- Sarkar CA, Dodevski I, Kenig M et al (2008) From the cover: directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci U S A* 105:14808–14813. doi:10.1073/pnas.0803103105
- Sevastyanovich YR, Alfasi SN, Cole JA (2010) Sense and nonsense from a systems biology approach to microbial recombinant protein production. *Biotechnol Appl Biochem* 55:9–28. doi:10.1042/BA20090174
- Shaw AZ, Miroux B (2003) A general approach for heterologous membrane protein expression in *Escherichia coli*: the uncoupling protein, UCP1, as an example. *Methods Mol Biol* 228:23–35. doi:10.1385/1-59259-400-X:23
- Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185:60–89
- Supply P, Wach A, Thinès-Sempoux D, Goffeau A (1993) Proliferation of intracellular structures upon overexpression of the PMA2 ATPase in *Saccharomyces cerevisiae*. *J Biol Chem* 268:19744–19752
- Tate CG (2012) A crystal clear solution for determining G-protein-coupled receptor structures. *Trends Biochem Sci* 37:343–352. doi:10.1016/j.tibs.2012.06.003
- Tifrea DF, Sun G, Pal S et al (2011) Amphipols stabilize the *Chlamydia* major outer membrane protein and enhance its protective ability as a vaccine. *Vaccine* 29:4623–4631. doi:10.1016/j.vaccine.2011.04.065
- Von Meyenburg K, Jorgensen BB, Van Deurs B (1984) Physiological and morphological effects of overproduction of membrane-bound ATP synthase in *Escherichia coli* K-12. *EMBO J* 3:1791–1797
- Wagner S, Baars L, Ytterberg AJ et al (2007) Consequences of membrane protein overexpression in *Escherichia coli*. *Mol Cell Proteomics* 6:1527–1550. doi:10.1074/mcp.M600431-MCP200
- Wagner S, Klepsch MM, Schlegel S et al (2008) Tuning *Escherichia coli* for membrane protein overexpression. *Proc Natl Acad Sci U S A* 105:14371–14376. doi:10.1073/pnas.0804090105
- Walker J, Miroux B (1999) Selection of *Escherichia coli* hosts that are optimized for the overexpression of proteins. In: Demain AL, Davies JE (eds) *Manual of industrial microbiology and biotechnology (MIMB2)*, 2nd edn. ASM, Washington DC
- Walse B, Dufe VT, Svensson B et al (2008) The structures of human dihydroorotate dehydrogenase with and without inhibitor reveal conformational flexibility in the inhibitor and substrate binding sites. *Biochemistry* 47:8929–8936. doi:10.1021/bi8003318
- Warschawski DE (2013) Membrane proteins of known structure determined by NMR. <http://www.drorlist.com/nmr/MPNMR.html>. Accessed 30 Aug 2013
- Way M, Pope B, Gooch J et al (1990) Identification of a region in segment 1 of gelsolin critical for actin binding. *EMBO J* 9:4103–4109
- Weiner JH, Lemire BD, Elmes ML et al (1984) Overproduction of fumarate reductase in *Escherichia coli* induces a novel intracellular lipid-protein organelle. *J Bacteriol* 158:590–596
- White S (2013) Membrane proteins of known 3D structure determined by X-ray crystallography. <http://blanco.biomol.uci.edu/mpstruc/>. Accessed 30 Aug 2013

- Wilkison WO, Walsh JP, Corless JM, Bell RM (1986) Crystalline arrays of the *Escherichia coli* sn-glycerol-3-phosphate acyltransferase, an integral membrane protein. *J Biol Chem* 261:9951–9958
- Wright R, Basson M, D’Ari L, Rine J (1988) Increased amounts of HMG-CoA reductase induce “karmellae”: a proliferation of stacked membrane pairs surrounding the yeast nucleus. *J Cell Biol* 107:101–114
- Zoonens M, Miroux B (2010) Expression of membrane proteins at the *Escherichia coli* membrane for structural studies. *Methods Mol Biol* 601:49–66. doi:10.1007/978-1-60761-344-2_4

Chapter 5

Lactococcus lactis: Recent Developments in Functional Expression of Membrane Proteins

Sana Bakari, François André, Daphné Seigneurin-Berny, Marcel Delaforge,
Norbert Rolland and Annie Frelet-Barrand

5.1 Introduction

5.1.1 Membrane Proteins

Membrane proteins (MPs) are key molecules in the cell and therefore are important targets for pharmaceutical drugs (Lundstrom 2007). They are encoded by approximately 30% of the genome (Wallin and von Heijne 1998); however, only about 400 3D unique structures are today available (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html) compared to more than 40,000 nonredundant structures of soluble proteins. The low number of 3D structures accessible compared to those obtained for soluble proteins is due to the difficulty in obtaining sufficient amounts of functionally folded proteins. This can be explained by the features of MPs: (1) they display various topologies from peripheral to intrinsic polytopic proteins with a high number of transmembrane (TM) helices, (2) their surface is relatively hydrophobic, (3) they need detergents for extraction from cell membrane, and sometimes thus need to be reconstituted into proteoliposomes, (4) they are also often flexible

A. Frelet-Barrand (✉) · S. Bakari · F. André · M. Delaforge
Oxidative Stress and Detoxification Laboratory, UMR-CNRS 8221, Institute of Biology and
Technology of Saclay, SB2SM, and Centre for Nuclear Studies and Université Paris-Suds,
Gif-sur-Yvette, France
e-mail: annie.barrand-frelet@cea.fr

D. Seigneurin-Berny · N. Rolland
Cell & Plant Physiology Laboratory, Institute of Sciences Research and Technologies,
UMR-CNRS 5168, National Institute of Agronomical Research, Centre for Nuclear Studies,
Université Grenoble Alpes, Grenoble, France

I. Mus-Veteau (ed.), *Membrane Proteins Production for Structural Analysis*,
DOI 10.1007/978-1-4939-0662-8_5, © Springer Science+Business Media New York 2014

and unstable, (5) they have to be targeted to cell membrane for a correct folding, (6) they are expressed at very low levels in the cell, and/or (7) they can be arranged in membranes in multimeric species or be functional as a monomer (Junge et al. 2008). Despite high-throughput screens of recent years, prediction of successful MP expressions in heterologous systems still remains a challenge (Lacapère et al. 2007). Different systems from prokaryotic (*Escherichia coli* and *Lactococcus lactis*) and eukaryotic origins (yeasts *Saccharomyces cerevisiae*, *Pichia pastoris*; mammalian or insect cells) as well as cell-free systems revealed to be efficient to produce sufficient amounts of functional MPs (Junge et al. 2008; Freigassner et al. 2009; Marreddy et al. 2011a).

5.1.2 Prokaryotic Expression Systems

The first choice for preliminary trials of heterologous expression is often to use prokaryotic systems. Indeed, they are easy to handle and relatively inexpensive compared to eukaryotic systems. Genetic methods and vector systems are well established. Among them, *E. coli* can be considered as the traditional and oldest bacterial system. It has been developed for many years and a wide variety of plasmids and strains are available, mostly based on isopropyl β -D-1-thiogalactopyranoside (IPTG) induction (Gordon et al. 2008). However, yields of functional MP expressed are often unsatisfactory; this is generally due to the formation of inclusion bodies, the production of endotoxins and proteases by the bacteria, and/or the rapid translation rate (for review, see Schlegel et al. 2010).

5.1.3 *Lactococcus lactis*

Therefore, instead of optimizing *E. coli* for overexpression of problematic MPs, a viable alternative strategy is to look for other expression hosts with distinct properties which may perform better for the protein of interest (Chen 2012). One promising bacterial host is *Lactococcus lactis*. This Gram-positive bacterium has emerged in the past decades as a good alternative for functional expression of prokaryotic and eukaryotic MPs (Kunji et al. 2003; Midgett and Madden 2007; Junge et al. 2008; Morello et al. 2008). This aerotolerant lactic acid bacterium (LAB) grows at 30 °C with a doubling time of 35–60 min (Gasson and de Vos, 1994). Already largely used in the food industry for production of fermented food, its potential as a host for the overexpression of homologous and heterologous proteins has also been explored (Mierau et al. 2005; Morello et al. 2008). It is easy and inexpensive to grow; genetic methods and vector systems are available and well developed. *L. lactis* started to be an interesting alternative expression host, especially for eukaryotic MPs, because of its moderate proteolytic activity, the absence of inclusion body formation and of endotoxin production (Kunji et al. 2003, 2005), and the efficient targeting of MPs into

a single glycolipid cytoplasmic membrane (Kunji et al. 2003; Monné et al. 2005; Bernaudat et al. 2011). Moreover, it allows performing functional studies on intact bacteria and membrane vesicles (Kunji et al. 2003; Mierau and Kleerebezem 2005).

However, the relatively small genome size of *L. lactis* (only 50% as compared to *E. coli*) is correlated to the lack of specific chaperone systems and other auxiliary factors that could be necessary for targeting and correct folding of particular MPs, such as disulfide isomerases (Kunji et al. 2003). Furthermore, the *L. lactis* codon usage has an approximate 65% bias for AT base pairs. This may cause limitations, especially for the high-level expression of larger mammalian MPs as shown for the Erd2, the human receptor of the peptide sequence KDEL (composed of the respective K,D,E,L amino acids; Kunji et al. 2005). On the other hand, this could be helpful if the gene encoding the protein of interest possesses comparable guanine–cytosine (GC) content, as suggested by Mierau and Kleerebezem (2005). Finally, one last difficulty when working with *L. lactis* is in the cloning steps as pointed out by Surade et al. (2006). In order to facilitate and obtain larger number of recombinant clones, different strategies have been developed in the past years in addition to the classical one (see below).

The expression of heterologous proteins in *L. lactis* has been facilitated both by advances in genetic knowledge and by new developments in molecular biology techniques. Using these tools, various vectors containing either constitutive or inducible promoters have been developed to obtain increased levels of proteins and to control their production. They currently constitute the basis of all expression systems in *L. lactis* and other LAB (Pontes et al. 2011).

5.1.4 The NICE System

Among them, the tightly regulated nisin-controlled gene expression (NICE) system is the most commonly used (Mierau and Kleerebezem 2005). This promising and powerful expression system developed for LAB is based on genes involved in the biosynthesis and regulation of the antimicrobial peptide, nisin (product of the *nisA* gene). This 34-amino acid bacteriocin is produced by several strains of *L. lactis* or *Streptococcus uberis* (Lubelski et al. 2008) and extensively used as a natural food preservative (Delves-Broughton et al. 1996). For the exploitation of the auto-induction mechanism of nisin for gene expression, genes for the signal transduction system *nisK* and *nisR* were isolated from the nisin gene cluster and inserted into the chromosome of *L. lactis* subsp. *cremoris* MG1363 (nisin negative), creating the strain NZ9000 (Kuipers et al. 1998; Hasper et al. 2004). When a gene of interest is subsequently placed behind the inducible promoter *PnisA* on a plasmid (de Ruyter et al. 1996a), expression of that gene can be induced by the addition of subinhibitory amounts of nisin (0.1–5 ng/ml) to the culture medium (de Ruyter et al. 1996b; Fig. 5.1).

Being well characterized and highly versatile, the NICE system has been widely used over the last decade for overexpression and functional and structural studies

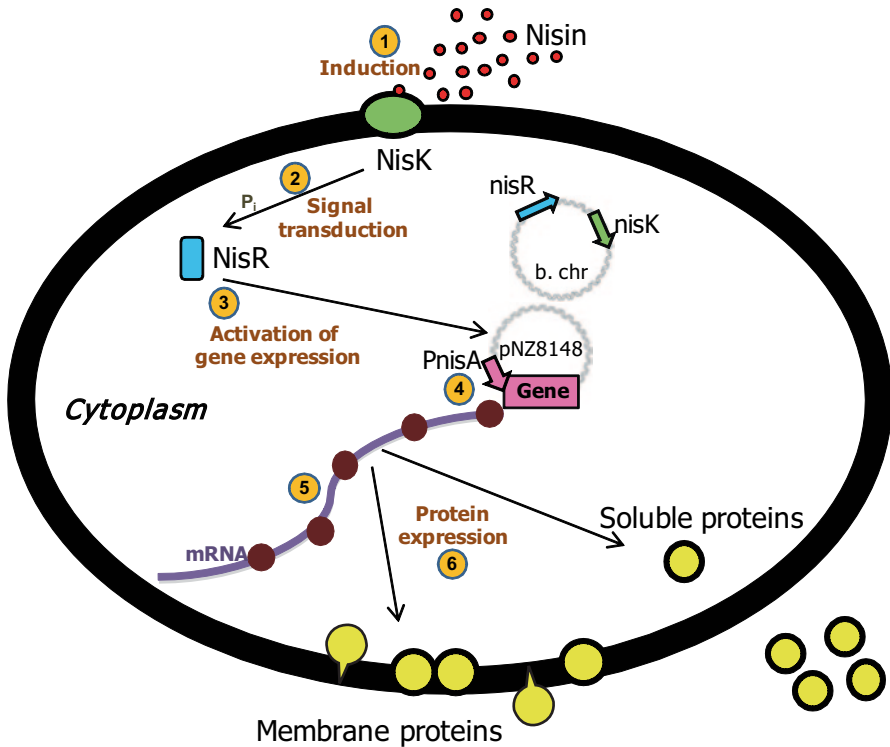


Fig. 5.1 Nisin-controlled gene expression (NICE) system in *L. lactis*. After detection of nisin by the membrane-located sensor protein (*NisK*), this histidine protein kinase autophosphorylates and transfers its phosphate group to activate the cytoplasmic response-regulator *NisR*. Activated *NisR* subsequently induces transcription controlled by *PnisA* promoter. Depending on the presence or absence of the corresponding targeting signals, the protein is either expressed into the cytoplasm or the cell envelope, or secreted into the external medium. *B. chr* bacterial chromosomes. (Adapted from Pontes et al. 2011)

of homologous and heterologous, soluble and membrane proteins (for reviews, see Kunji et al. 2003; Mierau et al. 2005; Zhou et al. 2006). First developed for *L. lactis*, it was subsequently implemented to other Gram-positive bacteria species such as *Leuconostoc lactis* (Kleerebezem et al. 1997), *Streptococcus pyogenes* (Eichenbaum et al. 1998), *Lactobacillus plantarum* (Pavan et al. 2000), or *Lactobacillus brevis* (Avall-Jaaskelainen et al. 2002). Recently, a new NICE system, zinc-regulated expression system (ZIREX), has been described by combining zinc and nisin for better control of the expression of genes; it allows the combination of both inducible promoters for the expression of different proteins at different times in the cell growth (Mu et al. 2013).

5.2 Cloning of cDNA-Encoding Membrane Proteins in *L. lactis*

Prior to expression, the cDNA coding for the protein of interest has to be cloned into the appropriate plasmid suitable for expression, i.e. pNZ8048 or its derivatives for *L. lactis*. In addition to the classical cloning approaches using restriction enzymes, new strategies have been developed recently to overcome the problem of poor cloning efficiencies in *L. lactis* (Geertsma and Poolman 2007).

5.2.1 *Lactococcus* Strains and Plasmids

Table 5.1 gives an overview of *L. lactis* host strains and plasmids used for cloning and expression of cDNAs with the NICE system. All the strains are derivative of *L. lactis* subsp. *cremoris* MG1363, a plasmid-free progeny of the dairy starter strain NCDO712 (Gasson 1983). The nisin-producing strain NZ9700 (Kuipers et al. 1993, 1998) has been obtained by conjugation of the nisin–sucrose transposon Tn5276 of the nisin-A-producer NIZO B8 with MG1464, a rifampicin and streptomycin derivative of MG1363 (Gasson 1983). The early host strain of the NICE system NZ9800, a non-nisin-producing derivative of NZ9700, possesses the necessary regulatory genes (*nisK* and *nisR*). Presently, the most commonly used host is NZ9000 with *nisK* and *nisR* integrated into the *pepN* gene of MG1363 (de Ruyter et al. 1996a; Kuipers et al. 1998) and transcribed from their own constitutive promoter.

Several plasmids have been constructed for translational and transcriptional fusions. The plasmids used for expression in *L. lactis* are based on the pSH71 replicon carrying the chloramphenicol resistance gene (de Ruyter et al. 1996a). pNZ8048 is the most commonly used plasmid for translational fusions. Genes are directly fused to the *NcoI* site, which contains the ATG start codon directly upstream of the *PnisA* promoter. Recently, two variants of pNZ8048 have been constructed: pNZ8148 and pNZ8150. In pNZ8148, a small 60-bp DNA fragment from *Bacillus subtilis*, the initial cloning host of the pSH series of plasmids (de Vos 1987), has been removed. In pNZ8150, the *NcoI* site has been replaced by a *ScaI* site situated directly upstream of the ATG start codon. This improved version of pNZ8148 turns away the obligate use of the *NcoI* site (or other enzymes with compatible overhangs such as *AflIII* and *BspHI*) for translational coupling. Other plasmids and strains are available for protein secretion or for other purposes (Mierau and Kleerebezem 2005; Zhou et al. 2006).

5.2.2 Cloning Strategies

5.2.2.1 Classical Cloning

cDNAs are amplified by polymerase chain reaction (PCR) using the appropriate primers, i.e. containing the *NcoI* restriction site at the start codon of the cDNA in

Table 5.1 Bacterial strains and plasmids commonly used in the NICE system for overexpression of MPs

		Characteristics	References
<i>Strains</i>			
<i>L. lactis</i>	NZ9700	Progeny of the conjugation between nisin producer strain NIZO B8 and MG1614 (Rif ^R Strp ^R derivative of MG1363). Nisin producer strain for induction experiments	Kuipers et al. 1993, 1998 Kunji et al. 2003
	NZ9800	Derivative of NZ9700 with deletion of 4 bp in <i>nisA</i> gene No nisin production but <i>nisRK</i> transcribed. Host of the NICE system	Kuipers et al. 1993, 1998
	NZ9000	MG1363 strain with <i>nisRK</i> integrated into <i>pepN</i> gene. Most commonly used host for NICE system	Kuipers et al. 1998
<i>Plasmids</i>			
	pNZ8048	<i>NcoI</i> site used for translational fusions, Cm ^R	Kuipers et al. 1998
	pNZ8148	pNZ8048 with deletion of 60-bp DNA from <i>B. subtilis</i> , Cm ^R	Mierau and Kleerebezem 2005
	pNZ8150	pNZ8148 with <i>ScaI</i> site used for translational fusions, Cm ^R	Mierau and Kleerebezem 2005

nisA, *nisRK* genes of the nisin operon; *Rif^R*, *Strp^R*, and *Cm^R* resistance to rifampicin, streptomycin, and chloramphenicol, respectively

the forward direction and another site of the multi-cloning site (MCS) of the plasmid (pNZ8048 or derivatives) placed after the stop codon in the reverse direction. Then, after digestion with the same restriction enzymes, cDNAs can be directly ligated into the *L. lactis* expression plasmid (Frelet-Barrand et al. 2010a). This unidirectional cloning allows obtaining a higher number of recombinant clones after transformation. Nevertheless, the MCS site is relatively small, containing less than 10 restriction sites that are not rare and that can be found in the cDNA, requiring, if necessary, partial digestion or mutagenesis of that site (Fig. 5.2).

5.2.2.2 New Strategies to Facilitate Cloning

In the past years, new strategies of cloning have been developed in order to overcome the problem of low efficiency of gene manipulation in *L. lactis* and of instability of *L. lactis*-*E. coli* shuttle vectors (Kok et al. 1984; de Vos and Simons 1994). In order to obtain higher insert-containing plasmids after transformation, Berlec and Strukelj (2012) have developed a TA-cloning expression plasmid. On the other hand, Geertsma and Poolman developed a generic cloning strategy compatible with high-throughput manipulations, suitable for other organisms besides *L. lactis* and applicable to genes coding for soluble proteins (Geertsma and Poolman 2007). Their method involves ligation-independent cloning (LIC) in an

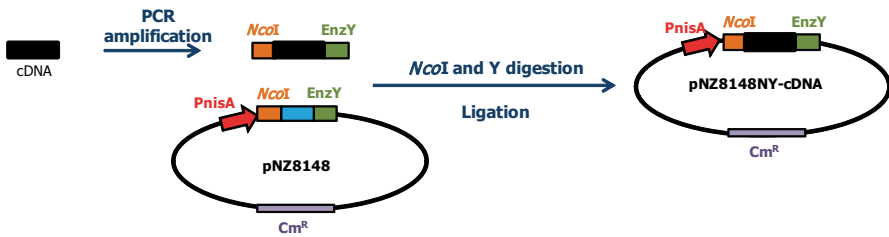


Fig. 5.2 Classical cloning of cDNA in pNZ8148. After amplification by PCR, the cDNA is digested by *NcoI* and enzyme Y (another enzyme from the multi-cloning site). Then, the digested fragment is ligated into pNZ8148 previously digested with the same endonucleases, giving rise to pNZ8148NY-cDNA

intermediary *E. coli* vector (pRExLIC-geneX), which is rapidly converted via vector backbone exchange (VBEx) into an organism-specific plasmid ready for high-efficiency transformation, i.e., pNZxLIC-geneX for *L. lactis*. In order to overcome problems with restriction sites, rare restriction sites (*SwaI* and *SfiI*) were used in both LIC and VBEx procedures. This strategy allowed the successful expression of MPs from prokaryotic and eukaryotic origins (Groeneveld et al. 2010; Erkens et al. 2011; Steen et al. 2011).

Other laboratories developed strategies based on the Gateway technology (Invitrogen). Nowadays, this technology is widely used to simplify cloning of cDNA into many different expression systems from bacteria to eukaryotic systems (Hartley et al. 2000). Douillard et al. (2011) built Gateway-compatible vectors but only for the expression of soluble proteins. On the other hand, this technology has already been used for high-throughput expression screening of integral MPs (Eshaghi et al. 2005). Several libraries are currently available in Gateway-compatible vectors (Yashiroda et al. 2008). However, the *L. lactis* vectors (pNZ8048 or derivatives) cannot be converted into Gateway destination vectors because of the lack of *L. lactis* strains able to propagate Gateway vectors. A strategy with preservation of the correct reading frame has then been established for the rapid transfer of cDNA from Gateway entry vectors into *L. lactis* nisin-inducible vectors (Fig. 5.3; Frelet-Barrand et al. 2010b). First, cDNAs coding for the proteins of interest have to be cloned into Gateway entry vectors by a BP reaction (pDONR-cDNA). Then, a second recombination reaction (LR) allows transferring of the cDNAs into the destination vector (pBS-RfA). After excision of cDNAs by digestion with *EcoRV*, they are ligated into pNZ8148NK (plasmid after *NcoI* digestion and treatment with Klenow). More importantly, for the Gateway recombination sequences being translated in N-terminus of the recombinant proteins, all proteins share the same N-terminal sequence. Therefore, this totally abolishes known impact of the diversity of the very first codons on the production level and stability of the produced recombinant proteins (Grisshammer and Tate 1995; Kunji et al. 2003). The presence of *EcoRV* restriction site(s) in cDNAs could easily be circumvented by either partial enzymatic digestion of shuttle plasmids or mutagenesis of the restriction site, depending

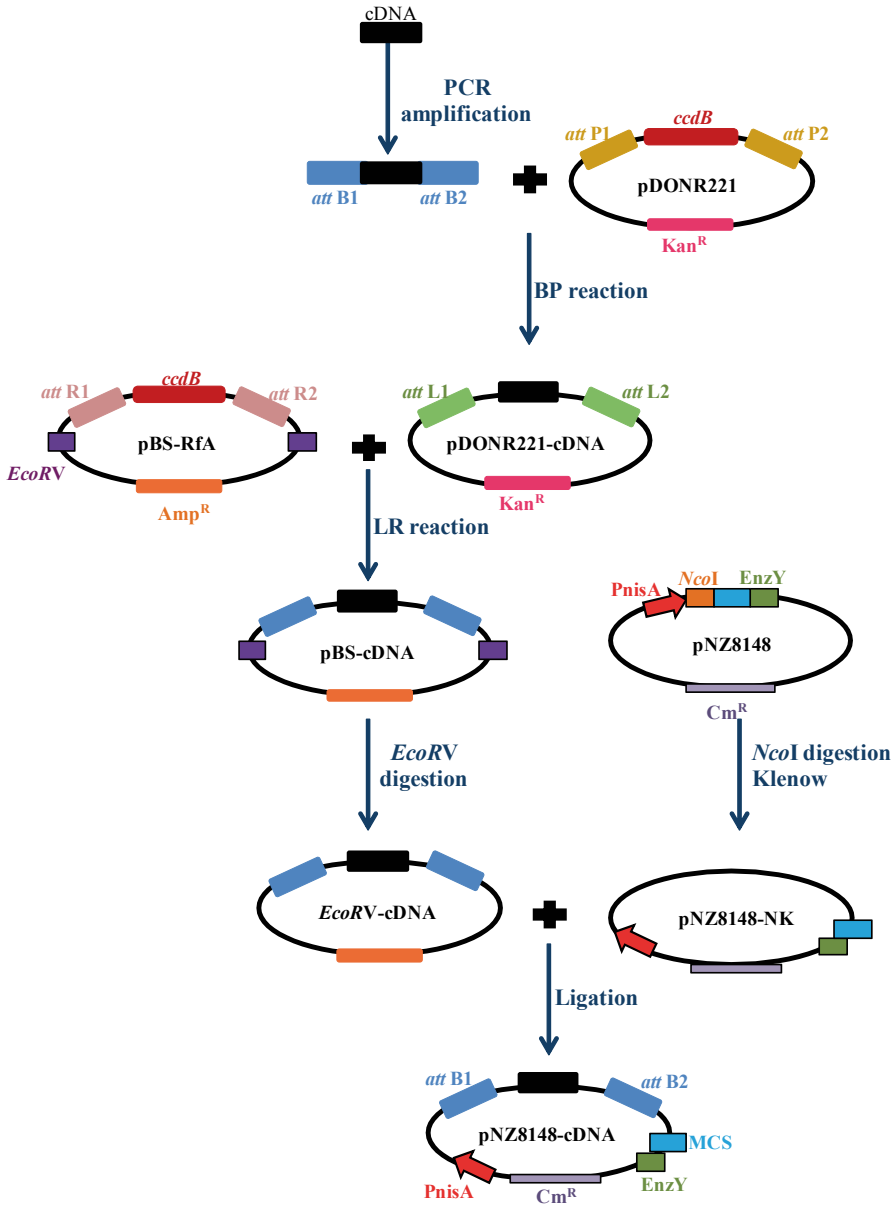


Fig. 5.3 Strategy compatible with Gateway to clone cDNA into pNZ8148. After PCR amplification, the cDNA fragment is inserted into the entry vector pDONR221 by a BP reaction. Afterwards, the cDNA is transferred by an LR reaction into pBS-RfA, and the “shuttle” vector pBS-RfA-cDNA is generated. Then, the cDNA is excised from the pBS-RfA-cDNA by digestion with *EcoRV* and ligated into pNZ8148NK. *Amp^R*, *Kan^R*, and *Cm^R* resistance to ampicillin, kanamycin, and chloramphenicol, respectively; *att X* recombination sites of the Gateway technology; *MCS* multi-cloning site of pNZ8148. (Adapted from Frelet-Barrand et al. 2010b)

on the location of restriction sites on the cDNAs. This strategy allows the successful expression of MPs from prokaryotic and eukaryotic origins (Frelet-Barrand et al. 2010b; Bernaudat et al. 2011).

5.3 Expression of Membrane Proteins Using the NICE System

The NICE system has proved to be highly versatile, and its use in pharmaceutical, medical, and bio- and food-technology fields is most promising (Hugenholtz et al. 2002; Zhou et al. 2006; Morello et al. 2008). Since its construction, numerous homologous or heterologous soluble and membrane proteins, bacteriocins, and antigens have been expressed in *L. lactis* with this system (for review, see Kunji et al. 2003; Zhou et al. 2006). Using pNZ8048 or its derivatives, almost 100 MPs from diverse origins (prokaryotic or eukaryotic), topologies, and sizes have been successfully expressed in the last decade (Kunji et al. 2003; Monné et al. 2005; Zhou et al. 2006; Bernaudat et al. 2011). This system also allows the expression of MP complexes (homodimers or heterodimers; Kunji et al. 2003). Tables 5.2, 5.3 and 5.4 display non-exhaustive lists of prokaryotic MPs (homologous or heterologous expression) and eukaryotic MPs expressed in *L. lactis* with the NICE system. They do not include studies of functionally active proteins in which expression yields were not determined (for example, those of Trip et al. 2013). In addition, in some cases, the functionality of the protein after expression in *L. lactis* could not be examined because of a deficiency in the knowledge of the protein function and/or problems in the ratio of correctly folded protein produced. The tables also do not display the percentage of functional proteins out of the proteins expressed. This information is seldom reported, since such a ratio is difficult to measure and necessitates isolating native proteins as controls.

5.3.1 Expression of Prokaryotic MPs

In Tables 5.2 and 5.3, successful expression of 20 homologous and 25 heterologous MPs is reported, respectively, using the NICE system. Compared to eukaryotic MPs, expression yields of prokaryotic MPs were the highest obtained, up to 30% of total MPs (TMP) with heterologous (HorA and MsbA) and homologous (LmrA) expression (Tables 5.2 and 5.3; Kunji et al. 2003). The proteins expressed possess up to 13 TM helices and, even with high TM helix content, they have been produced with expression yields of up to 20% TMP (BcaP and XylP). Most endogenous MP expression studies have been focused on proteins belonging to the families of amino acid and ATP-binding cassette (ABC) transporters, probably related to the specificities of the laboratories working with this system. Marreddy et al. (2011a) could also overexpress ABC transporters with their corresponding ATP-binding proteins with yields ranging between 1 and 5%. Similar levels of expression could also be

Table 5.2 List of homologous MPs expressed in *L. lactis* using the NICE system

Protein	Function	Size (kDa) ^a	TM helices ^b	Organism ^c	Expression level ^d (%)	Functional ^e	References
MscL	large-conductance mechanosensitive channel	13.8	2	<i>L. lactis</i>	5–10	yes	Folgering et al., 2005
GlnP	ABC transporter	78.5	4	<i>L. lactis</i>	<1	yes	Schuurman-Wolters and Poolman, 2005
BioY	biotin transporter	20.5	5	<i>L. lactis</i>	5	yes	Berntsson et al., 2012
ChoS	glycine betaine ABC transporter permease	55.1	5	<i>L. lactis</i>	2	nd	Marreddy et al., 2011a
GlnP	glutamine transport system permease	78.5	5	<i>L. lactis</i>	2–5	yes	Schuurman-Wolters and Poolman, 2005
ThiT	thiamine transporter	20	6	<i>L. lactis</i>	2	yes	Erkens et al., 2011
RibU	riboflavin transporter	23	6	<i>L. lactis</i>	5	yes	Duurkens et al., 2007
OppC	ABC transporter with OpuB,D,F	32.3	6	<i>L. lactis</i>	<1	yes	Kunji et al., 2003; Doeven et al., 2008
OppB	ABC transporter with OpuC,D,F	35.1	6	<i>L. lactis</i>	<1	yes	
LmrA	ABC efflux pump	65	6	<i>L. lactis</i>	30	yes	Venter et al., 2003
GlnQ	glutamine transporter ATP-binding	27	8	<i>L. lactis</i>	2–5	yes	Schuurman-Wolters and Poolman, 2005
OpuABC	ABC transporter with OpuAA	63	8	<i>L. lactis</i>	10	yes	Kunji et al., 2003
MleP	MFS transporter	46.7	11	<i>L. lactis</i>	1–2	yes	
LmrP	MFS efflux pump	45	12	<i>L. lactis</i>	5	yes	Schaedler et al., 2012
CmbT	MFS transporter	50	12	<i>L. lactis</i>	<1	yes	Filipic et al., 2013
BcaP	branched-chain amino acid permease	50	12	<i>L. lactis</i>	20	nd	Pinto et al., 2011
DtpT	di-/tripeptide transporter	54.8	12	<i>L. lactis</i>	10	yes	Kunji et al., 2003
LmrCD	ABC transporter	63+73.7	6+6	<i>L. lactis</i>	5–10	yes	Seeger et al., 2012
CitP	citrate sodium symporter	48.6	13	<i>L. lactis</i>	1–2	yes	Pudlik and Lolkema, 2012

Species, size, expression yields, and functions are given for each protein; the classification of MPs has been sorted according to the protein complexity in terms of TM helix numbers. Uniprot (<http://www.uniprot.org/>) is used as reference for protein information in addition to literature

^a Protein sizes are given in kDa and for full proteins

^b The number of TM helices listed here has been either already demonstrated or predicted with software (such as TMHMM or psipred) with the FASTA sequence published in Uniprot

^c *L. lactis* (*Lactococcus lactis*)

^d The expression yields are given as a percentage of the recombinant protein compared to the total membrane proteins (TMP)

^e The protein is functionally active in *L. lactis* (yes) or the activity has not been determined (nd)

Table 5.3 List of heterologous prokaryotic MPs

Protein	Function	Size (kDa) ^a	TM helices ^b	Organism ^c	Expression level ^d (%)	Functional ^e	References
CYP201A2	cytochrome-mono-oxygenase	49.7	p	<i>R. palustris</i>	1.5	nd	Bernaudat et al., 2011
NapC	cytochrome-electron transfer	25.6	1	<i>R. sphaeroides</i>	0.5	nd	
MreC	peptidoglycan synthesis	32	1	<i>S. pneumoniae</i>	1	nd	
SP_1241	AA ABC transporter, AA-binding protein/permease	78.4	3	<i>S. pneumoniae</i>	<1	nd	Marreddy et al., 2011a
SAR1949	putative extracellular glutamine-binding protein	53.1	4	<i>S. aureus</i>	1	nd	
L276	mitochondrial carrier-like	27.3	6	<i>A. polyphaga</i>	5	yes	Monné et al., 2007
Lmo2250	ABC transporter	53.1	6	<i>L. monocytogenes</i>	2	nd	Marreddy et al., 2011a
Lin0840	ABC transporter	53.2	6	<i>L. innocua</i>	<1	nd	
Lin2352	ABC transporter	53.4	6	<i>L. innocua</i>	1	nd	
ProWX	ABC transporter permease-choline transporter	55.5	6	<i>S. pneumoniae</i>	2-3	nd	
Lmo1422	binding-protein-dependent transport system permease	55.7	6	<i>L. monocytogenes</i>	1	nd	
Lin1461	binding-protein-dependent transport system permease	55.7	6	<i>L. innocua</i>	2	nd	
SP_0453	AA ABC transporter, AA-binding protein/permease	57.4	6	<i>S. pneumoniae</i>	<1	nd	
CA_C2849	proline/glycine betaine ABC-type transport system, permease	57.6	6	<i>C. acetobutylicum</i>	2	nd	
Jhp0757	putative osmoprotection binding protein	62.6	6	<i>H. pylori</i>	1	nd	
MsbA	lipid A export ATP-binding/permease	64.5	6	<i>E. coli</i>	20-30	yes	Woebking et al., 2005
Sav1866	multidrug export ATP-binding/permease	64.8	6	<i>S. aureus</i>	20-25	yes	Velamakanni et al., 2008

Table 5.3 (continued)

Protein	Function	Size (kDa) ^a	TM helices ^b	Organism ^c	Expression level ^d (%)	Functional ^e	References
abcB	ABC transporter	66	6	<i>B. breve</i>	5–10	yes	Margolles et al., 2006
abcA	ABC transporter	70	6	<i>B. breve</i>	1	yes	
DctA	C4-dicarboxylate transport	45.4	8	<i>B. subtilis</i>	0.5–1	yes	Groeneveld et al., 2010
XyIP	xylose-proton symporter	52.7	12	<i>Lb. pentosus</i>	20	yes	Kunji et al., 2003
LacS	MFS transporter	56.6	12	<i>S. thermophilus</i>	1–2	yes	
T1cA,B,C	ATP/ADP translocator	56.8	12	<i>R. prowazekii</i>	5–10	nd	
HorA	multidrug transporter	64.2	12	<i>L. brevis</i>	30	yes	Sakamoto et al., 2001
TM287/288	ABC transporter	60+60	6+6	<i>T. maritima</i>	0.5–1	yes	Hohl et al., 2012

^a Protein sizes are given in kDa and for full proteins

^b The number of TM helices listed here has been either already demonstrated or predicted with software (such as TMHMM or psipred) with the FASTA sequence published in Uniprot. p for peripheral proteins

^c *R. palustris* (*Rhodospseudomonas palustris*); *R. sphaeroides* (*Rhodobacter sphaeroides*); *S. pneumoniae* (*Streptococcus pneumoniae*); *S. aureus* (*Staphylococcus aureus*); *L. brevis* (*Lactobacillus brevis*); *E. coli* (*Escherichia coli*); *A. polyphaga* (*Acanthamoeba polyphaga*); *L. monocytogenes* (*Listeria monocytogenes*); *L. innocua* (*Listeria innocua*); *C. acetobutylicum* (*Clostridium acetobutylicum*); *H. pylori* (*Helicobacter pylori*); *B. breve* (*Bifidobacterium breve*); *B. subtilis* (*Bacillus subtilis*); *Lb. pentosus* (*Lactobacillus pentosus*); *S. thermophilus* (*Streptococcus thermophilus*); *R. prowazekii* (*Rickettsia prowazekii*); *T. maritima* (*Thermotoga maritima*)

^d The expression yields are given as a percentage of the recombinant protein compared to the total membrane proteins (TMP)

^e The protein is functionally active in *L. lactis* (yes) or the activity has not been determined (nd)

Table 5.4 List of heterologous eukaryotic MPs

Protein	Function	Size (kDa) ^a	TM helices ^b	Organism ^c	Expression level ^d (%)	Functional ^e	References
MPC1/ MPC2	mitochondrial pyruvate carrier	12.3+14.3	2x2	<i>M. musculus</i>	<1	yes	Herzig et al., 2012
SAM5	mitochondrial S-adenosyl methionine carrier	30.9	4	<i>S. cerevisiae</i>	<1	yes	Monné et al., 2005
Mdl1	mitochondrial ATP-dependent permease	76	5	<i>S. cerevisiae</i>	<0.1	yes	Hofäcker et al., 2007
MIR1	mitochondrial phosphate carrier protein	32.8	6	<i>S. cerevisiae</i>	<1	yes	Monné et al., 2005
CTP1	tricarboxylate transport protein	32.9	6	<i>S. cerevisiae</i>	5	yes	
DIC1	mitochondrial dicarboxylate transporter	33	6	<i>S. cerevisiae</i>	10	yes	
GGC1	mitochondrial GTP/GDP carrier protein	33.2	6	<i>S. cerevisiae</i>	4	yes	
PIC2	mitochondrial phosphate carrier protein 2	33.5	6	<i>S. cerevisiae</i>	1–2	yes	Vest et al., 2013
AAC3	mitochondrial ADP/ATP carrier protein 3	33.7	6	<i>S. cerevisiae</i>	5	yes	Kunji et al., 2003
ODC2	mitochondrial 2-oxodicarboxylate carrier 2	34	6	<i>S. cerevisiae</i>	10	yes	Monné et al., 2005
AAC1	mitochondrial ADP/ATP carrier protein 1	34.1	6	<i>S. cerevisiae</i>	<1	yes	
ODC1	mitochondrial 2-oxodicarboxylate carrier 1	34.2	6	<i>S. cerevisiae</i>	8	yes	
AAC2	mitochondrial ADP/ATP carrier protein 2	34.4	6	<i>S. cerevisiae</i>	<1	yes	
AAC1	mitochondrial ADP/ATP carrier protein 1	34	6	<i>H. sapiens</i>	0.5–1	yes	Mifsud et al., 2013
ceQORH	quinone oxidoreductase - electron transfer	33.1	p	<i>A. thaliana</i>	30	yes	Frelet-Barrand et al., 2010b
LPR1	multi-copper oxidase	60.5	p	<i>A. thaliana</i>	<0.1	nd	Bernaudeau et al., 2011
PHF	phosphate transport regulation	42.4	1	<i>A. thaliana</i>	1.5	nd	
AAC hyd	hydrogenosomal carrier	33.9	6	<i>N. patriciarum</i>	<1	yes	Kunji et al., 2003
AHMA1	heavy metal transporter	80.1	6	<i>A. thaliana</i>	3	nd	Frelet-Barrand et al., 2010b
AHMA3	heavy metal transporter	81.4	8	<i>A. thaliana</i>	1	nd	
AHMA6	heavy metal transporter	100	8	<i>A. thaliana</i>	3	yes	
AHMA4	heavy metal transporter	126.7	8	<i>A. thaliana</i>	0.75	nd	Bernaudeau et al., 2011

Table 5.4 (continued)

Protein	Function	Size (kDa) ^a	TM helices ^b	Organism ^c	Expression level ^d (%)	Functional ^e	References
NNT1	chloroplast ADP/ATP transporter	57.5	12	<i>A. thaliana</i>	0.2	yes	Frelet-Barrand et al., 2010b
SUT1	sucrose transporter	54.8	12	<i>S. tuberosum</i>	1–2	nd	Marreddy et al., 2011b
Bcl-Xl	apoptosis regulation	24.7	1	<i>H. sapiens</i>	1	nd	Bernaumat et al., 2011
CYP3A4	cytochrome-mono-oxygenase	57.4	1	<i>H. sapiens</i>	5	nd	Bakari et al., unpublished
MGST1	microsomal glutathione S-transferase 1	17.6	4	<i>H. sapiens</i>	3	nd	
ABCG2	breast cancer resistance protein	72	6	<i>H. sapiens</i>	0.5–1	yes	Janvilisri et al., 2003
Erd2	KDEL receptor	24.4	7	<i>H. sapiens</i>	<0.1	yes	Kunji et al., 2003
CXCR4	chemokine receptor type 4	37.9	7	<i>H. sapiens</i>	<0.1	nd	Bernaumat et al., 2011
CCR5	chemokine receptor type 5	38.7	7	<i>H. sapiens</i>	<0.1	nd	
PS1Δ9	human alpha secretase component	55	9	<i>H. sapiens</i>	0.1–0.2	nd	Marreddy et al., 2011b
CFTR	cystic fibrosis transmembrane conductance regulator	168	12	<i>H. sapiens</i>	<0.1	nd	Steen et al., 2011

^a Protein sizes are given in kDa and for full proteins, i.e., including the transit peptide for mitochondrial and chloroplastic MP (truncated for heterologous expression)

^b The number of TM helices listed here has been either already demonstrated or predicted with software (such as TNHMM or pspired) with the FASTA sequence published in Uniprot. p for peripheral proteins

^c *M. musculus* (*Mus musculus*); *S. cerevisiae* (*Saccharomyces cerevisiae*); *H. sapiens* (*Homo sapiens*); *A. thaliana* (*Arabidopsis thaliana*); *S. tuberosum* (*Solanum tuberosum*); *N. patriciarum* (*Neocallimastix patriciarum*)

^d The expression yields are given as a percentage of the recombinant protein compared to the total membrane proteins (TMP)

^e The protein is functionally active in *L. lactis* (yes) or the activity has not been determined (nd)

reached with heterologous proteins from other bacteria (Table 5.3). In addition to the above-mentioned amino acid and ABC transporters, other various MPs have been heterologously expressed, belonging to diverse families such as cytochrome, permease, and mitochondrial proteins (Table 5.3; Kunji et al. 2003). The relatively high expression yields obtained with heterologous prokaryotic MPs could be explained by the fact that the codon usage is compatible with AT-rich codon bias (Schleifer et al. 1985), as in the case of the transporters from the prokaryotes *Lactobacillus* and *Rickettsia* (Kunji et al. 2003).

5.3.2 Expression of Eukaryotic MPs

In 2003, Kunji and collaborators first reported expression of four eukaryotic MPs in *L. lactis*. Since then, several other eukaryotic MPs have been expressed, with levels from 0.1 to 10% (Table 5.4), in particular from the mitochondrial carrier superfamily (Monné et al. 2005; Herzig et al. 2012; Mifsud et al. 2013; Vest et al. 2013) but also from other families (Zhou et al. 2006; Bernaudat et al. 2011).

5.3.2.1 Plants

Ten MPs from three plant species, i.e. *Arabidopsis thaliana*, *Symphytum tuberosum*, and *Neocallimastix patriciarum*, have been successfully expressed in *L. lactis*. They belong to different families, from oxidase to transporters (heavy metal, adenosine triphosphate (ATP)/adenosine diphosphate (ADP) or sucrose), and their topologies span from peripheral to intrinsic 12 TM helices (Kunji et al. 2003; Frelet-Barrand et al. 2010b; Bernaudat et al. 2011; Marreddy et al. 2011b). The levels of expression obtained were relatively high, up to 30% (Table 5.4), without modifications of the sequence. These relatively high expression yields allowed performing functional studies on some of them (see below).

5.3.2.2 Yeast (*S. cerevisiae*)

Ten MPs from yeast have also been successfully expressed in *L. lactis*. Two main studies on the mitochondrial carriers (Kunji et al. 2003; Monné et al. 2005) revealed that all the MPs tested could be expressed using the NICE system with yields from 0.5 to 10% (Table 5.4); they were all targeted to the cytoplasmic membrane and were functional. For some of them, expression yields were even improved by rational design of the N-terminus (replacing or truncating these regions or by addition of lactococcal signal peptides; Monné et al. 2005).

5.3.2.3 Human MPs

Human mitochondrial ADP/ATP translocator AAC1 was also expressed and found to be active in *L. lactis* (Mifsud et al. 2013). Other human MPs from diverse families and topologies (1–12 TM helices) have been expressed with yields from almost undetectable (<0.1%) to 1% (Bcl-X1; Table 5.4; Janvilisri et al. 2003; Kunji et al. 2003; Bernaudat et al. 2011; Marreddy et al. 2011b). The famous human ABC transporter cystic fibrosis transmembrane conductance regulator (CFTR) has also been expressed in *L. lactis* at very low levels (below 0.1% of TMP; Steen et al. 2011). This full ABC transporter possesses 12 TM helices and is today the larger-size protein produced in *L. lactis* (168 kDa; Table 5.4). In our laboratory, we could express two human MPs involved in liver detoxification, the cytochrome P450 3A4 (CYP3A4) and the microsomal glutathione S-transferase (MGST1), with expression yields from 2 to 5%, respectively (Bakari et al. unpublished data).

The expression yields obtained for expression of MPs in *L. lactis* are generally lower than those obtained for overexpression of same MPs in *E. coli* (Surade et al. 2006; Bernaudat et al. 2011; Marreddy et al. 2011a). On the other hand, in some cases, expression in *L. lactis* allowed an enhancement of the expression, or the expression of proteins which were in general produced in inclusion bodies in *E. coli*. In most cases, for proteins produced with both bacterial expression systems, yields were almost 10 times lower after expression in *L. lactis* compared to *E. coli*. The reason could reside in a limitation of amino acid import, especially for branched amino acids. This problem could be overcome by supplying the cells with an alternative path, a medium containing the appropriate dipeptides, or by engineering the transport capacity for branched-chain amino acids (Marreddy et al. 2010). Other strategies have been implemented using improved strains engineered for enhancing the recombinant MP expression (Zhou et al. 2006; Linares et al. 2010; Noreen et al. 2011; Pinto et al. 2011) or based on optimization of functional expression, i.e. control of transcription rate, nutrient availability, gene optimization, and/or fusion tags (for review, see Marreddy et al. 2011a).

5.4 Functional Characterization of Membrane Proteins

In this part, we will only point out some examples from either prokaryotic or eukaryotic origins for which functional and/or structural analysis of MPs have been performed after expression in *L. lactis* using the NICE system. These functional characterizations could be performed on: (1) whole bacteria using radioactive substrates, (2) membrane vesicles, (3) proteoliposomes after reconstitution with phospholipids (Kunji et al. 2003, 2005; Mierau and Kleerebezem 2005), and/or (4) solubilized/purified proteins (Quick and Javitch 2007). Nuclear magnetic resonance (NMR) studies were also possible; indeed, supplementation of labeled amino acids could be performed by replacing the M17 medium with a chemically defined medium (CDM; Mierau and Kleerebezem 2005). *L. lactis* strains are auxotrophic for

a number of amino acids that can be added in a labeled form and then integrated into recombinant proteins (Kunji et al. 2003). In addition, selenomethionine could be incorporated into proteins expressed in *L. lactis* (Berntsson et al. 2009) for X-ray structural studies. Moreover, *L. lactis* is able to express MPs in their oligomeric state, homodimer or heterodimer, or even larger oligomers (Kunji et al. 2003). *L. lactis* presents two major advantages over *E. coli* for functional expression: (1) it possesses only one membrane and (2) it does not form inclusion bodies (Kunji et al. 2003). In addition, the genome of MG1363 has been completely sequenced and annotated (Wegmann et al. 2007), allowing deleted or mutated strain generation. For these reasons, *L. lactis* can be considered as a good expression system complementary to others.

5.4.1 Prokaryotic MPs

Since 2003 and the first study of expression of MPs performed by Kunji et al., several prokaryotic MPs from *L. lactis* and other bacteria have been successfully expressed in a functional state using the NICE system in *L. lactis*. Several families have been studied: ABC transporters, secondary transporters, and other amino acid transporters. Transport or ATPase activities could be assayed with radioactive or nonradioactive compounds on intact cells or on detergent-purified proteins with or without reconstitution in proteoliposomes. In some cases, mutants were designed and studied to assign the role of certain amino acids in the proper function of the proteins.

The ABC half-transporter LmrA (65 kDa, 6 TM helices) from *L. lactis* has been homologously expressed in high levels, up to 30% of TMP (Table 5.2; Venter et al. 2003). Shilling et al. (2005) could assign the critical role of a carboxylate group in proton conduction to secondary-active transporters. Studies were also performed on mutated versions expressed either in wild-type strains or in strains deleted of LmrA homologs (LmrCD; Lubelski et al. 2007). Moreover, using site-directed spin labeling and pulsed electron–electron double resonance (PELDOR/DEER) spectroscopy, Hellmich et al. (2012) could probe the reorientation of the nucleotide-binding domains and of the TM helix 6, and revealed the conformational changes occurring between the different steps of ATP binding and hydrolysis.

A second half ABC transporter, MsbA, from *E. coli*, has also been expressed in *L. lactis*. The expression yield obtained was a little bit lower than those obtained with the homologous expression of LmrA (20–30%; Table 5.3). This homodimeric transporter of 6 TM helices and 64 kDa is involved in lipid A export in *E. coli* (Woebking et al. 2005). Functional studies similar to LmrA have demonstrated that the substrate binding to MsbA dimer caused nucleotide-binding domain (NBD) dimerization (Woebking et al. 2008; Doshi et al. 2010; Doshi and van Veen 2013). Functional studies on ABC transporters were in fact facilitated in *L. lactis* as compared to *E. coli* because of the availability of strains deleted in *LmrACD*.

A heterodimeric ABC exporter TM287/288, from *Thermotoga maritima*, has also been expressed in *L. lactis*. TM287 and TM288, 60 kDa and 6 TM helices each, form a functional heterodimer sharing 36 % of sequence identity with LmrCD, the well-characterized heterodimeric ABC exporter from *L. lactis* (Lubelski et al. 2006), therefore facilitating their functional studies. Despite the low expression yield of 1 % (Table 5.3), a crystal structure of the heterodimer in its inward-facing state has been obtained. In contrast to previous studies, this structure revealed that the NBDs are partially associated, remaining in contact through an interface involving conserved motifs that connect the two ATP hydrolysis sites (Hohl et al. 2012).

The thiamine high-affinity ABC transporter ThiT (20 kDa, 6 TM helices), belonging to the family of energy coupling factors (ECFs), has been characterized in *L. lactis*. The expression yield in *L. lactis* was around 1–2 % (Table 5.2; Erkens and Slotboom 2010). Mutagenesis studies allowed the determination of some amino acids interacting with the energizing module, necessary for vitamin translocation (Erkens et al. 2011). In 2013, electronic paramagnetic resonance (EPR) performed on purified ThiT and molecular dynamic studies allowed detailed description of the conformational changes of the protein during binding and coupling with the energizing module (Majsnerowska et al. 2013).

In addition to ABC transporters, secondary transport proteins were also expressed in *L. lactis* and characterized (Ter Horst and Lolkema 2010; Trip et al. 2013). These proteins are involved in the transport of amino acids (Trip et al. 2013), and organic or inorganic anions, through symport or exchange processes (Ter Horst et al. 2010). While the quantity of protein produced in these studies was not determined, biological activities of proteins were however detected using substrates specific to the transporters.

5.4.2 Eukaryotic MPs

Several eukaryotic MPs have been successfully expressed in *L. lactis*. The origin of MPs varies from plant to human proteins and yields range from almost undetectable (<0.1 %) to 30 % (Table 5.4; Kunji et al. 2003).

5.4.2.1 Plants

Expression in *L. lactis* using the NICE system proved to be efficient for functional expression of several plant MPs involved in different chloroplast pathways, i.e. the ceQORH, HMA6, and NTT1 proteins from *Arabidopsis thaliana*. ceQORH is a peripheral protein which interacts with the chloroplast envelope through electrostatic interactions (Miras et al. 2002). In *E. coli*, it was mainly produced in inclusion bodies (Miras et al. 2002); in contrary, in *L. lactis*, it was expressed at almost 30 % of TMP (Table 5.4; Frelet-Barrand et al. 2010b), a relatively high expression yield, similar to those obtained for homologous expression of prokaryotic MPs in *L. lactis*

(Kunji et al. 2003). Functional characterization was performed on purified proteins reconstituted in proteoliposomes. It revealed that the ceQORH protein produced in *L. lactis* has nicotinamide adenine dinucleotide phosphate reduced (NADPH)-dependent dehydrogenase activity, and that this activity requires a lipid environment. Moreover, the ceQORH protein produced in *L. lactis* thus behaves as the natural chloroplast envelope protein and appears to interact with the bacterial membrane through electrostatic interactions (Frelet-Barrand et al. 2010b).

Other chloroplast MPs from the P_{1B}-type ATPase family have also been successfully expressed in *L. lactis* using the NICE system with expression yields from 0.7 to 3% of TMP (Table 5.4; Frelet-Barrand et al. 2010b; Bernaudat et al. 2011). These multispinning MPs (6–8 TM helices) translocate ions across plasma or organelle membranes at the expense of ATP consumption (Kühlbrandt 2004). They are involved in the control of metal homeostasis within the cell. Among the eight P_{1B}-type ATPases encoded by the *Arabidopsis* genome, four have been successfully expressed in *L. lactis* (Frelet-Barrand et al. 2010b). Previous attempts for expressing these ATPases in different expression systems (*E. coli*, insect cells, and yeast) were either not successful or disappointing because of the very low production level obtained (Eren et al. 2006; Seigneurin-Berny et al. 2006; Bernaudat et al. 2011). Biochemical characterization using phosphorylation assays were performed on *L. lactis* membrane expressing HMA6 and allowed identification of this protein as a high-affinity Cu⁺ transporter of the chloroplast envelope (Catty et al. 2011).

The NTT1 protein is one of the adenylate translocators identified in the chloroplast which imports ATP in exchange with ADP. This transporter has already been functionally characterized after expression in *S. cerevisiae* and *E. coli* (Neuhaus et al. 1997; Tjaden et al. 1998). In *L. lactis*, NTT1 was produced at a very low expression yield, around 0.2% of TMP. Despite this low expression, uptake assays of radioactive nucleotides could be performed on intact recombinant *L. lactis* cells and showed a time-dependent uptake of ATP with a rate similar to the one measured in *E. coli* cells (Frelet-Barrand et al. 2010b).

To conclude, *L. lactis* appears to be an appropriate expression system for functional characterization of *Arabidopsis* MPs, especially for chloroplast MPs. This could be explained by the fact that the *L. lactis* membrane contains glycolipids (Oliveira et al. 2005) like the inner membrane of the chloroplast envelope (Block et al. 2007), in contrast to *E. coli* membranes (Ingram 1977). The importance of the lipid composition of host cells in overexpression of functional MPs has already been underlined (Opekarova and Tanner 2003; Junge et al. 2008).

5.4.2.2 Mitochondrial MPs

First, two mitochondrial carriers from *S. cerevisiae*, CTP1 and AAC3, have been successfully expressed with yields of 5% and were shown to be functionally active (Kunji et al. 2003). Subsequently, 10 other carriers from *S. cerevisiae* have been successfully expressed with yields ranging from 1 to 10% and activities varying depending on the substrate and the protein studied (Monné et al. 2005). In 2013,

Vest et al. have successfully expressed a functional mitochondrial phosphate carrier PIC2. The relatively high expression yields obtained could be linked to the presence of cardiolipin in the membrane of *L. lactis* (32%; Mifsud et al. 2013). Indeed, the expression is facilitated and the presence of the appropriate lipids could help to drive the protein folding to the right conformation.

In addition, two proteins, MPC1 and MPC2 from *Mus musculus*, have been co-expressed in *L. lactis* and reconstituted as a functional heterodimer. This mitochondrial carrier was able to transport pyruvate across the membrane in intact recombinant bacteria (Herzig et al. 2012). This uptake was sensitive to the mitochondrial pyruvate carrier inhibitor UK5099 and to 2-deoxyglucose, which collapses the proton electrochemical gradient. Moreover, artificially increasing the membrane potential by lowering the pH in the import buffer from 7.2 to 6.2 significantly increased pyruvate uptake. Thus, co-expression of mMPC1 and mMPC2 in the membrane of *L. lactis* was sufficient to allow import of pyruvate with properties similar to the mitochondrial pyruvate carrier (Halestrap 1978).

5.4.2.3 Human MPs

The first human MP produced in *L. lactis* was the KDEL receptor, Erd2. This protein of 7 TM helices is involved in the retrieval of proteins of the endoplasmic reticulum (ER) at later stages of the secretory pathway. While expressed at a very low level, the protein could still bind its specific peptide and conserve the pH-dependent activities as in rat Golgi membranes (Kunji et al. 2003).

The human ATP/ADP translocator (AAC1), which displays TM helix number and size features similar to the other mitochondrial carriers from *S. cerevisiae*, has also been studied after expression in *L. lactis*. The protein was expressed at 0.5–1 % of TMP and was found sensitive to the same inhibitors as its yeast orthologs (Mifsud et al. 2013).

5.4.3 Structure Resolved from Membrane Proteins Expressed in *L. lactis*

The first structure of a polytopic MP from *L. lactis*, ThiT, obtained after expression in *L. lactis* using the NICE system has been published in 2011. Structures have been obtained with both the wild-type and a selenomethionine-labeled protein. This protein is an ECF transporter and the S component involved in thiamin transport. The crystal structure has been obtained, even with an expression yield not better than 2 % of TMP (Erkens et al. 2011). In 2012, the same group resolved the structure of BioY, another *L. lactis* MP from the ECF family involved in biotin transport (Berntsson et al. 2012). Despite the low sequence identity (16 %) between the two proteins, the structures revealed conserved domains, and explained the specificity and high affinity for their corresponding substrates. This opens the road to other

structures in the future since the expression yields obtained for almost all the proteins produced are close to 1–2% or higher and the ability to label with SelenoMet allows resolving of the diffraction data (Berntsson et al. 2009).

5.5 Conclusion

Over the last decade, *Lactococcus lactis* has emerged as an expression system alternative to other bacterial systems. Using the tightly regulated nisin-controlled system (NICE), the number of prokaryotic and eukaryotic MPs from diverse topologies, origins, and functions (up to 100 in 2013) to be successfully expressed in *L. lactis* is exponentially increasing. In some cases, the expression yields obtained, even low, have allowed functional and structural characterizations. Prokaryotic MPs were found easier to express than eukaryotic ones, but preliminary studies performed with them helped to overcome expression bottlenecks. Finally, the crystal structures of two MPs after homologous expression in *L. lactis* have been resolved during the past 2 years and opened the road to others in the future.

Acknowledgments This work has been supported by the Region Ile de France (DIM SeNt, PhD fellowship to SB AAP2010-3-10T6). We thank Edmund Kunji for his kind help and critical reading of the chapter.

References

- Avall-Jääskeläinen S, Kylä-Nikkilä K, Kahala M, Miikkulainen-Lahti T, Palva A (2002) Surface display of foreign epitopes on the *Lactobacillus brevis* S-layer. *Appl Environ Microbiol* 68:5943–5951
- Berlec A, Štrukelj B (2012) Generating a custom TA-cloning expression plasmid for *Lactococcus lactis*. *Biotechniques* 52:51–53
- Bernaudeau F, Frelet-Barrand A, Pochon N, Dementin S, Hivin P, Boutigny S, Rioux JB, Salvi D, Seigneurin-Berny D, Richaud P, Joyard J, Pignol D, Sabaty M, Desnos T, Pebay-Peyroula E, Darrouzet E, Vernet T, Rolland N (2011) Heterologous expression of membrane proteins: choosing the appropriate host. *PLoS ONE* 6:e29191
- Berntsson RP, Alia Oktaviani N, Fusetti F, Thunnissen AM, Poolman B, Slotboom DJ (2009) Selenomethionine incorporation in proteins expressed in *Lactococcus lactis*. *Protein Sci* 18:1121–1127
- Berntsson RP, ter Beek J, Majsnerowska M, Duurkens RH, Puri P, Poolman B, Slotboom DJ (2012) Structural divergence of paralogous S components from ECF-type ABC transporters. *Proc Natl Acad Sci U S A* 109:13990–13995
- Block MA, Douce R, Joyard J, Rolland N (2007) Chloroplast envelope membranes: a dynamic interface between plastids and the cytosol. *Photosynth Res* 92:225–244
- Catty P, Boutigny S, Miras R, Joyard J, Rolland N, Seigneurin-Berny D (2011) Biochemical characterization of AthMA6/PAA1, a chloroplast envelope Cu(I)-ATPase. *J Biol Chem* 286:36188–36197
- Chen R (2012) Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnol Adv* 30:1102–1107

- Delves-Broughton J, Blackburn P, Evans RJ, Hugenholtz J (1996) Applications of the bacteriocin, nisin. *Antonie Van Leeuwenhoek* 69:193–202
- de Ruyter PG, Kuipers OP, Beerthuyzen MM, Alen-Boerrigter I, de Vos WM (1996a) Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. *J Bacteriol* 178:3434–3439
- de Ruyter PG, Kuipers OP, de Vos WM (1996b) Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Appl Environ Microbiol* 62:3662–3667
- de Vos WD (1987) Gene cloning and expression in lactic streptococci. *FEMS Microbiol Lett* 46:281–295
- de Vos WM, Simons GFM (1994) Gene cloning and expression systems in *Lactococci*. In: Gasson MJ, de Vos WM (eds) *Genetics and biotechnology of lactic acid bacteria*. Blackie Academic and Professional, London
- Doeven MK, van den Bogaart G, Krasnikov V, Poolman B (2008) Probing receptor-translocator interactions in the oligopeptide ABC transporter by fluorescence correlation spectroscopy. *Biophys J* 94:3956–3965
- Doshi R, Woebking B, van Veen HW (2010) Dissection of the conformational cycle of the multi-drug/lipidA ABC exporter MsbA. *Proteins* 78:2867–2872
- Doshi R, van Veen HW (2013) Substrate binding stabilizes a pre-translocation intermediate in the ATP-binding cassette transport protein MsbA. *J Biol Chem* 288:21638–21647
- Douillard FP, Mahony J, Campanacci V, Cambillau C, van Sinderen D (2011) Construction of two *Lactococcus lactis* expression vectors combining the Gateway and the NIsin controlled expression systems. *Plasmid* 66:129–135
- Duurkens RH, Tol MB, Geertsma ER, Permentier HP, Slotboom DJ (2007) Flavin binding to the high affinity riboflavin transporter RibU. *J Biol Chem* 282:10380–10386
- Eichenbaum Z, Federle MJ, Marra D, de Vos WM, Kuipers OP, Kleerebezem M, Scott JR (1998) Use of the lactococcal nisA promoter to regulate gene expression in gram-positive bacteria: comparison of induction level and promoter strength. *Appl Environ Microbiol* 64:2763–2769
- Eren E, Kennedy DC, Maroney MJ, Argüello JM (2006) A novel regulatory metal binding domain is present in the C terminus of Arabidopsis Zn²⁺-ATPase HMA2. *J Biol Chem* 281:33881–33891
- Erkens GB, Slotboom DJ (2010) Biochemical characterization of ThiT from *Lactococcus lactis*: a thiamin transporter with picomolar substrate binding affinity. *Biochemistry* 49:3203–3212
- Erkens GB, Berntsson RP, Fulyani F, Majsnerowska M, Vujičić-Žagar A, Ter Beek J, Poolman B, Slotboom DJ (2011) The structural basis of modularity in ECF-type ABC transporters. *Nat Struct Mol Biol* 18:755–760
- Eshaghi S, Hedrén M, Nasser MI, Hammarberg T, Thornell A, Nordlund P (2005) An efficient strategy for high-throughput expression screening of recombinant integral membrane proteins. *Protein Sci* 14:676–683
- Filipic B, Golic N, Jovcic B, Tolinacki M, Bay DC, Turner RJ, Antic-Stankovic J, Kojic M, Topisirovic L (2013) The cmbT gene encodes a novel major facilitator multidrug resistance transporter in *Lactococcus lactis*. *Res Microbiol* 164:46–54
- Folgering JH, Moe PC, Schuurman-Wolters GK, Blount P, Poolman B (2005) *Lactococcus lactis* uses MscL as its principal mechanosensitive channel. *J Biol Chem* 280:8784–8792
- Freigassner M, Pichler H, Glieder A (2009) Tuning microbial hosts for membrane protein production. *Microb Cell Fact* 8:69
- Frelet-Barrand A, Boutigny S, Kunji ER, Rolland N (2010a) Membrane protein expression in *Lactococcus lactis*. *Method Mol Biol* 601:67–85
- Frelet-Barrand A, Boutigny S, Moyet L, Deniaud A, Seigneurin-Berny D, Salvi D, Bernaudat F, Richaud P, Pebay-Peyroula E, Joyard J, Rolland N (2010b) *Lactococcus lactis*, an alternative system for functional expression of peripheral and intrinsic Arabidopsis membrane proteins. *PLoS ONE* 5:e8746
- Gasson MJ (1983) Genetic transfer systems in lactic acid bacteria. *Antonie Van Leeuwenhoek* 49:275–282
- Gasson MJ, de Vos WM (1994) *Genetics and biotechnology of lactic acid bacteria*. Blackie, London

- Geertsma ER, Poolman B (2007) High-throughput cloning and expression in recalcitrant bacteria. *Nat Method* 4:705–707
- Gordon E, Horsefield R, Swarts HG, de Pont JJ, Neutze R, Snijder A (2008) Effective high-throughput overproduction of membrane proteins in *Escherichia coli*. *Protein Expr Purif* 62:1–8
- Grisshammer R, Tate CG (1995) Overexpression of integral membrane proteins for structural studies. *Q Rev Biophys* 28:315–422
- Groeneveld M, Weme RG, Duurkens RH, Slotboom DJ (2010) Biochemical characterization of the C4-dicarboxylate transporter DctA from *Bacillus subtilis*. *J Bacteriol* 192:2900–2907
- Halestrap AP (1978) Stimulation of pyruvate transport in metabolizing mitochondria through changes in the transmembrane pH gradient induced by glucagon treatment of rats. *Biochem J* 172:389–398
- Hartley JL, Temple GF, Brasch MA (2000) DNA cloning using in vitro site-specific recombination. *Genome Res* 10(11):1788–1795
- Hasper HE, de Kruijff B, Breukink E (2004) Assembly and stability of nisin-lipid II pores. *Biochemistry* 43:11567–11575
- Hellmich UA, Lyubenova S, Kaltenborn E, Doshi R, van Veen HW, Prisner TF, Glaubitz C (2012) Probing the ATP hydrolysis cycle of the ABC multidrug transporter LmrA by pulsed EPR spectroscopy. *J Am Chem Soc* 134:5857–5862
- Herzig S, Raemy E, Montessuit S, Veuthey JL, Zamboni N, Westermann B, Kunji ER, Martinou JC (2012) Identification and functional expression of the mitochondrial pyruvate carrier. *Science* 337:93–96
- Hofacker M, Gompf S, Zutz A, Presenti C, Haase W, van der Does C, Model K, Tampé R (2007) Structural and functional fingerprint of the mitochondrial ATP-binding cassette transporter Mdl1 from *Saccharomyces cerevisiae*. *J Biol Chem* 282:3951–3961
- Hohl M, Briand C, Grütter MG, Seeger MA (2012) Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nat Struct Mol Biol* 19:395–402
- Hugenholtz J, Sybesma W, Groot MN, Wisselink W, Ladero V, Burgess K, van Sinderen D, Piard JC, Eggink G, Smid EJ, Savoy G, Sesma F, Jansen T, Hols P, Kleerebezem M (2002) Metabolic engineering of lactic acid bacteria for the production of nutraceuticals. *Antonie Van Leeuwenhoek* 82:217–235
- Ingram LO (1977) Changes in lipid composition of *Escherichia coli* resulting from growth with organic solvents and with food additives. *Appl Environ Microbiol* 33:1233–1236
- Janvilisri T, Venter H, Shahi S, Reuter G, Balakrishnan L, van Veen HW (2003) Sterol transport by the human breast cancer resistance protein (ABCG2) expressed in *Lactococcus lactis*. *J Biol Chem* 278:20645–20651
- Junge F, Schneider B, Reckel S, Schwarz D, Dötsch V, Bernhard F (2008) Large-scale production of functional membrane proteins. *Cell Mol Life Sci* 65:1729–1755
- Kleerebezem M, Beerthuyzen MM, Vaughan EE, de Vos WM, Kuipers OP (1997) Controlled gene expression systems for lactic acid bacteria: transferable nisin-inducible expression cassettes for *Lactococcus*, *Leuconostoc*, and *Lactobacillus* spp. *Appl Environ Microbiol* 63:4581–4584
- Kok J, van der Vossen JM, Venema G (1984) Construction of plasmid cloning vectors for lactic streptococci which also replicate in *Bacillus subtilis* and *Escherichia coli*. *Appl Environ Microbiol* 48:726–731
- Kühlbrandt W (2004) Biology, structure and mechanism of P-type ATPases. *Nat Rev Mol Cell Biol* 5:282–295
- Kuipers OP, Beerthuyzen MM, Siezen RJ, de Vos WM (1993) Characterization of the nisin gene cluster nisABTCIPR of *Lactococcus lactis*. Requirement of expression of the nisA and nisI genes for development of immunity. *Eur J Biochem* 216:281–291
- Kuipers OP, de Ruyter PGGA, Kleerebezem M, de Vos WM (1998) Quorum sensing-controlled gene expression in lactic acid bacteria. *J Biotechnol* 64:15–21
- Kunji ERS, Slotboom DJ, Poolman B (2003) *Lactococcus lactis* as host for overproduction of functional membrane proteins. *Biochim Biophys Acta* 1610:97–108

- Kunji ERS, Chan KW, Slotboom DJ, Floyd S, O'Connor R, Monné M (2005) Eukaryotic membrane protein overproduction in *Lactococcus lactis*. *Curr Opin Biotechnol* 16:546–551
- Lacapere JJ, Pebay-Peyroula E, Neumann JM, Etchebest C (2007) Determining membrane protein structures: still a challenge. *Trends Biochem Sci* 32:259–270
- Linares DM, Geertsma ER, Poolman B (2010) Evolved *Lactococcus lactis* strains for enhanced expression of recombinant membrane proteins. *J Mol Biol* 401:45–55
- Lubelski J, de Jong A, van Merkerk R, Agustiandari H, Kuipers OP, Kok J, Driessen AJ (2006) LmrCD is a major multidrug resistance transporter in *Lactococcus lactis*. *Mol Microbiol* 61:771–781
- Lubelski J, Konings WN, Driessen AJ (2007) Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. *Microbiol Mol Biol Rev* 71:463–476
- Lubelski J, Rink R, Khusainov R, Moll GN, Kuipers OP (2008) Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin. *Cell Mol Life Sci* 65:455–476
- Lundstrom K (2007) Structural genomics and drug discovery. *J Cell Mol Med* 11:224–238
- Majsnerowska M, Hänelt I, Wunnicke D, Schäfer LV, Steinhoff HJ, Slotboom DJ (2013) Substrate-induced conformational changes in the S-component ThiT from an energy coupling factor transporter. *Structure* 21:861–867
- Margolles A, Flórez AB, Moreno JA, van Sinderen D, de los Reyes-Gavilán CG (2006) Two membrane proteins from *Bifidobacterium breve* UCC2003 constitute an ABC-type multidrug transporter. *Microbiology* 152:3497–3505
- Marreddy RK, Geertsma ER, Permentier HP, Pinto JP, Kok J, Poolman B (2010) Amino acid accumulation limits the overexpression of proteins in *Lactococcus lactis*. *PLoS ONE* 5:e10317
- Marreddy RKR, Geertsma ER, Poolman B (2011a) Recombinant membrane protein production: past, present and future. In: Brnjas-Kraljević J, Pifat-Mrzljak G (eds) *Supramolecular structure and function 10*. Springer, Heidelberg
- Marreddy RK, Pinto JP, Wolters JC, Geertsma ER, Fusetti F, Permentier HP, Kuipers OP, Kok J, Poolman B (2011b) The response of *Lactococcus lactis* to membrane protein production. *PLoS ONE* 6:e24060
- Midgett CR, Madden DR (2007) Breaking the bottleneck: eukaryotic membrane protein expression for high-resolution structural studies. *J Struct Biol* 160:265–274
- Mierau I, Kleerebezem M (2005) 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. *Appl Microbiol Biotechnol* 68:705–717
- Mierau I, Olieman K, Mond J, Smid EJ (2005) Optimization of the *Lactococcus lactis* nisin-controlled gene expression system NICE for industrial applications. *Microb Cell Fact* 4:16
- Mifsud J, Ravaud S, Krammer EM, Chipot C, Kunji ER, Pebay-Peyroula E, Dehez F (2013) The substrate specificity of the human ADP/ATP carrier AAC1. *Mol Membr Biol* 30:160–168
- Miras S, Salvi D, Ferro M, Grunwald D, Garin J, Joyard J, Rolland N (2002) Non-canonical transit peptide for import into the chloroplast. *J Biol Chem* 277:47770–47778
- Monné M, Chan KW, Slotboom DJ, Kunji ERS (2005) Functional expression of eukaryotic membrane proteins in *Lactococcus lactis*. *Protein Sci* 14:3048–3056
- Monné M, Robinson AJ, Boes C, Harbour ME, Fearnley IM, Kunji ER (2007) The mimivirus genome encodes a mitochondrial carrier that transports dATP and dTTP. *J Virol* 81:3181–3186
- Morello E, Bermúdez-Humarán LG, Llull D, Solé V, Miraglio N, Langella P, Poquet I (2008) *Lactococcus lactis*, an efficient cell factory for recombinant protein production and secretion. *J Mol Microbiol Biotechnol* 14:48–58
- Mu D, Montalbán-López M, Masuda Y, Kuipers OP (2013) Zirex: a novel zinc-regulated expression system for *Lactococcus lactis*. *Appl Environ Microbiol* 79:4503–4508
- Neuhaus HE, Thom E, Möhlmann T, Steup M, Kampfenkel K (1997) Characterization of a novel eukaryotic ATP/ADP translocator located in the plastid envelope of *Arabidopsis thaliana* L. *Plant J* 11:73–82
- Noreen N, Hooi WY, Baradaran A, Rosfarizan M, Sieo CC, Rosli MI, Yusoff K, Raha AR (2011) *Lactococcus lactis* M4, a potential host for the expression of heterologous proteins. *Microb Cell Fact* 10:28

- Oliveira AP, Nielsen J, Förster J (2005) Modelling *Lactococcus lactis* using a genome-scale flux model. *BMC Microbiol* 5:39
- Opekarova M, Tanner W (2003) Specific lipid requirements of membrane proteins—a putative bottleneck in heterologous expression. *Biochim Biophys Acta* 1610:11–22
- Pavan S, Hols P, Delcour J, Geoffroy MC, Grangette C, Kleerebezem M, Mercenier A (2000) Adaptation of the nisin-controlled expression system in *Lactobacillus plantarum*: a tool to study in vivo biological effects. *Appl Environ Microbiol* 66:4427–4432
- Pinto JP, Kuipers OP, Marreddy RK, Poolman B, Kok J (2011) Efficient overproduction of membrane proteins in *Lactococcus lactis* requires the cell envelope stress sensor/regulator couple CesSR. *PLoS ONE* 6(7):e21873
- Pontes DS, de Azevedo MS, Chatel JM, Langella P, Azevedo V, Miyoshi A (2011) *Lactococcus lactis* as a live vector: heterologous protein production and DNA delivery systems. *Protein Expr Purif* 79:165–175
- Pudlik AM, Lolkema JS (2012) Rerouting citrate metabolism in *Lactococcus lactis* to citrate-driven transamination. *Appl Environ Microbiol* 78:6665–6673
- Quick M, Javitch JA (2007) Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proc Natl Acad Sci U S A* 104:3603–3608
- Sakamoto K, Margolles A, van Veen HW, Konings WN (2001) Hop resistance in the beer spoilage bacterium *Lactobacillus brevis* is mediated by the ATP-binding cassette multidrug transporter HorA. *J Bacteriol* 183:5371–5375
- Schaedler TA, Tong Z, van Veen HW (2012) The multidrug transporter LmrP protein mediates selective calcium efflux. *J Biol Chem* 287:27682–27690
- Schlegel S, Klepsch M, Gialama D, Wickström D, Slotboom DJ, de Gier JW (2010) Revolutionizing membrane protein overexpression in bacteria. *Microb Biotechnol* 3:403–411
- Schleifer KH, Kraus J, Dvorak C, Kilpper-Bälz R, Collins MD, Fischer W (1985) Transfer of *Streptococcus lactis* and related streptococci to the genus *Lactococcus* gen. nov. *Syst Appl Microbiol* 6:183–195
- Schuurman-Wolters GK, Poolman B (2005) Substrate specificity and ionic regulation of GlnPQ from *Lactococcus lactis*. An ATP-binding cassette transporter with four extracytoplasmic substrate-binding domains. *J Biol Chem* 280:23785–23790
- Seeger MA, Mittal A, Velamakanni S, Hohl M, Schauer S, Salaa I, Grütter MG, van Veen HW (2012) Tuning the drug efflux activity of an ABC transporter in vivo by in vitro selected DAR-Pin binders. *PLoS ONE* 7:e37845
- Seigneurin-Berny D, Gravot A, Auroy P, Mazard C, Kraut A, Finazzi G, Grunwald D, Rappaport F, Vavasseur A, Joyard J, Richaud P, Rolland N (2006) HMA1, a new Cu-ATPase of the chloroplast envelope, is essential for growth under adverse light conditions. *J Biol Chem* 281:2882–2892
- Shilling R, Federici L, Walas F, Venter H, Velamakanni S, Woecking B, Balakrishnan L, Luisi B, van Veen HW (2005) A critical role of a carboxylate in proton conduction by the ATP-binding cassette multidrug transporter LmrA. *FASEB J* 19:1698–1700
- Steen A, Wiederhold E, Gandhi T, Breitling R, Slotboom DJ (2011) Physiological adaptation of the bacterium *Lactococcus lactis* in response to the production of human CFTR. *Mol Cell Proteomics* 10:M000052MCP200
- Surade S, Klein M, Stolt-Bergner PC, Muenke C, Roy A, Michel H (2006) Comparative analysis and “expression space” coverage of the production of prokaryotic membrane proteins for structural genomics. *Protein Sci* 15:2178–2189
- Ter Horst R, Lolkema JS (2010) Rapid screening of membrane topology of secondary transport proteins. *Biochim Biophys Acta* 1798:672–680
- Trip H, Mulder NL, Lolkema JS (2013) Cloning, expression, and functional characterization of secondary amino acid transporters of *Lactococcus lactis*. *J Bacteriol* 195:340–350
- Tjaden J, Schwöppe C, Möhlmann T, Quick PW, Neuhaus HE (1998) Expression of a plastidic ATP/ADP transporter gene in *Escherichia coli* leads to a functional adenine nucleotide transport system in the bacterial cytoplasmic membrane. *J Biol Chem* 273:9630–9636

- Velamakanni S, Yao Y, Gutmann DA, van Veen HW (2008) Multidrug transport by the ABC transporter Sav1866 from *Staphylococcus aureus*. *Biochemistry* 47:9300–9308
- Venter H, Shilling RA, Velamakanni S, Balakrishnan L, Van Veen HW (2003) An ABC transporter with a secondary-active multidrug translocator domain. *Nature* 426:866–870
- Vest KE, Leary SC, Winge DR, Cobine PA (2013) Copper import into the mitochondrial matrix in *Saccharomyces cerevisiae* is mediated by Pic2, a mitochondrial carrier family protein. *J Biol Chem* 288(33):23884–23892, Accessed 11 July 2013
- Wallin E, von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* 7:1029–1038
- Wegmann U, O’Connell-Motherway M, Zomer A, Buist G, Shearman C, Canchaya C, Ventura M, Goesmann A, Gasson MJ, Kuipers OP, van Sinderen D, Kok J (2007) Complete genome sequence of the prototype lactic acid bacterium *Lactococcus lactis* subsp. *cremoris* MG1363. *J Bacteriol* 189:3256–3270
- Woecking B, Reuter G, Shilling RA, Velamakanni S, Shahi S, Venter H, Balakrishnan L, van Veen HW (2005) Drug-lipid A interactions on the *Escherichia coli* ABC transporter MsbA. *J Bacteriol* 187:6363–6369
- Woecking B, Velamakanni S, Federici L, Seeger MA, Murakami S, van Veen HW (2008) Functional role of transmembrane helix 6 in drug binding and transport by the ABC transporter MsbA. *Biochemistry* 47:10904–10914
- Yashiroda Y, Matsuyama A, Yoshida M (2008) New insights into chemical biology from OR-Feome libraries. *Curr Opin Chem Biol* 12:55–59
- Zhou XX, Li WF, Ma GX, Pan YJ (2006) The nisin-controlled gene expression system: construction, application and improvements. *Biotechnol Adv* 24:285–295

Chapter 6

Overexpression of Membrane Proteins in *Saccharomyces cerevisiae* for Structural and Functional Studies: A Focus on the Rabbit Ca²⁺-ATPase Serca1a and on the Yeast Lipid “Flippase” Complex Drs2p/Cdc50p

Cédric Montigny, Hassina Azouaoui, Aurore Jacquot, Marc le Maire, Christine Jaxel, Philippe Champeil and Guillaume Lenoir

6.1 Introduction

The so-called baker’s yeast, *Saccharomyces cerevisiae*, has been used for centuries by humans to produce food and beverages like bread and beer (or wine). Since its complete genome sequencing in 1996 (Goffeau et al. 1996), we know that it contains about 6,000 genes and that about 25% of them encode putative membrane proteins (MPs). For comparison, human cells contain about 25,000 genes and the same proportion of putative MPs as in yeast (Lander et al. 2001; Venter et al. 2001). Yeast cells share with human cells similar protein synthesis mechanisms, maturation machinery, and membrane-trafficking pathways. As a result, *S. cerevisiae* is probably the best characterized and the most widespread model for studying eukaryotic cell biology.

Genetic engineering of *S. cerevisiae* started in the 1970s with a description of homologous recombination mechanisms, discovery of numerous yeast plasmids (Gunge 1983), and the development of recombinant DNA technologies. This led to its use as a powerful biotechnology tool for protein overexpression, prominent examples being overproduction of recombinant human insulin precursors in 1987 (Novolin R®, Novo Nordisk) or production of the first effective vaccine against human viral infections by hepatitis B (Diers et al. 1991; McAleer et al. 1984). Due to the simplicity of its genome and its ability to perform homologous recombination,

Cédric Montigny and Hassina Azouaoui contributed equally to this work.

G. Lenoir (✉) · C. Montigny · H. Azouaoui · A. Jacquot · M. le Maire · C. Jaxel · P. Champeil
Laboratory of Membrane Proteins, Institute of Biology and Technology of Saclay, UMR-CNRS
8221, Centre for Nuclear Studies and Université Paris-Sud, Gif-sur-Yvette, France
e-mail: guillaume.lenoir@cea.fr

S. cerevisiae genetic manipulation is now straightforward (Wach 1996). Isolation of genes and design of disruption cassettes allowed obtaining thousands of mutants that are available for the international community (EUROpean Saccharomyces Cerevisiae ARchive for Functional Analysis data bank, <http://web.uni-frankfurt.de/fb15/mikro/euroscarf/>). Whole genome and proteome information are available on the Saccharomyces Genome Database website (<http://www.yeastgenome.org/>).

Since the 1990s, *S. cerevisiae* has been used as a heterologous expression system for MPs (for review, Grisshammer and Tate 1995). Yet, despite intense efforts, only a limited amount of high-resolution structures of yeast-expressed eukaryotic MPs have been solved so far (see Table 6.1; White 2013). There is still a long way to go but recent technical advances should help overcome barriers in MP structure determination (Carpenter et al. 2008; Lee and Stroud 2010). Our goal in this chapter is not to give an exhaustive inventory of yeast expression systems, which have been excellently reviewed elsewhere (Britton et al. 2011), but rather to focus on systems used or specifically developed for eukaryotic MP overexpression and which led to the determination of high-resolution structures.

6.2 *S. cerevisiae* as a Host for Heterologous Expression of MPs

S. cerevisiae seems to be an appropriate expression system for large-scale production of MPs. This organism is able to perform most of the eukaryotic posttranslational modifications such as disulfide bond formation, glycosylation, and proteolytic maturation. From a cell-biological point of view, yeast contains almost all internal organelles that are found in mammalian cells (instead of lysosomes, they contain a related compartment called vacuole, which is also found in plants), and expressed MPs may be targeted to these specific membranes (Zimmermann et al. 2011; Weis et al. 2013). As mentioned in the introduction, molecular biology techniques are well documented and fairly simple compared to other eukaryotic expression systems; thus, this organism is well suited for cloning and mutagenesis using classical molecular biology tools. In addition, yeasts are easy to cultivate; growth conditions are well described and scaling up cultures from small volumes to some tenth of liters is relatively straightforward. However, it is important to be aware that *S. cerevisiae* is not a perfect model for mammalian cells. Differences in membrane lipid composition or in the glycosylation machinery may turn out to be critical for MP folding and function. As an example, *N*-glycan structures generated in the Golgi of *S. cerevisiae* are made exclusively of one sugar type, mannose, while mammalian *N*-glycans are much more diverse.

Two main cloning methods have been used for protein overexpression in yeast. The target gene may be cloned on an episomal plasmid that replicates separately from the yeast chromosome; in this case, the number of plasmids per cells may be very low, and therefore selection pressure has to be maintained tightly to prevent plasmid loss during prolonged culture of the cells. Another method consists in taking advantage of the homologous recombination ability of *S. cerevisiae*, by intro-

Table 6.1 (continued)

Name of the protein	ATP synthase	Ca ²⁺ -ATPase	Monoamine oxidase A	H ⁺ -PPase	CAAX protease Ste24p	PiPT	Vex Ip
<i>Crystallization</i>							
Method	Microbatch in paraffin oil	Hanging/sitting drops	Hanging drops	Sitting drops	Hanging drops	Sitting drops	Hanging drops (lipidic cubic phase)
Detergent	DDM	C ₁₂ E ₈	FC-12, APO 10	DDM	DDM	C ₁₂ E ₇	DDM
Lipids	–	DOPC	–	–	–	–	Monolein
Ligands	–	Ca ²⁺ and ATP analogs	–	Inhibitors	Substrate analog	–	Ca ²⁺ and Mn ²⁺
Resolution, Å	3.9	3.0–3.5	3.2	2.2	2.35	3.1	2.7/4–6
PDB entry code	1QO1	–	1O5W	2Z5X	4AO1	4IL3	4J05/- 4K1C

ducing the gene of interest on a plasmid that will subsequently be integrated into the yeast genome (Wach 1996; Decottignies et al. 1998; Nagy et al. 2006). Auxotrophy markers such as *URA3* (uracil), *HIS3* (histidine), or *LEU2* (leucine) are classically used as selection markers, either for maintaining episomal plasmids or for selecting yeasts with modified genomic DNA (Brachmann et al. 1998). Markers conferring resistance to a toxic molecule like geneticin (G418) or hygromycin B may be used too (Decottignies et al. 1998).

6.2.1 Lipids

One general concern with heterologous expression of MPs is the lipid composition of the membrane they are embedded in: interaction of specific lipids has been shown to play an important role in the function of these MPs (Lee and East 1998; Powl et al. 2008; Kapri-Pardes et al. 2011; Whorton and MacKinnon 2011; Haviv et al. 2013). Most of the glycerophospholipids and their derivatives are present in *S. cerevisiae*, although not in identical proportion as in mammalian cells (Blagovic et al. 2001; Blagovic et al. 2005; van Meer et al. 2008; Canadi Juresic and Blagovic 2011).

One of the striking differences between mammalian cells and *S. cerevisiae* is the presence of ergosterol in yeast, a cholesterol analog. For expression of MPs of higher eukaryotes, this may be a disadvantage because the function of several MPs is known to be affected by the nature of the sterol in the membrane phase. For instance, Lagane et al. (2000) demonstrated that ergosterol and cholesterol have opposite effects with respect to the ligand-binding function of a recombinant μ -opioid receptor. However, only a few such examples were reported and some yeast strains have been now engineered to be able to synthesize cholesterol instead of ergosterol (Kitson et al. 2011). Another significant difference resides in the amount of phosphatidylserine (PS) found in the plasma membrane. PS is a minor lipid in most *S. cerevisiae* membranes, except in plasma membrane where it accounts for as much as 30% of the total phospholipids, a higher proportion than that in plasma membranes of mammalian cells (van Meer et al. 2008).

It is important to note that lipid metabolism in yeast is highly variable, depending on the strain used and on culture conditions (Daum et al. 1999). In some cases, overexpression of MPs has also been reported to stimulate membrane biosynthesis and activate quality control mechanisms like unfolded protein response (UPR) and endoplasmic reticulum-associated degradation (ERAD). Such activation suggests that MP trafficking through the secretory pathway may become overwhelmed, leading to accumulation of misfolded neosynthesized MPs and activation of degradation pathways (Griffith et al. 2003; Meusser et al. 2005). For example, Griffith and coworkers showed that an increase in the UPR response is correlated with a dramatic decrease in specific activity of a recombinant P2 adenosine-proton transporter (Griffith et al. 2003). This phenotype is not specific to MP overexpression since it has also been described for a soluble protein (Bleve et al. 2011). Hence, for the sake of increasing the amount of functional protein, the use of adjustable yeast expression systems is probably highly desirable.

6.2.2 Glycosylation

As in other eukaryotes, glycosylation in *S. cerevisiae* starts in the endoplasmic reticulum (ER), where a core structure is transferred to nascent polypeptide chains, and continues in the Golgi apparatus, where this core structure is subsequently modified (Helenius and Aebi 2001). While the core structure is highly similar in all eukaryotes, the nature of *N*-glycan structures found in the Golgi may vary significantly. Glycosylation in yeast is implicated in folding, stability, and function of MPs as in other eukaryotes (Celik and Calik 2012). Because baker's yeast lacks Golgi mannosidases, it is characterized by its ability to hypermannosylate proteins (Gemmill and Trimble 1999), resulting in the addition of more than 40 mannose residues per glycosylation sites. When hypermannosylation occurs, glycosylation is heterogeneous and may affect protein function (Celik and Calik 2012). In mammalian cells, when proteins reach the Golgi apparatus, an appropriate mannosidase removes some of the mannose residues which are then replaced by specific carbohydrates. One option for controlling glycosylation is to remove glycosylation sites by site-directed mutagenesis, but this may result in the biosynthesis of an inactive protein. Several approaches have been attempted to substitute *S. cerevisiae* glycosylation pathways for human pathways and recent advances in the expression of glycoproteins with humanized glycan structures in engineered yeast are promising (Wildt and Gerngross 2005; Chiba and Jigami 2007; Chiba and Akeboshi 2009).

6.2.3 Inventory of Crystal Structures Obtained After Expression in *S. cerevisiae*

In this section, we present an overview of eukaryotic MPs purified from *S. cerevisiae* membranes, used either as a natural source or as a host for overexpression (Table 6.1).

Most of the already crystallized MPs were obtained from natural sources. One famous result was obtained with adenosine triphosphate (ATP) synthase (Stock et al. 1999) from *S. cerevisiae*, an enzyme responsible for ATP synthesis and which is very abundant in mitochondria. In this case, it was not necessary for researchers to design a particular expression system: they simply purified the protein from a block of baker's yeast conventionally used for cooking. Here, the challenge was to purify and crystallize a large MP complex, as ATP synthase is composed of 17 subunits for the complete assembly of the membrane part (F_0), the catalytic part (F_1), and the peripheral stalk (the latter being lost during purification and crystallization procedures). More recently, the *S. cerevisiae* strain was modified to integrate a hexahistidine tag at the N-terminus of four different mutants of the β subunit (Muellet et al. 2004), leading to the crystallization of the F_1 ATPase sub-domain alone (Arsenieva et al. 2010). Crystallization of the whole complex with its F_0 membranous part is in progress (Pagadala et al. 2011).

After heterologous overexpression in *S. cerevisiae*, structures have been obtained for rat and human monoamine oxidase A (rMAOA and hMAOA). The structure of a close isoform, hMAOB, had been previously resolved but after truncation of the C-terminal domain (Binda et al. 2002). Ma and colleagues succeeded in resolving the structure of the whole protein and revealed that the C-terminal 29 residues fold as a helix responsible for anchoring the protein to the mitochondrial outer membrane (Ma and Ito 2002; Ma et al. 2004a, b; Son et al. 2008). Here, yeast cells were transformed with an episomal plasmid carrying the MAOA gene under the control of a galactose-inducible promoter. A classical purification step on a nickel affinity-chromatography resin allowed the authors to obtain about 10 mg of active protein per liter of culture. In this case, particular care was taken with respect to the choice of the detergent for protein solubilization, stabilization, and crystallization: FC-12 proved to be the most efficient detergent for solubilization and stabilization whereas high-resolution diffracting crystals were obtained in the presence of dimethyldecylphosphine. The rMAOA structure represented the first structure of a eukaryotic monotopic MP. Rat and human recombinant enzymes were shown to be isomorphous and displayed close kinetic parameters (Son et al. 2008).

The structure of a plant vacuolar proton-translocating pyrophosphatase was also recently obtained after expression in *S. cerevisiae* (Lin et al. 2012). Expression was performed using a galactose-inducible promoter. The authors chose to overexpress the protein fused to a hexahistidine tag for purification purposes. Note that the same strategy for expression and purification was then used successfully for determining the high-resolution structure of two other PPases from *Thermotoga maritima* (Kellouso et al. 2012a, b), suggesting that this system is particularly efficient for expressing this class of proteins, regardless of the evolutionary distance between them.

Even more recently, Pryor and coworkers resolved the structure of the Ste24p CAAX protease from *S. mikatae*, which is involved in the maturation of the mating pheromone “a-factor” in yeast (Pryor et al. 2013). The gene encoding Ste24p was cloned downstream of an *ADH2* promoter, a strong constitutive promoter subjected to a tight catabolite repression: its transcription level is repressed several 100-fold in the presence of glucose. In the alternative classical *GAL1*–*GAL10*-inducible system, glucose is also a strong repressor of transcription, but in that case, after reaching the stationary phase, galactose must be added to the medium to activate transcription of downstream genes. But it is difficult to know exactly when glucose has been completely consumed and in fact, only traces of remaining glucose are sufficient to repress the *GAL1*–*GAL10* promoter, even in the presence of large amounts of galactose. The use of a constitutive *ADH2* promoter allows to overcome this problem, because expression will start automatically when the glucose concentration is sufficiently low, without any requirement for addition of an extra compound in the medium (Lee and DaSilva 2005). Screening of nine different orthologs of Ste24p identified the one from *S. mikatae* as the most stable and suitable for crystallization trials. For purification of Ste24p, the authors fused an immunoglobulin G (IgG)-binding domain to the C-terminus of Ste24p, as IgG resins have proven to be highly specific compared to Ni²⁺-chelating resins (Waugh 2005).

The above inventory of those crystal structures of integral eukaryotic MPs which have been obtained after expression in *S. cerevisiae* is rather short, and it is therefore not possible to define a general path leading to success. But this is not specific to overexpression in *S. cerevisiae*, since no general rule emerged either from other expression systems (Bill et al. 2011). Another trail could be to investigate new screening strategies. For example, Li and coworkers developed a medium-throughput pipeline to test homologous expression, solubilization, and purification of 384 MPs from *S. cerevisiae* (Li et al. 2009). This system was subsequently adapted to heterologous expression of MPs from higher eukaryotes (Li et al. 2009). A key feature of this approach is to identify MPs for downstream structural studies on the basis of their size-exclusion profiles, which may reflect protein stability, a critical parameter for crystallization trials. The first step consists in selecting targets on the basis of their primary sequence and their putative number of transmembrane helices, in order to focus on integral transmembrane proteins. The corresponding genes are then cloned under the control of a galactose-inducible *GALI* promoter, with N-terminal FLAG and C-terminal decahistidine tags. As MP overexpression sometimes turns out to be toxic, some targets are already eliminated after expression trials (Osterberg et al. 2006). For cell lysis, the authors chose to use glass beads in a bead-beater, and for solubilization tests dodecylmaltoside (DDM) was preferred, as this detergent previously proved to be stabilizing for several MPs (Prive 2007). On the 384 targets that Li and coworkers tested, 6 were pushed into the intensive production phase and were readily purified and stable (Li et al. 2009). So far, this strategy resulted in the publication of high-resolution structures for the fungal phosphate/H⁺ symporter PiPT (Pedersen et al. 2013) and for the Ca²⁺/H⁺ antiporter Vcx1p from *S. cerevisiae* (Waight et al. 2013).

For decades, our laboratory has been interested in unraveling the relationships between structure and function in the sarco(endo)plasmic reticulum Ca²⁺-ATPase (Serca1a) Ca²⁺ pump. Among different approaches, we started developing a heterologous expression system of this integral membrane transporter. The subsequent section focuses on the strategy for heterologous overexpression in *S. cerevisiae* our laboratory developed for the Serca1a Ca²⁺ pump, and which led in 2005 to a high-resolution structure.

6.3 Expression in *S. cerevisiae* and Purification of Sarcoplasmic Reticulum (SR) Ca²⁺-ATPase Serca1a

6.3.1 Structure and Transport Cycle of the SR Ca²⁺-ATPase Serca1a

Serca1a belongs to the P-type ATPase family of membrane pumps, and plays a critical role in muscle relaxation and Ca²⁺ homeostasis, by transporting Ca²⁺ uphill from the cytosol into the SR lumen (Hasselbach and Makinose 1961; Ebashi and

Lipmann 1962). The P-type ATPase family (organized in five subfamilies called P1, P2, P3, P4, and P5, with the P2 subfamily including for instance the Na⁺, K⁺- and H⁺, K⁺-ATPases present in animal cells) is an evolutionarily conserved large family of proteins which is characterized by the formation of an obligatory aspartyl-phosphate intermediate during the pump cycle (Degani and Boyer 1973). P-type ATPases are chemi-osmotic pumps that convert the chemical energy of this aspartyl-phosphate intermediate into active transport of ions.

Probably because of its natural abundance in skeletal muscle, the SR Ca²⁺-ATPase was the first P-type ATPase for which the three-dimensional (3D) structure was solved (Toyoshima et al. 2000; Toyoshima and Nomura 2002). Since then, X-ray structures of analogs of the most relevant intermediate states in the Ca²⁺-ATPase transport cycle have been obtained (some of which are shown in Fig. 6.1), providing insights into the molecular events leading to Ca²⁺ translocation across the ER/SR membrane (e.g., Toyoshima et al. 2000, 2007; Toyoshima and Nomura 2002; Sorensen et al. 2004; Olesen et al. 2007). These high-resolution structures confirmed most of the previous suggestions based on site-directed mutagenesis (Andersen 1995), spectroscopic studies (Bigelow and Inesi 1992), or low-resolution structures (Dux and Martonosi 1983; Toyoshima et al. 1993; Zhang et al. 1998). In these structures, the cytosolic portion of the pump is made of three distinct domains: the nucleotide-binding (N) domain, the phosphorylation (P) domain, and the actuator (A) domain (Fig. 6.1). The N domain binds the nucleotide and positions the γ -phosphoryl of ATP for a nucleophilic attack on the conserved aspartate residue located in the P domain, while the A domain contains a strongly conserved TGES motif involved in the dephosphorylation of the aspartate. These cytosolic domains are connected to a transmembrane (M) domain consisting of 10 α -helices (for Serca1a), via a linker region which is crucial for transmitting events from the cytosolic catalytic portion of the pump to the cation-binding sites where transmembrane transport does actually occur.

The transport cycle of P-type ATPases from the P2 subfamily (cation transporters) has been worked out in detail. Briefly, and as depicted in Fig. 6.1, it starts with binding of the ion(s) to be transported (two Ca²⁺ ions in the case of Serca1a) to a conformation called “E1,” which has high-affinity binding sites for Ca²⁺ within the membrane domain, and accessible from the cytosolic side. Binding of the ion(s) allows the aspartic acid to become phosphorylated from Mg²⁺-ATP. The resulting high-energy “Ca₂E1~P” intermediate is then converted to an “E2P” intermediate, this step being accompanied by the release of the ion to the exoplasmic side. The transport site now has high affinity for a counter-transported ion species (two to three protons per ATP hydrolytic cycle in the case of Serca1a). Hydrolysis of the phosphorylated aspartic acid drives the enzyme back to the “E2” conformation.

Shortly after cloning the *SERCA1a* gene (Brandl et al. 1986), various host cell systems (for example COS cells or insect cells) have been developed for heterologous expression of Serca1a, in order to dissect by site-directed mutagenesis the molecular mechanism for Ca²⁺ transport (Maruyama and MacLennan 1988; Zhang et al. 2000; Miras et al. 2001). These systems have been very successful for

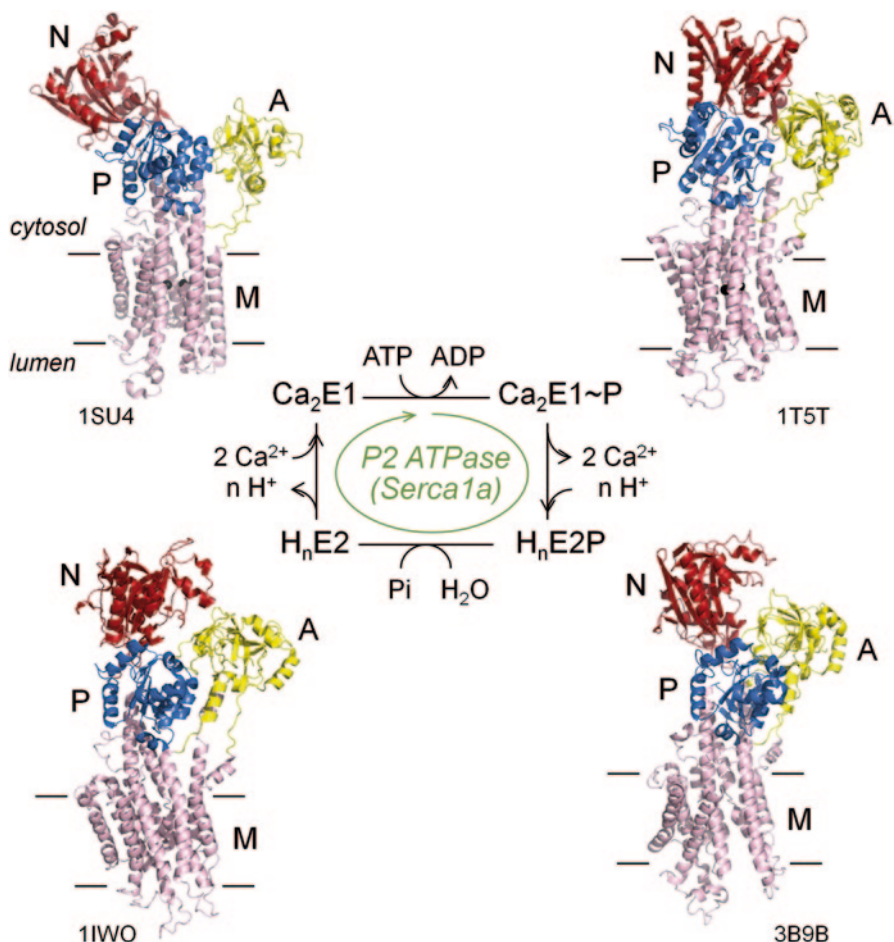


Fig. 6.1 Key intermediates in the catalytic cycle of Serca1a and 3D structures of their analogs. The cycle starts with the exchange of n protons with two Ca^{2+} ions in the cytosol. The Ca^{2+} -bound ATPase is then phosphorylated from ATP to give the $Ca_2E1\sim P$ state, which is converted to the $E2P$ state after exchange of two Ca^{2+} ions with n protons. Hydrolysis of the phosphorylated aspartic acid drives the enzyme back to the $E2$ conformation. The structures are shown as cartoons, with the N domain in red, the P domain in blue, the A domain in yellow, and the M domain in light pink. Ca^{2+} ions are shown as black spheres. The approximate boundaries of the membrane bilayer are indicated by solid horizontal lines. Protein Data Bank (PDB) accession codes: 1SU4 (Ca_2E1 form, obtained in the presence of Ca^{2+}); 1T5T ($Ca_2E1\sim P$ form; obtained in the presence of Ca^{2+} , Mg^{2+} , K^+ , ADP, and AlF_4^-); 3B9B ($E2P$ form, obtained in the presence of Mg^{2+} and BeF_3^-); 1IWO ($E2$ form, obtained in the presence of Thapsigargin, a specific inhibitor of SERCAs)

identification of residues or regions of the protein critical for transport and catalysis (Clarke et al. 1989; Andersen and Vilsen 1992; Strock et al. 1998). However, in spite of high expression levels, these cells are generally grown in too small quantities to allow purification at a reasonable cost of amounts compatible with crystallization trials.

6.3.2 Optimization of *Serca1a* Expression in *S. cerevisiae*

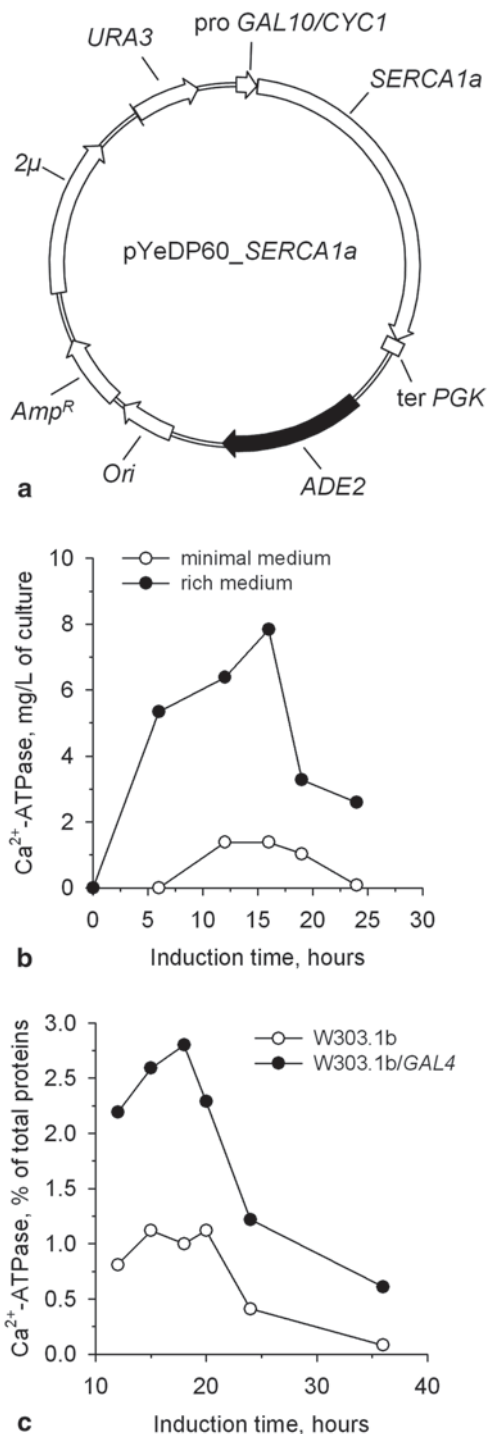
In this context, rabbit *Serca1a* heterologous expression in our laboratory began in the early 1990s, by using the yeast *S. cerevisiae* (Centeno et al. 1994). The initial strategy consisted in inserting the *SERCA1a* complementary DNA (cDNA) in a high-copy-number plasmid, pYeDP1/8–10 (Pompon 1988), under the control of an inducible promoter, with the rationale that delaying the expression phase until the end of the growth phase should allow to overcome potential toxicity problems associated with overexpression of a foreign protein. In pYeDP1/8–10, the strong *GAL10*–*CYC1* hybrid promoter is fully repressed by glucose present in the medium and when glucose is exhausted (corresponding here to the end of the exponential phase, as deduced from glucose titration assays), addition of galactose triggers induction of the target gene. The *GAL10*–*CYC1* hybrid promoter contains a sequence called the *GAL* upstream activating sequence (UAS_G) which is required for the induction of the *GAL10* gene in the presence of the product of the *GAL4* regulatory gene (Guarente et al. 1982; West et al. 1984). In addition, this hybrid promoter contains functional TATA elements of the *CYC1* gene that are DNA regions essential for transcriptional initiation by RNA polymerase II (Li and Sherman 1991), as well as the transcription initiation region of the *CYC1* gene itself (Guarente et al. 1982).

Using this expression system, the heterologously expressed Ca^{2+} -ATPase was recovered in the plasma membrane fraction (and to a lesser extent in the microsomes), yielding ~0.3% of total proteins in the plasma membrane (Centeno et al. 1994). Despite this rather low expression level, the total amount produced was quite substantial since about 1 mg Ca^{2+} -ATPase could be obtained per liter of culture. In addition, the expressed Ca^{2+} pump was functional, as judged by its ability to form a phosphoenzyme form [γ - ^{32}P]ATP and by the fact that its specific ATP hydrolysis activity was similar to that of native SR Ca^{2+} -ATPase (Centeno et al. 1994).

A significant step forward was then made by optimizing the expression plasmid as well as the host strain.

Optimization of the expression plasmid was achieved by using the pYeDP60 plasmid designed by Pompon and colleagues (Pompon et al. 1996), a plasmid identical to the pYeDP1/8–10 plasmid except for the presence of an additional selection marker, *ADE2* (Fig. 6.2a). This selection marker is of special interest since even rich culture media become rapidly deprived of adenine upon yeast growth at high cell densities. Thus, in contrast to yeast cells transformed with pYeDP1/8–10, yeast cells auxotroph for adenine and transformed with pYeDP60 could grow in rich medium (where the cell density can rise up to a level about fivefold to tenfold higher than in minimal medium) with minimal plasmid loss. Adding the *ADE2* selection marker turned out to be critical as, thanks to the increased maximal cell density, the expression level reached ~8 mg Ca^{2+} -ATPase per liter of rich culture medium (compared with ~1.5 mg Ca^{2+} -ATPase per liter of minimal medium, as shown in Fig. 6.2b; Lenoir et al. 2002). This corresponded to production of *Serca1a* up to about 0.7% of yeast total proteins (not shown), or even 1% of yeast total proteins after performing two successive galactose inductions (open circles in Fig. 6.2c).

Fig. 6.2 Optimization of *Serca1a* expression. **a** Map of the pYeDP60 plasmid used for the expression of *SERCA1a*. *ADE2* auxotrophy selection marker for adenine, represented as a *black arrow* for emphasis; *Ori* bacterial replication origin; *Amp^R* gene conferring resistance to ampicillin; *2μ* yeast replication origin; *URA3* auxotrophy selection marker for uracil; *pro GAL10/CYC1* strong hybrid galactose-inducible promoter; *ter PGK* sequence of the phosphoglycerate kinase used for termination of transcription. **b** Comparison of *Serca1a* expression from yeast growing in rich (●) or minimal (○) medium. The W303.1b yeast strain was grown at 28 °C in either rich or minimal medium. At time zero, expression was triggered by addition of 2% galactose. **c** Optimization of the host strain. An additional copy of the *GAL4* gene was inserted at the *Trp* locus of W303.1b yeast strain, resulting in the W303.1b/*GAL4* strain. W303.1b (○) and W303.1b/*GAL4* (●) yeast strains were transformed with pYeDP60 and grown in rich medium. Two galactose inductions (2% w/v each time) were performed here; a first one at time zero and a second one 13 h later. Panels **b** and **c** adapted from Lenoir et al. (2002)



Further improvement of Ca^{2+} -ATPase expression was achieved by manipulating the W303.1b host strain. A common limitation to the use of recombinant *GAL* promoters for the controlled high-level expression of foreign proteins is the inherent low expression level of the *GAL4* gene product, Gal4p, which mediates expression of proteins that are under the control of *GAL* promoters (Romanos et al. 1992). To overcome this limitation, we integrated a hybrid gene in the yeast chromosome, consisting of the *GAL10* promoter fused to the *GAL4* gene (Schultz et al. 1987; Pedersen et al. 1996). In this manner, the expression of Gal4p is increased in the presence of galactose, as well as that of the foreign protein. This is especially helpful for expression of proteins from multi-copy plasmids (2μ -based plasmids). The resulting yeast strain, W303.1b/*GAL4* was compared with the W303.1b parental strain for its ability to overexpress *Serca1a*. As illustrated by the black circles in Fig. 6.2c, the expression level of *Serca1a* reached a maximum of almost 3% of total yeast proteins, thus about ten times more than that reported by Centeno and colleagues (Centeno et al. 1994), and 30 mg Ca^{2+} -ATPase could be recovered for 1 L of culture (Lenoir et al. 2002).

Another critical parameter was the control of temperature during the expression phase. As already observed in bacteria and other yeast expression systems, decreasing the temperature from 28 to 18 °C significantly increased the amount of properly folded and active protein (Lenoir et al. 2002). One obvious reason for this is that the yeast metabolism is slowed down, thereby facilitating integration by the translocon of newly synthesized proteins into the ER membrane. When after yeast breakage differential centrifugation was performed to roughly separate a “heavy” membrane fraction (sedimenting at rather low-speed centrifugation, ~20,000 g), from a “light” membrane fraction (recovered after high-speed centrifugation, ~100,000 g), we observed that the temperature influenced the final destination of the expressed *Serca1a* toward the “light” membrane fraction. The “heavy” fraction is poorly solubilized by mild detergents, with the solubilized fraction containing low amounts of active *Serca1a*, while the “light” membrane fraction is more easily solubilized and contains higher amounts of active protein. In sum, although shifting temperature to 18 °C during the expression phase decreases *Serca1a* overexpression, the quality of the protein expressed at 18 °C is enhanced (Lenoir et al. 2002). In the “light” fraction, recombinant *Serca1a* finally accounts for ~1.5–2% of total MPs (~6–8 mg *Serca1a* per liter of culture).

6.3.3 Affinity Purification of *Serca1a*

Initial *SERCA1a* constructs were cloned in frame with a C-terminal hexahistidine tag, for subsequent affinity purification by immobilized metal ion affinity chromatography. The yield of this purification step was about 20%, allowing the recovery of about 1 mg Ca^{2+} -ATPase from 1 L of yeast culture, and with a purity estimated to be about 50%. Although this proved to be sufficient for functional studies of wild-type and mutated Ca^{2+} -ATPases (Lenoir et al. 2004, 2006), the relatively low purity

of the sample was not compatible with crystallization trials. Toward this goal, we resorted to the extraordinarily high affinity of avidin for biotin.

Biotin (vitamin H) is a small coenzyme synthesized by plants, most bacteria, and some fungi, which is bound to a specific lysine at the active site of biotinylated proteins *via* an amide linkage. The biotinylation reaction is performed by biotin protein ligases (Chapman-Smith and Cronan 1999). The biotinylation reaction catalyzed by biotin protein ligases is very specific, and there are only a scarce number of biotinylated proteins (only one in *E. coli* and up to five in most other organisms). Biotinylated proteins are multi-subunit enzymes that play roles in metabolic carboxylation/decarboxylation and transcarboxylation reactions, such as the oxaloacetate decarboxylation reaction. The biotinylation domain is strongly conserved among biotinylated proteins (Schwarz et al. 1988). This region is usually located at the C-terminus of the polypeptide chain and the lysine residue that becomes modified with biotin is generally located 34 or 35 residues upstream from the C-terminus (Chapman-Smith and Cronan 1999).

Using various biotinylated proteins, it has been shown that the minimum size of the segment required for biotinylation is 75 amino acids long, starting from the protein C-terminus. This sequence, when added to the C-terminus of β -galactosidase, is indeed sufficient to promote biotinylation of the fusion protein expressed in *E. coli* (Cronan 1990). Importantly, that study also showed that a bacterial biotin-acceptor sequence also functions in yeast, paving the way for expression of proteins fused to biotin-acceptor domain (BAD) in eukaryotic expression systems (Cronan 1990).

The specific and high affinity between tetrameric streptavidin and biotin ($K_d \sim 10^{-14}$ mol/L) can be used to purify biotinylated proteins. Such strategy has been successfully used for the purification in a single step of various proteins, including MPs such as lactose permease expressed in *E. coli* (Consler et al. 1993; Pouny et al. 1998), a plant sucrose carrier (Stolz et al. 1995) and human P-glycoprotein (Howard and Roepe 2003), both expressed in *S. cerevisiae*. However, because of the extremely tight binding between tetrameric streptavidin and biotin, harsh conditions are required to disrupt their interaction, often leading to denaturation of the protein of interest.

We therefore used an alternative procedure for purification purposes. We fused the last 93 residues of the BAD of the α -subunit of oxaloacetate decarboxylase, a biotinylated enzyme from *Klebsiella pneumoniae*, to the C-terminus of Serca1a, but to circumvent problems associated with the subsequent release of the fusion protein from avidin beads (even using monomeric avidin, which has a reduced affinity for biotin), a thrombin cleavage site was inserted between Serca1a and BAD (Fig. 6.3a; Jidenko et al. 2006). The recombinant Serca1a-BAD was then expressed using the above-described pYedP60 plasmid, in the *S. cerevisiae* strain W303.1b/*GAL4*. Serca1a-BAD proved to experience biotinylation *in vivo*, and after solubilization of yeast membranes with *n*-dodecyl β -D-maltoside (DDM), the biotinylated protein was retained specifically onto monomeric avidin-coupled resin (Fig. 6.3b, lane “R₀”; Jidenko et al. 2006). Note that naturally occurring yeast biotinylated proteins were also retained onto avidin beads, but the treatment of the resin with thrombin (lane “R₆₀”) allowed to specifically elute Serca1a (lane “E”), without its BAD

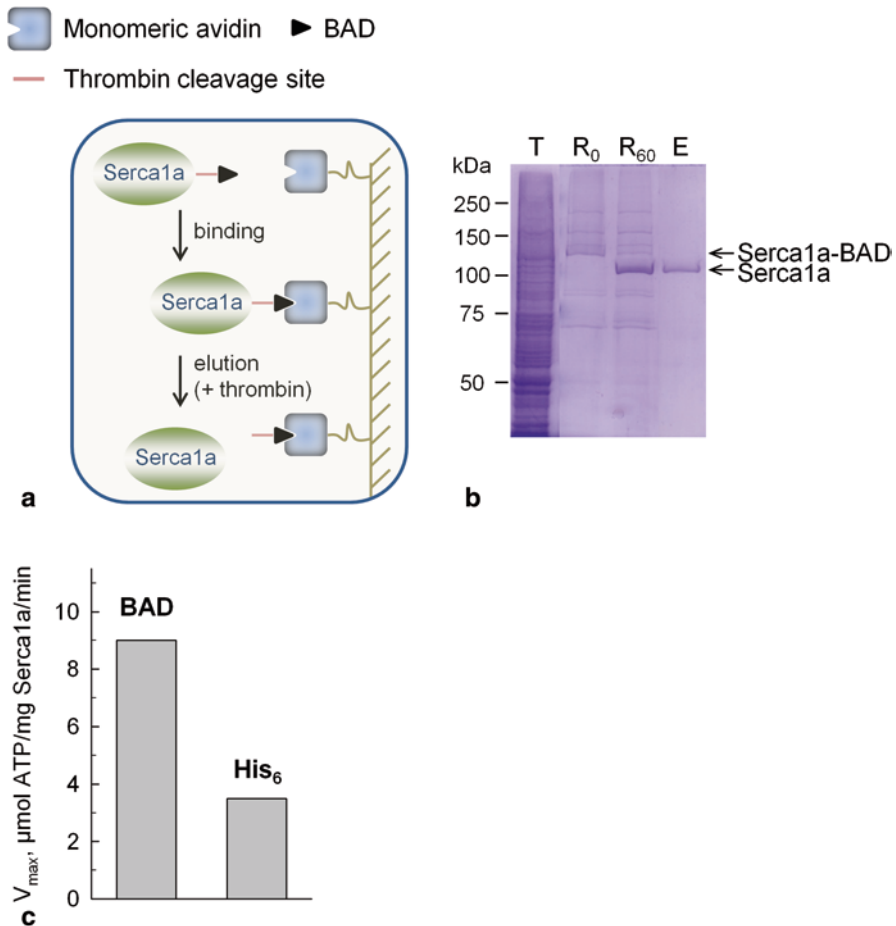


Fig. 6.3 Affinity purification of *Serca1a* in a functional form. **a** Purification scheme. The C-terminally *BAD*-tagged *Serca1a* (*Serca1a*-*BAD*) is solubilized from yeast membranes with DDM and applied onto a *monomeric avidin* resin. A *thrombin cleavage site* located between *Serca1a* and the *BAD* tag allows elution of *Serca1a* from the resin. **b** SDS-PAGE analysis of *Serca1a* purification. *T* total—starting yeast membrane fraction, expressing *Serca1a*-*BAD*; *R*₀ resin—aliquot of the *monomeric avidin* resin before treatment with thrombin, showing proteins bound to the resin; *R*₆₀ aliquot of the *avidin* resin after a 60-min incubation with thrombin; *E* eluate fraction recovered after thrombin cleavage. Proteins were run onto an 8% Laemmli-type acrylamide gel and stained with Coomassie blue. **c** Maximum velocity of Ca^{2+} -dependent ATP hydrolysis by *Serca1a* purified either from an initial *BAD*-tagged construct or from a *His*₆-tagged construct. ATPase activity was assayed spectrophotometrically at 30 °C. (Adapted from Jidenko et al. 2006)

moiety. At this stage, *Serca1a* was in a relatively pure form (Fig. 6.3b) and the yield of purification was estimated to be about 6%. As much as 0.3 mg *Serca1a* was recovered from 1 L yeast culture. Beyond the fact that purification on *monomeric avidin* yielded a more homogeneous sample than affinity purification on Ni^{2+} -NTA

beads, the purified (initially BAD-tagged) enzyme displayed a maximal rate of ATP hydrolysis of about 9 $\mu\text{mol/mg}$ Serca1a/min at 30°C, about twofold to threefold higher than that of the His-tagged enzyme (Fig. 6.3c), suggesting either that the six histidines located at the C-terminus of Serca1a are deleterious for enzyme activity (as the His₆ tag was not removed after purification) or that active ATPases are somehow preferentially selected by the avidin purification procedure (Jidenko et al. 2006). The latter hypothesis is consistent with the idea that only correctly folded biotinylation domains will be biotinylated, ensuring that misfolded biotinylation domains will not be purified, whereas using histidine tags will not make it possible to sort correctly folded from misfolded fusion proteins.

6.3.4 Crystallization and Structure Determination of Heterologously Expressed Wild-Type and Mutated Serca1a

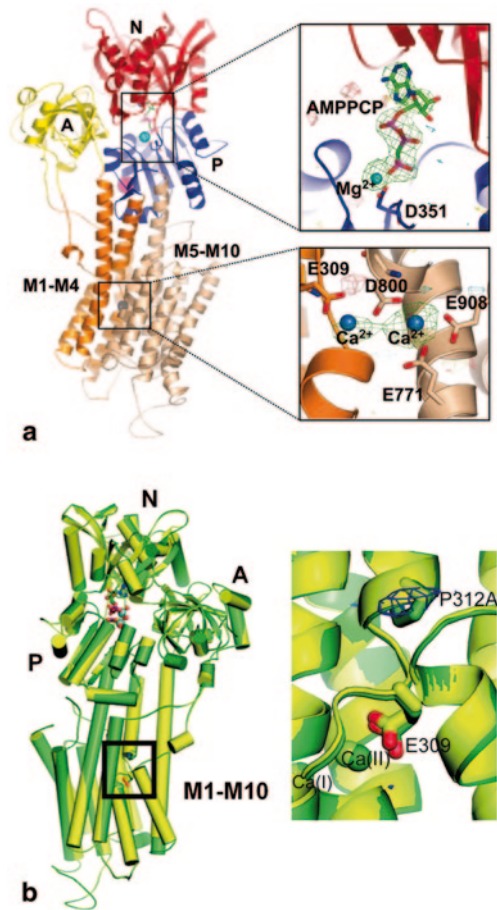
This system has been the cornerstone for purification of Serca1a, after its heterologous expression in yeast. However, a major improvement for subsequent structural characterization came from the introduction of a second purification step, namely size-exclusion chromatography (SEC-HPLC). Such extra purification step made it possible to eliminate the small amounts of aggregates and contaminants (including the thrombin protease), to increase the purity of the sample, and, most importantly, to simultaneously allow for detergent exchange (Jidenko et al. 2005). Indeed, the type of detergent chosen for crystallization attempts is a critical parameter, because it will influence both the stability of the protein and its capability to form crystals (Sorensen et al. 2006). As octaethylene glycol mono-*n*-dodecyl ether (C₁₂E₈) had previously proven superior to DDM for crystallogenesis of several P-type ATPases (Toyoshima et al. 2000; Sorensen et al. 2004), DDM was exchanged for C₁₂E₈ during the above-mentioned SEC-HPLC step (Jidenko et al. 2005).

After concentration of the purified sample to ~10 mg/mL and relipidation with 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC), crystallization conditions were explored using a screen based on conditions identified with native Serca1a (Sorensen et al. 2004). Crystals obtained in the presence of AMPPCP, a nonhydrolyzable ATP analog, and Ca²⁺, diffracted beyond 3.1 Å, and the diffraction properties were similar to those of native Serca1a, without any significant structural difference between the recombinant Serca1a (Fig. 6.4a) and its native counterpart crystallized under the same conditions. Binding of AMPPCP takes place at a highly conserved region located at the interface between N and P domains (inset of Fig. 6.4a) and AMPPCP can be found in very close proximity to the invariant phosphorylated aspartate found in all P-type ATPases (D351 in the case of Serca1a). The two Ca²⁺ ions that are transported at the expense of one ATP molecule per ATPase cycle are coordinated by residues located in the membrane domain (M; inset of Fig. 6.4a). This structure, together with that of a mammalian voltage-dependent K⁺ channel, represent the first two crystal structures for mammalian integral transmembrane proteins produced in a heterologous system (Jidenko et al. 2005; Long et al. 2005; Midgett and Madden 2007).

Fig. 6.4 X-ray 3D structures of recombinant Serca1a expressed in yeast.

a Structure of recombinant rabbit wild-type Serca1a. The structure, obtained in the presence of Ca^{2+} and AMP-PCP, is shown as cartoon with the *N* domain displayed in red, the *P* domain in blue, the *A* domain in yellow, the first four TM helices in orange, and the last six TM helices in light pink. Two critical regions involved in nucleotide binding and autophosphorylation, and in Ca^{2+} binding, are magnified. Reproduced from Jidenko et al. (2005).

b Structure of recombinant rabbit P312A mutant of Serca1a. The structure was obtained in the presence of ADP and the phosphate analog AlF_4^- . On the left, the crystal structure of P312A (yellow) is superimposed with the wild-type structure in the same conditions (green). On the right, magnification of the Ca^{2+} -binding site in the transmembrane region underlining the fact that no significant differences are observed between wild-type and P312A structures. P312 or A312, and E309 are indicated. (Reproduced from Marchand et al. 2008)



Obtaining the structure of the heterologously expressed wild-type Serca1a has paved the way toward analysis of the structure of mutated proteins. Indeed, many mutations have been reported to interfere with the Ca^{2+} -ATPase catalytic cycle (Andersen 1995), but whether the effect of these mutations reflects a genuine alteration of partial reactions of the catalytic cycle or whether it is the result of a global structural change remained to be determined. Along those lines, the P312A mutation in Serca1a, which had been previously reported as interfering with ATP hydrolysis and Ca^{2+} transport (Vilsen et al. 1989), has been selected. P312 is located in transmembrane segment M4, next to the $^{308}\text{PEGL}$ motif containing the E309 residue involved in Ca^{2+} binding. Studying the partial reactions of P312A catalytic cycle

revealed that the rate of the $\text{Ca}_2\text{E1} \sim \text{P} \rightarrow \text{E2P}$ transition was dramatically decreased (Vilsen et al. 1989). The P312A mutant was expressed and purified using our procedure, and subsequently crystallized (Marchand et al. 2008). Crystallization was performed in the presence of AlF_4^- and adenosine diphosphate (ADP) together with 10 mM Ca^{2+} in order to mimic the ADP-sensitive phosphoenzyme $\text{Ca}_2\text{E1} \sim \text{P}$. The crystal structure of the P312A mutant displayed a high degree of isomorphism with that of the wild-type enzyme (Fig. 6.4b), and it was thus concluded that the severe functional consequences of this mutation were associated with only minor structural changes (Marchand et al. 2008). It was proposed in that work that replacing P312 by an alanine does stabilize the $\text{Ca}_2\text{E1} \sim \text{P}$ conformation by facilitating hydrogen bonding between P308 and A312. This mutation would thus relieve an inbuilt constrained region at site I (Marchand et al. 2008).

These achievements with expressed Sercala were the basis for attempting to overexpress in yeast (and purify) other MPs, as exemplified below by the lipid “flippase” Drs2p/Cdc50p complex.

6.4 Coordinated Overexpression of the Lipid “Flippase” Complex Drs2p/Cdc50p

6.4.1 *Transbilayer Phospholipid Asymmetry in Eukaryotic Cell Membranes*

Studying how membrane transport proteins termed “flippases” create and maintain phospholipid asymmetry in eukaryotic cell membranes is rapidly expanding, because of the influence of membrane lipid asymmetry on a multitude of cellular functions. Prime candidates for this transport activity are P-type ATPases from the P4 subfamily (hereafter referred to as P4-ATPases), but the mechanism for phospholipid transport by P4-ATPases remains largely elusive (Poulsen et al. 2008b; Coleman et al. 2012).

Let us first recall that eukaryotic cells contain thousands of different lipid structures. This diversity reflects numerous different functions for cellular lipids: in addition to their role in cell compartmentalization, they are used for energy storage, as important regulatory molecules, or they may act as first or second messengers in signal transduction (van Meer et al. 2008). In membranes of the late secretory pathway, this range of functions is accompanied by a particular distribution of lipids: The *trans*-Golgi network (TGN), as well as the plasma and endosomal membranes, display an asymmetric transbilayer distribution with PS and phosphatidylethanolamine (PE) primarily restricted to the cytosolic leaflet, and sphingomyelin (SM) and glycosphingolipids (GSLs) mainly residing in the noncytosolic leaflet (Op den Kamp 1979; Daleke 2007). Such an asymmetric distribution has important functional consequences. For instance, the strong interaction between GSLs and sterols ensures a high stability and impermeability of the plasma membrane; conversely,

the enrichment of aminophospholipids in the cytosolic leaflet of the plasma membrane and on the surface of endocytic and exocytic vesicles may help to keep these membranes in a fusion-competent state (Kinnunen and Holopainen 2000).

6.4.2 P4-ATPases and Cdc50 Proteins as Prime Candidates for Phospholipid Translocation

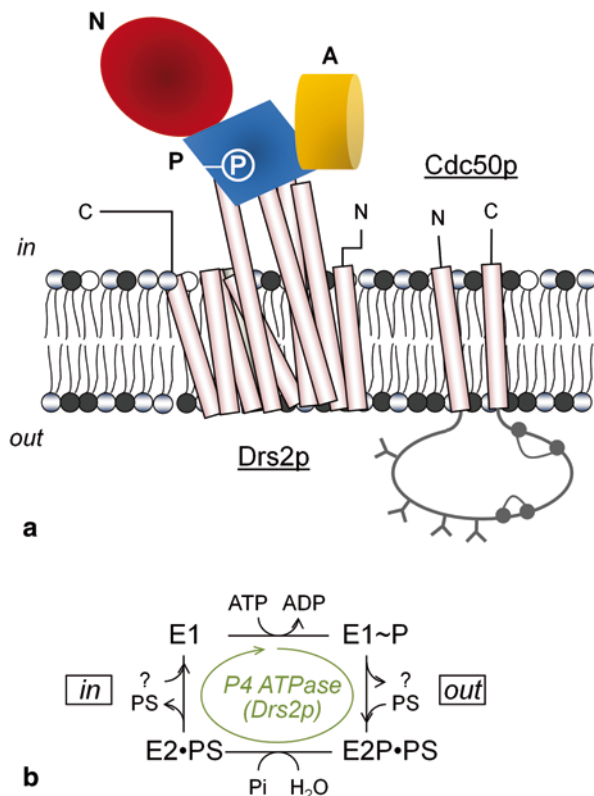
P4-ATPases are thought to play a prominent role in creating and maintaining this phospholipid asymmetry, by selectively translocating lipids (instead of translocating ions, like members of most P-type ATPase subfamilies), mainly PS and PE, from the luminal leaflet to the cytosolic leaflet of plasma membranes as well as of TGN membranes (Graham 2004; Daleke 2007). This lipid-translocation activity is thought to be dependent on the intimate coupling of P4-ATPases with accessory subunits known as Cdc50 proteins (Poulsen et al. 2008a; Lenoir et al. 2009; Lopez-Marques et al. 2011).

In plasma membranes, loss of lipid asymmetry and the resulting cell-surface exposure of PS triggers recognition of apoptotic cells by macrophages, activation of the blood coagulation cascade, or virus entry (Rosing et al. 1980; Fadok et al. 2000; Mercer and Helenius 2008). But much remains to be discovered about other physiological roles of lipid asymmetry and the relevance for eukaryotic cells to spend energy to maintain such transbilayer asymmetry in other contexts. This is especially true for unicellular organisms like yeast cells, as they certainly do not need to deal with PS exposure for triggering blood clotting signaling. So, what is lipid translocation also made for? An appealing hypothesis has been provided by the observation that the yeast P4-ATPase Drs2p is required for budding of post-Golgi exocytic vesicles (Gall et al. 2002) and for the formation of clathrin-coated vesicles (Chen et al. 1999). The hypothesis is that by expanding one of the leaflets of the bilayer while reducing phospholipid number in the other one, flippases might drive membrane bending and thereby provide a major contribution to vesicle formation and ultimately vesicle-mediated protein transport (Devaux et al. 2008).

Of note, P4-ATPases are associated with several inherited disorders. For instance, one human disease, progressive familial intrahepatic cholestasis (PFIC1, or Byler disease), has been directly correlated with mutations in a gene encoding a P4-ATPase (*ATP8B1*). PFIC1 is an autosomal recessive disorder for which individuals manifest cholestasis in infancy and which progresses toward end-stage liver disease before adulthood (Bull et al. 1998).

Collectively, these data highlight the crucial importance of P4-ATPases in health and disease and underscore the need for structural and mechanistic insights into the molecular mechanism by which P4-ATPases catalyze lipid transport. We therefore aimed at adapting our overexpression system to P4-ATPases. A specific impetus was that although the wealth of P-type ATPase 3D structures has tremendously expanded in the recent years (Morth et al. 2007; Pedersen et al. 2007; Shinoda et al. 2009; Gourdon et al. 2011), P4-ATPases still lag behind.

Fig. 6.5 Predicted topology of yeast Drs2p and Cdc50p and putative lipid transport cycle. **a** Predicted topology of Drs2p and Cdc50p. Transmembrane helices of Drs2p and Cdc50p are shown as *light pink cylinders*. The phosphorylation site is indicated in the *P* domain. The phosphorylation domain (*P*) is represented in *blue*, the nucleotide-binding domain (*N*) is represented in *red*, and the actuator (*A*) domain is shown in *yellow*. N- and C-termini are indicated on the cytosolic (“in”) side of the bilayer. Cysteines of Cdc50p involved in the formation of disulfide bridges and previously identified by site-directed mutagenesis (Puts et al. 2012) are symbolized by *gray dots*. Predicted sites for *N*-glycosylation are also represented on Cdc50p ectodomain. **b** Simplified scheme of the transport cycle for P4-ATPases. Whether phospholipid transport by P4-ATPases (from the outside to the inside) is coupled to transport of another substrate in the opposite direction is currently unknown



There are five P4-ATPase members in yeast: Dnf1p and Dnf2p which reside at the plasma membrane, Drs2p and Dnf3p which reside in the TGN, and Neo1p localized in the endosomes (Gall et al. 2002; Hua et al. 2002). Deletion of Dnf1p and Dnf2p inhibits ATP-dependent transport of fluorescent analogs (nitrobenzoxadiazole (NBD)-labeled) of PC, PS, and PE (Pomorski et al. 2003), while removal of Drs2p and Dnf3p abolishes NBD-PS/PE and NBD-PC/PE transport, respectively (Natarajan et al. 2004; Alder-Baerens et al. 2006). Concerning accessory proteins, genetic disruption of yeast Cdc50 proteins was found to phenocopy *dnf* and *drs2* deletions. This family includes three proteins in yeast, namely Cdc50p, Lem3p, and Crf1p. It is now clear that Cdc50 proteins associate with P4-ATPases and that this association is of primary importance. Indeed, Cdc50 proteins have been shown to be required for stability and export of P4-ATPases from the ER, both in yeast and in mammalian cells (Saito et al. 2004; Chen et al. 2006; Furuta et al. 2007; Paulusma et al. 2008;

van der Velden et al. 2010). Recent evidence also indicates that Cdc50 proteins play an intimate role in the transport cycle catalyzed by P4-ATPases (Poulsen et al. 2008a; Lenoir et al. 2009; Bryde et al. 2010).

Because yeast Drs2p is the one for which the most convincing data suggesting implication in lipid transport have been obtained, we decided to focus on this particular P4-ATPase. Drs2p is predicted to contain ten transmembrane spans and to have an overall domain organization similar (but with long N- and C-terminal extensions) to that of other P-type ATPases from the P2 and P3 subfamilies (Fig. 6.5a). It binds specifically to the Cdc50 protein called Cdc50p. The Cdc50p polypeptide chain is predicted to span the membrane twice, with a large ectodomain protruding toward the TGN lumen (Fig. 6.5a). Two disulfide bridges have been identified in this ectodomain (Puts et al. 2012), the latter also containing four consensus sequences for N-linked glycosylation (Fig. 6.5a). As previously discussed, all P-type ATPases, including P4-ATPases, display a clear conservation of the key residues known in P2-ATPases to be involved in ATP binding, transient phosphorylation (the phosphorylated residue is Asp560 in Drs2p), and phosphoenzyme hydrolysis, as well as a common organization of transmembrane helices (Lenoir et al. 2007). Assuming that P2 and P4-ATPases also share similar mechanisms of energy transduction and reaction schemes, PS binding to Drs2p after phosphorylation from ATP, by analogy with proton binding to Serca1a or K⁺ binding to Na⁺, K⁺-ATPase, should stimulate dephosphorylation of the pump, in parallel to its transport to the other leaflet of the bilayer (Fig. 6.5b).

6.4.3 *Functional Co(over)Expression of Drs2p and Cdc50p in Yeast Membranes*

As most P4-ATPases appear to function as protein complexes, we devised a high-yield co-expression system for the yeast P4-ATPase Drs2p and its accessory subunit Cdc50p.

To facilitate detection of Drs2p and Cdc50p as well as for future purification of the complex, a BAD and a decahistidine tag were added at the C-terminus of Drs2p and Cdc50p, respectively. The fused genes were cloned independently in pYeDP60 expression plasmid, resulting in pYeDP60_*DRS2*-BAD and pYeDP60_*CDC50*-His₁₀ (Fig. 6.6). As described for Serca1a in the previous section of this chapter, *DRS2* and *CDC50* genes in this plasmid were both placed under the control of a strong galactose-inducible promoter. From the pYeDP60_*CDC50*-His₁₀ vector, a cassette containing the promoter, the *CDC50* coding sequence, and the PGK terminator was then amplified by polymerase chain reaction (PCR) and subsequently inserted in the pYeDP60_*DRS2*-BAD companion plasmid. We thus obtained pYeDP60_*DRS2*-BAD/*CDC50*-His₁₀, appropriate for co-expression of Drs2p-BAD and Cdc50p-His₁₀ (Fig. 6.6). The rationale for constructing this co-expression plasmid, rather than expressing the two proteins from different plasmids, was to avoid imbalanced transcription of *DRS2* and *CDC50* genes, because of an unequal number of plasmids in each cell, a frequent behavior of 2 μ -based plasmids.

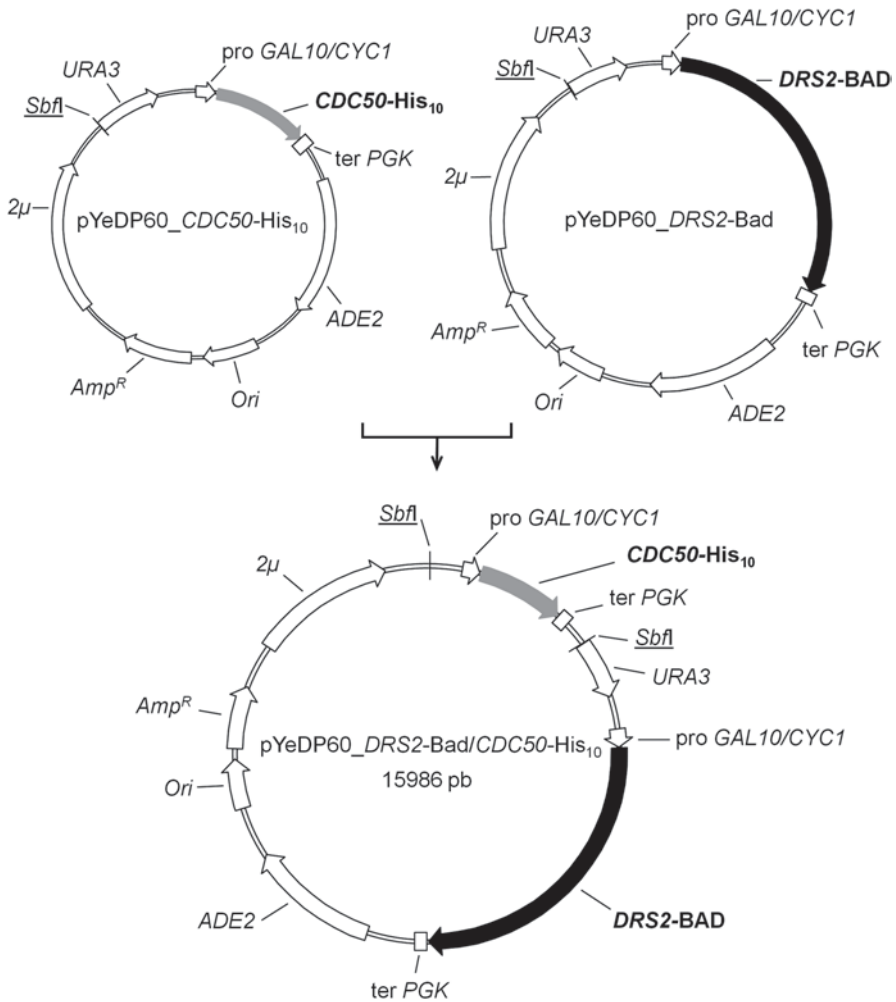


Fig. 6.6 Maps of the plasmids used for co-expression of Drs2p-BAD and Cdc50p-His₁₀: *CDC50* fused to a C-terminal decahistidine tag, His₁₀ (top left, light gray arrow), and *DRS2* fused to a C-terminal biotin-acceptor domain, BAD (top right, black arrow), were first cloned independently in pYeDP60, in both cases under the control of the inducible *GAL10/CYC1* hybrid promoter, using *EcoRI* and *SmaI* restriction sites. This resulted in pYeDP60_*DRS2-BAD* and pYeDP60_*CDC50-His₁₀* plasmids (top right and left, respectively). A cassette containing the *GAL10/CYC1* promoter, the *CDC50* gene, and the *PGK* terminator was then amplified by PCR from pYeDP60_*CDC50-His₁₀* using primers containing *SbfI* restriction sites, and inserted in pYeDP60_*DRS2-BAD*, which has a unique *SbfI* restriction site at position 11 (underlined). *ADE2* auxotrophy selection marker for adenine; *Ori* bacterial replication origin; *Amp^R* gene conferring resistance to ampicillin; *2μ* yeast replication origin; *URA3* auxotrophy selection marker for uracil

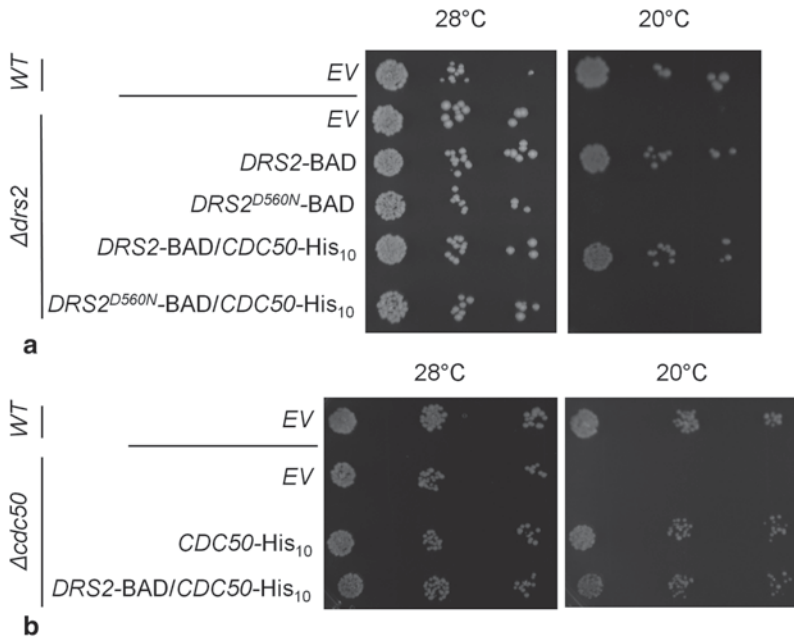


Fig. 6.7 Overexpressed Drs2p and Cdc50p restore the ability of $\Delta drs2$ and $\Delta cdc50$ yeast cells for growth at low temperatures. Yeast cells, either wild-type W303.1b/*GAL4-2* cells or $\Delta drs2$ or $\Delta cdc50$ mutants, were transformed with various vectors. Wild-type cells transformed with an empty vector (EV) were used as negative control. The $\Delta drs2$ mutant was transformed in order to express *DRS2* alone or in combination with *CDC50* (a), and the $\Delta cdc50$ mutant was transformed in order to express *CDC50* alone or in combination with *DRS2* (b). *DRS2* (wild type or mutated at its catalytic aspartate, *DRS2*^{D560N}) was tagged with a sequence coding for BAD at its C-terminus, and *CDC50* was tagged with a decahistidine tag (His₁₀) at its C-terminus. Serial dilutions of yeast cells were spotted onto a medium containing 2% galactose (and 2% fructose) for growth at either 28°C (left) or at the restrictive temperature of 20°C (right), for 3–5 days.

We checked the functionality of our constructs by taking advantage of the fact that $\Delta drs2$ yeast cells exhibit a cold-sensitive growth phenotype (Chen et al. 1999). As can be seen from Fig. 6.7a, BAD-tagged Drs2p restored the ability of $\Delta drs2$ cells for growth at 20°C in a galactose-containing medium, whereas a “dead” mutant of Drs2p, for which the catalytic Asp560 residue had been swapped for an Asn, did not, thus confirming that catalytic activity of Drs2p is required for growth at low temperature. Drs2p-BAD, expressed together with Cdc50p-His₁₀ thanks to our co-expression plasmid also restored growth of $\Delta drs2$ at 20°C. Similarly, Fig. 6.7b shows that constructs for Cdc50p, tagged with ten histidines at its C-terminus, restored growth of $\Delta cdc50$ cells at 20°C. The same was true if Cdc50p-His₁₀ was expressed from the co-expression plasmid (Fig. 6.7b).

We then turned to large-scale production of Drs2p and Cdc50p. Yeasts transformed with the pYeDP60_*DRS2*-BAD/*CDC50*-His₁₀ plasmid were first precultured in minimal medium, at 28°C (Fig. 6.8a). Yeast cells were then grown in a

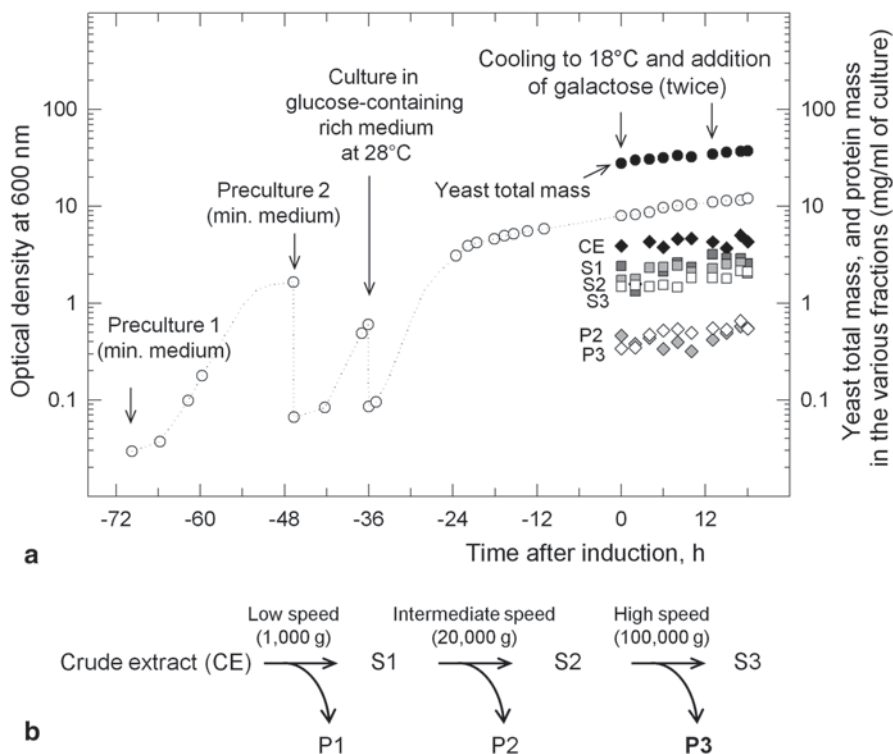


Fig. 6.8 Follow-up of yeast growth and membrane fractionation. **a** Yeast growth and expression phase, as monitored by turbidity measurements (*open circles*). Two precultures in minimal medium were performed (between -70 and -36 h) before yeast growth in rich medium (between -36 h and time zero) and the subsequent induction of expression (between time zero and 18 h). After various times of induction, the protein content of crude extract (CE) fractions recovered after yeast breakage with beads is indicated by *black diamonds*. **b** Schematic outline of the membrane fractionation procedure. In panel **a**, the total protein content of the *S1* low-speed supernatant (*dark gray squares*) as well as that of the *S2* (*light gray squares*) and *S3* (*white squares*) supernatants, and of the *P2* (*light gray diamonds*) and *P3* (*white diamonds*) pellets, are also indicated.

(glucose-containing) rich medium for 36 h, until a slower phase was reached, revealing glucose depletion from the medium (Fig. 6.8a). The biomass reached an absorbance at 600 nm of about 10 (using a Novaspec II Pharmacia spectrophotometer) at that stage, corresponding to roughly 30–40 g yeast cells per liter of culture. Expression of Drs2p-BAD and Cdc50p-His₁₀ was then induced by adding galactose (time zero in Fig. 6.8a) and simultaneously lowering temperature to 18°C, to improve protein folding and sorting. After 13 h, a second galactose addition was performed, for 5 more hours. Yeast aliquots were sampled during the expression phase, and after yeast breakage, various fractions were recovered by differential centrifugations, among which membrane pellets recovered at moderate and high

speed, P2 and P3, respectively (Fig. 6.8b). This protocol made it possible to collect about 0.5 g P3 membranes per liter of culture at the end (white diamonds in Fig. 6.8a). The various fractions were analyzed for their contents in Drs2p-BAD and Cdc50p-His₁₀, using a biotin probe and a His probe, respectively. Drs2p-BAD expression (Fig. 6.9a) increased progressively after induction by galactose, to reach a quasi-plateau after 17–18 h in both P3 and P2 membranes. Endogenous yeast biotinylated proteins (among which Pyc1/2p and Arc1p) were detected as faint bands. Cdc50p-His₁₀ (Fig. 6.9b) was detected as several bands corresponding to various glycosylation levels, the fastest one corresponding to core unglycosylated Cdc50p-His₁₀ (Jacquot et al. 2012) and the others to Cdc50p-His₁₀ glycosylated to various degrees.

Remarkably, the pattern of Cdc50p-His₁₀ glycosylation exhibited significant differences in P2 membranes *versus* P3 membranes (Fig. 6.9b and c). As a function of time, fair glycosylation of Cdc50p was observed in P2 membranes after a few hours of induction, but the ratio between the mature (glycosylated) and the non-mature forms of Cdc50p was clearly in favor of the non-mature one at the end of the expression period, whereas Cdc50p in P3 membranes remained properly glycosylated. This suggested that Cdc50p in the P3 fraction has reached compartments where more complete maturation has occurred. We tested whether expression of Drs2p and Cdc50p in P3 or P2 membranes would result in different functional properties, taking advantage of the fact that P-type ATPases generally form a stable phosphorylated intermediate during their catalytic cycle. P3 and P2 membranes were thus subjected to phosphorylation from [γ -³²P]ATP and the amount of phosphoenzyme formed at steady state was measured, either in the presence or in the absence of vanadate, a potent inhibitor of P-type ATPases. As displayed in Fig. 6.9d, vanadate-sensitive phosphorylation was threefold to fourfold higher in P3 membranes than in P2 membranes, suggesting that the most active Drs2p was recovered in P3 membranes (Jacquot et al. 2012).

To estimate the concentration of Drs2p-BAD in P3 membranes, we made use of two additional samples, namely P3 membranes from yeast expressing Serca1a-BAD (~119 kDa; Cardi et al. 2010a) and P3 membranes from nontransformed yeast, supplemented with 1.5% (w/w) of Serca1a-containing SR fragments (Fig. 6.10, “SR”) where Serca1a is known to be the predominant protein. Firstly, comparison of both samples using the Ab79 antibody directed against Serca1a indicated that yeast membranes are enriched to about 1.5–2% in Serca1a-BAD (Fig. 6.10, top blot). Secondly, comparison of the same Serca1a-BAD samples with Drs2p-BAD/Cdc50p-His₁₀-containing P3 membranes using a biotin probe indicated that for the same amount of total proteins present in P3 membranes, the amount of Drs2p-BAD is about twice that of Serca1a-BAD (Fig. 6.10, bottom blot). Thus, assuming that the BAD domains in Drs2p-BAD and Serca1a-BAD react similarly toward the biotin probe, Drs2p-BAD is expected to be enriched to about 3% (w/w) relative to the amount of total proteins in P3 membranes. Hence, out of the 0.5 g of total proteins recovered in the P3 fraction for 1 liter of culture, 15 mg correspond to Drs2p.

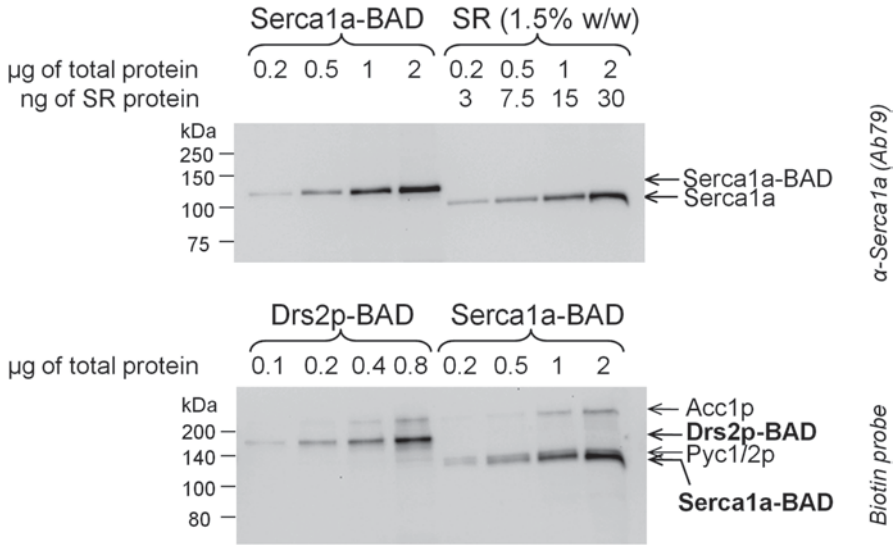


Fig. 6.10 Estimation of the concentration of Drs2p-BAD in P3 membranes of yeast overexpressing Drs2p-BAD/Cdc50p-His₁₀. The top gel compares P3 membranes derived from yeast cells expressing Serca1a-BAD with P3 membranes derived from nontransformed yeast cells, to which native sarcoplasmic reticulum (SR) membranes containing Serca1a were added (these SR membranes were mixed with the P3 membranes at a 1.5% w/w ratio). The Serca1a content in native SR membranes is known to be about 4–6 nmol ATPase/mg total protein. Immunoblotting of the top gel was performed using a α -Serca1a antibody (“Ab79”). The bottom gel shows comparison of the same P3 membranes derived from yeast expressing Serca1a-BAD but now together with P3 membranes containing Drs2p-BAD/Cdc50p-His₁₀. Immunoblotting of the bottom gel was performed using a biotin probe

6.4.4 Toward Purification of the Drs2p/Cdc50p Complex: Solubilization and Stability in Detergents

In view of future purification of the Drs2p/Cdc50p complex, we investigated the ability of various detergents to solubilize Drs2p-BAD and Cdc50p-His₁₀. DDM, which among nonionic mild detergents proved appropriate for solubilizing Serca1a

c Evolution of Cdc50p glycosylation over induction time in both P2 and P3 membranes. Glycosylated and core Cdc50p were quantified using the “Quantity One” software, and the glycosylation index shown is the ratio between the fully glycosylated and the nonglycosylated forms. **d** Phosphorylation from [γ -³²P]ATP of P3 and P2 membrane fractions co-expressing Drs2p-BAD and Cdc50p-His₁₀. Formation of the phosphoenzyme intermediate was measured after incubation of P3 or P2 membranes (at 0.5 mg total protein per mL) with 0.5 μ M [γ -³²P]ATP (0.25–1 mCi/ μ mol) for 25 s on ice, followed by acid quenching and filtration (see Jacquot et al. (2012) for detailed experimental procedures). Phosphorylation took place in the absence or presence of 1 mM vanadate (*open bars* and *gray bars*, respectively). Data are presented as the mean \pm SD of duplicates. Panel **d** has been reproduced from Jacquot et al. (2012)

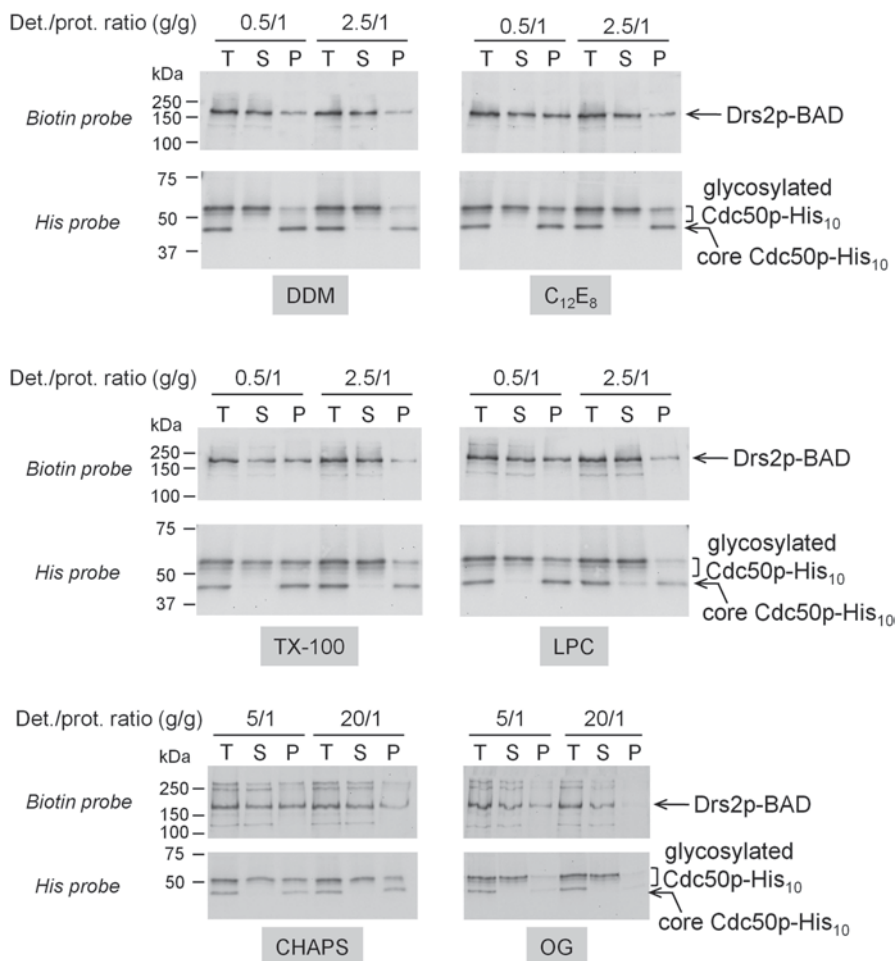


Fig. 6.11 Solubilization of P3 membranes with various detergents. P3 membranes containing Drs2p-BAD and Cdc50p-His₁₀ were diluted to 2 mg proteins per mL in an ice-cold solubilization buffer supplemented with detergent (DDM, C₁₂E₈, TX-100, LPC, CHAPS, or OG), at the indicated detergent to protein ratio. After 1 h of incubation at 4 °C, detergent-treated membranes (“T”) were centrifuged at 100,000 g for 1 h at 4 °C and soluble (“S”) material was separated from the insoluble (“P”) fraction. The “P” fraction was resuspended with solubilization buffer in the same volume as that of the “T” fraction. 0.5 μg of the “T” fraction was loaded onto SDS-PAGE and the same volumes of “S” and “P” fractions were loaded as well. Drs2p-BAD was detected using a biotin probe and Cdc50p-His₁₀ was detected using a His probe. (Figure adapted from Jacquot et al. 2012)

expressed in yeast membranes (Lenoir et al. 2002; Jidenko et al. 2006), and C₁₂E₈, which up to now remains the favorite detergent for crystallization of P-type ATPases (Toyoshima et al. 2000; Pedersen et al. 2007; Gourdon et al. 2011), were included in this screen (Fig. 6.11). As can be seen from Fig. 6.11, DDM allowed almost complete solubilization of Drs2p-BAD (compare “T” for total and “S” for supernatant lanes) at DDM-to-protein ratios as low as 0.5 g/g. Solubilization with C₁₂E₈

proved to also be efficient, although perhaps slightly less than DDM, with about 50% Drs2p-BAD solubilized for a 0.5 g/g detergent to protein ratio, while increasing this ratio to 2.5 g/g allowed almost complete solubilization of Drs2p-BAD. Triton X-100 (TX-100), another mild detergent, L- α -lysophosphatidylcholine (LPC), a zwitterionic detergent, or high critical micelle concentration (cmc) detergents like CHAPS and octyl-glucoside (OG) were also tested and turned out to solubilize Drs2p-BAD as efficiently as C₁₂E₈ (Fig. 6.11), although at higher concentrations for CHAPS and OG. Remarkably, the unglycosylated portion of Cdc50p-His₁₀ remained insoluble in all cases, while solubilization of the glycosylated, mature form displayed a similar dependence on detergent concentration as that of Drs2p-BAD (Fig. 6.11; Jacquot et al. 2012).

The ability of these detergents to solubilize both Drs2p-BAD and the glycosylated portion of Cdc50p-His₁₀ provided us with a tool for investigating whether both proteins interact with each other after solubilization. Indeed, incubating DDM-solubilized P3 membranes with Ni²⁺-NTA beads, to fish out Cdc50p *via* its C-terminal decahistidine tag, indicated that both Cdc50p-His₁₀ and Drs2p-BAD were retained to the beads and therefore that they do interact with each other after solubilization with detergent (not shown here, see (Jacquot et al. 2012)).

The choice of the detergent to be used for solubilization and purification of MPs remains largely empirical. A first approach consists in screening several detergents and trying to figure out at the end of the process which one is the most suitable for preserving the activity of the purified target. An alternative approach consists in setting up a functional assay already usable with crude membranes enriched with the protein of interest, to guide the screening of the many conceivable conditions for solubilization and purification. For this purpose, we resorted to phosphorylation of Drs2p from [γ -³²P]ATP. We first tested short-term effects of detergents; for such experiments, detergent was added to P3 membranes prepared from yeast co-expressing Cdc50p-His₁₀ and Drs2p-BAD, and phosphorylation from [γ -³²P]ATP was measured after only short incubation in the presence of these detergents. Open bars in Fig. 6.12a show the phosphorylation levels measured for wild-type Drs2p while gray bars show the level for the inactive D560N mutant, the difference being characteristic of active Drs2p. Some of the detergents tested (all at a concentration above the cmc) left the phosphorylation level at a steady state essentially unaltered (e.g., C₁₂E₈ and CHAPS) while others like TX-100, digitonin, and diC₇PC led to rather low levels of phosphorylation. The phosphorylation level measured in the presence of DDM was even slightly higher than that measured in native membranes (Fig. 6.12a).

As detergents are not only known to alter the steady-state level of phosphorylation of P-type ATPases but also to make these MPs more prone to time-dependent inactivation (e.g., Lund et al. 1989), membranes were incubated with detergent for various periods of time before the phosphorylation measurement took place (Fig. 6.12b). Both C₁₂E₈ and TX-100 turned out to inactivate Drs2p rather rapidly, since phosphorylatability of Drs2p was almost completely lost after 3 h of incubation in those detergents (Fig. 6.12b, left panel), at variance with DDM, which made it possible to keep phosphorylatable a substantial fraction of Drs2p, even after several hours incubation. Using CHAPS, either in the absence or in the presence of

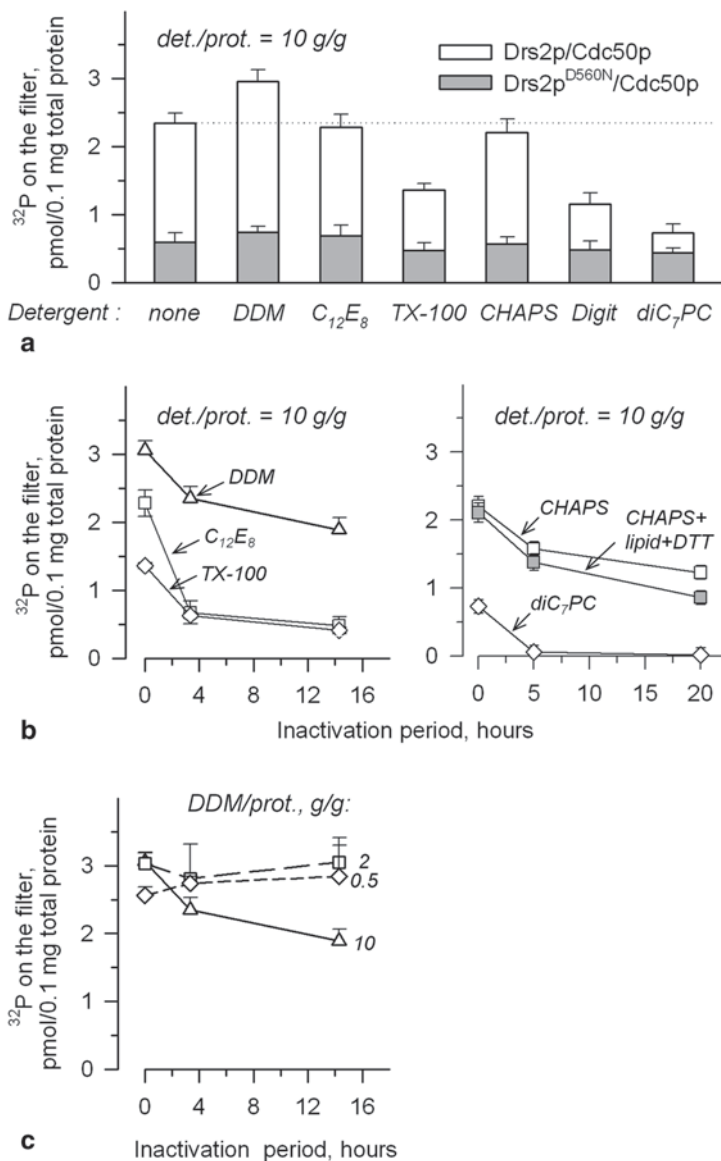


Fig. 6.12 Short-term and long-term effects of detergents on Drs2p/Cdc50p ability to become phosphorylated from ATP. P3 membranes containing either Drs2p-BAD/Cdc50p-His₁₀ (white bars) or Drs2p^{D560N}-BAD/Cdc50p-His₁₀ (gray bars) were suspended at 2 mg/ml in solubilization buffer. **a** Samples were incubated for 1 min in the presence of various detergents, all at 20 mg/mL. The detergents used are indicated (*Digit*, digitonin; *TX-100*, Triton X-100). Phosphorylation from [γ -³²P]ATP was then carried out as described in legend to Fig. 6.9. **b** For selected detergents, similar measurements were made after incubation on ice for various periods of time. **c** DDM was tested at various concentrations, resulting in various detergent-to-protein ratios. Data are presented as the mean \pm SD ($n=3$). (Figure adapted from Jacquot et al. 2012)

egg lecithin and DTT (Coleman et al. 2009), did not improve stability of the Drs2p/Cdc50p complex (Fig. 6.12b, right panel). The short-chain lipid diC₇PC, previously described not to be toxic in the case of a number of other MPs (e.g., see Hauser 2000), was very poor at preserving Drs2p from inactivation (Fig. 6.12b, right panel). Adding selected lipids together with detergent protected the ability of Drs2p to become phosphorylated (data not shown).

Although phosphorylation is not a true indicator of enzyme turnover, this suggested that DDM might be a good candidate for further purification of the complex. We thus explored whether reducing the detergent to protein ratio (from 10 to 2 or 0.5 g/g) would help maintaining Drs2p phosphorylation at its initial level for longer incubation periods. As displayed in Fig. 6.12c, lower detergent to protein ratios (but still solubilizing ones, see Fig. 6.11) indeed kept Drs2p stable over hours of incubation on ice, probably as a result of a less dramatic delipidation of the hydrophobic surface of the protein.

6.5 Conclusions

The overexpression system described in this chapter has thus been successful for crystallization of a mammalian MP, the Serca1a Ca²⁺-ATPase, as well as for the coordinated expression of the yeast Drs2p/Cdc50p complex (Jidenko et al. 2005; Jacquot et al. 2012). For the latter, the next stage will consist in purifying the complex in a functional form to gain insights into the molecular mechanism for lipid transport.

Beyond the examples of rabbit Serca1a and of yeast Drs2p/Cdc50p complex, overexpression in *S. cerevisiae* of other MPs has been attempted (see Table 6.2 for details). For instance, the human cardiac Ca²⁺-ATPase Serca2a was overexpressed using our system and purified by single-step affinity chromatography (Antaloea et al. 2013).

In our laboratory, another Ca²⁺-ATPase has been subjected to heterologous expression in *S. cerevisiae*, namely PfATP6, a Ca²⁺-ATPase from *Plasmodium falciparum*. After overexpression, the purification yield of PfATP6 is similar to that of the rabbit Serca1a, and although the two proteins share a rather high degree of sequence similarity they exhibit significant differences with respect to their sensitivity to known SERCA inhibitors, e.g., thapsigargin and cyclopiazonic acid (Cardi et al. 2010b; Arnou et al. 2011). At variance with studies from others, our work demonstrated that PfATP6 is probably not the target of the antimalarial drug artemisinin; recently, our purified recombinant PfATP6 was used for screening from a compound library inhibitors that might potentially be used as new antimalarial drugs (David-Bosne et al. 2013).

Also, we did not restrict our system to overexpression of P-type ATPases. A transporter that belongs to the mitochondrial carrier family (MCF), the human ADP/ATP translocase AAC1 (see, for instance, Nury et al. 2006 for review), was overexpressed in *S. cerevisiae*. The expression level of AAC1 is about twice higher than that of Serca1a and after purification, AAC1 is functional with respect to

Table 6.2 Membrane proteins successfully expressed using the pYeDP60/BAD system

Protein names and functions	Organism and organelle	Number of TM	MW, kDa	Specific features and comments	References
<i>Ca²⁺ transporters</i>					
Serca1a	Rabbit, ER	10	110	WT and mutants	Centeno et al. 1994; Lenoir et al. 2002; Jidenko et al. 2006; Cardi 2010a
Serca2a	Human, ER	10	115	WT	Antaloae et al. 2013
PfATP6	<i>P. falciparum</i> , ER	10 (?)	140	WT	Cardi et al. 2010b; Arnou et al. 2011; David-Bosne et al. 2013
<i>Lipid transporter</i>					
Drs2p/Cdc50p complex	Yeast, TGN	10/2 (?)	150/40	WT and mutants/ glycosylated (Cdc50p)	Jacquot et al. 2012
<i>ADP/ATP translocase</i>					
hAAC1	Human, Mitochondria	6	33		
<i>K⁺ channel</i>					
TREK-1	Mouse, PM	4–6 (?)	47	Glycosylated	Berrier et al. 2013

binding of substrates and inhibitors (unpublished results). The two-pore domain eukaryotic K⁺ channel TREK-1, that plays an essential role in setting the neuronal membrane potential, was also successfully expressed, purified, and reconstituted in proteoliposomes. Subsequent electrophysiological recordings on the reconstituted protein established that TREK-1 is a mechanosensitive channel directly sensitive to a change in membrane tension (Berrier et al. 2013).

At the present time, the high-resolution structure of these MPs (beyond Serca1a) has not been solved, but for some of them, crystallization trials are in progress.

Acknowledgments We wish to thank Raphaëlle Barry for her initial help in the “flippase” project, and specifically for construction of the co-expression plasmid, and Stéphanie David-Bosne for stimulating discussions.

This work was supported by the French Infrastructure for Integrated Structural Biology (FRISBI) and by grants from the Agence Nationale pour la Recherche and the Ile de France region (Domaine d’Intérêt Majeur Maladies Infectieuses, DIM MALINF).

References

- Alder-Baerens N, Lisman Q, Luong L, Pomorski T, Holthuis JC (2006) Loss of P4 ATPases Drs2p and Dnf3p disrupts aminophospholipid transport and asymmetry in yeast post-Golgi secretory vesicles. *Mol Biol Cell* 17:1632–1642
- Andersen JP (1995) Dissection of the functional domains of the sarcoplasmic reticulum Ca²⁺-ATPase by site-directed mutagenesis. *Biosci Rep* 15:243–261

- Andersen JP, Vilsen B (1992) Functional consequences of alterations to Glu309, Glu771, and Asp800 in the Ca²⁺-ATPase of sarcoplasmic reticulum. *J Biol Chem* 267:19383–19387
- Antaloe AV, Montigny C, le Maire M, Watson KA, Sorensen TL (2013) Optimisation of recombinant production of active human cardiac SERCA2a ATPase. *PLoS One* 8:e71842
- Arnou B, Montigny C, Morth JP, Nissen P, Jaxel C, Moller JV, Maire M (2011) The Plasmodium falciparum Ca²⁺-ATPase PfATP6: insensitive to artemisinin, but a potential drug target. *Biochem Soc Trans* 39:823–831
- Arsenieva D, Symersky J, Wang Y, Pagadala V, Mueller DM (2010) Crystal structures of mutant forms of the yeast F1 ATPase reveal two modes of uncoupling. *J Biol Chem* 285:36561–36569
- Berrier C, Pozza A, de Lacroix deLA, Chardonnet S, Mesneau A, Jaxel C, le Maire M, Ghazi A (2013) The purified mechanosensitive channel TREK-1 is directly sensitive to membrane tension. *J Biol Chem* 288(38):27307–27314
- Bigelow DJ, Inesi G (1992) Contributions of chemical derivatization and spectroscopic studies to the characterization of the Ca²⁺ transport ATPase of sarcoplasmic reticulum. *Biochim Biophys Acta* 1113:323–338
- Bill RM, Henderson PJ, Iwata S, Kunji ER, Michel H, Neutze R, Newstead S, Poolman B, Tate CG, Vogel H (2011) Overcoming barriers to membrane protein structure determination. *Nat Biotechnol* 29:335–340
- Binda C, Newton-Vinson P, Hubalek F, Edmondson DE, Mattevi A (2002) Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. *Nat Struct Biol* 9:22–26
- Blagovic B, Rupcic J, Mesaric M, Georgiu K, Maric V (2001) Lipid composition of brewer's yeast. *Food Technol Biotechnol* 39:175–181
- Blagovic B, Rupcic J, Mesaric M, Maric V (2005) Lipid analysis of the plasma membrane and mitochondria of brewer's yeast. *Folia Microbiol (Praha)* 50:24–30
- Bleve G, Di Sansebastiano GP, Grieco F (2011) Over-expression of functional *Saccharomyces cerevisiae* GUP1, induces proliferation of intracellular membranes containing ER and Golgi resident proteins. *Biochim Biophys Acta* 1808:733–744
- Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14:115–132
- Brandl CJ, Green NM, Korczak B, MacLennan DH (1986) Two Ca²⁺ ATPase genes: homologies and mechanistic implications of deduced amino acid sequences. *Cell* 44:597–607
- Britton Z, Young C, Can O, McNeely P, Naranjo A, Robinson A (2011) Membrane protein expression in *Saccharomyces cerevisiae*. In: Robinson A (ed) *Production of membrane proteins: strategies for production and isolation*. Wiley-VCH, Weinheim
- Bryde S, Hennrich H, Verhulst PM, Devaux PF, Lenoir G, Holthuis JC (2010) CDC50 proteins are critical components of the human class-1 P4-ATPase transport machinery. *J Biol Chem* 285:40562–40572
- Bull LN, van Eijk MJ, Pawlikowska L, DeYoung JA, Juijn JA, Liao M, Klomp LW, Lomri N, Berger R, Scharschmidt BF, Knisely AS, Houwen RH, Freimer NB (1998) A gene encoding a P-type ATPase mutated in two forms of hereditary cholestasis. *Nat Genet* 18:219–224
- Canadi Juresic G, Blagovic B (2011) The influence of fermentation conditions and recycling on the phospholipid and fatty acid composition of the brewer's yeast plasma membranes. *Folia Microbiol (Praha)* 56:215–224
- Cardi D, Montigny C, Arnou B, Jidenko M, Marchal E, le Maire M, Jaxel C (2010a) Heterologous expression and affinity purification of eukaryotic membrane proteins in view of functional and structural studies: the example of the sarcoplasmic reticulum Ca²⁺-ATPase. *Method Mol Biol* 601:247–267
- Cardi D, Pozza A, Arnou B, Marchal E, Clausen JD, Andersen JP, Krishna S, Moller JV, le Maire M, Jaxel C (2010b) Purified E255 L mutant SERCA1a and purified PfATP6 are sensitive to SERCA-type inhibitors but insensitive to artemisinins. *J Biol Chem* 285:26406–26416
- Carpenter EP, Beis K, Cameron AD, Iwata S (2008) Overcoming the challenges of membrane protein crystallography. *Curr Opin Struct Biol* 18:581–586

- Celik E, Calik P (2012) Production of recombinant proteins by yeast cells. *Biotechnol Adv* 30:1108–1118
- Centeno F, Deschamps S, Lompre AM, Anger M, Moutin MJ, Dupont Y, Palmgren MG, Villalba JM, Moller JV, Falson P et al (1994) Expression of the sarcoplasmic reticulum Ca^{2+} -ATPase in yeast. *FEBS Lett* 354:117–122
- Chapman-Smith A, Cronan JE Jr (1999) The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem Sci* 24:359–363
- Chen CY, Ingram MF, Rosal PH, Graham TR (1999) Role for Drs2p, a P-type ATPase and potential aminophospholipid translocase, in yeast late Golgi function. *J Cell Biol* 147:1223–1236
- Chen S, Wang J, Muthusamy BP, Liu K, Zare S, Andersen RJ, Graham TR (2006) Roles for the Drs2p-Cdc50p complex in protein transport and phosphatidylserine asymmetry of the yeast plasma membrane. *Traffic* 7:1503–1517
- Chiba Y, Akeboshi H (2009) Glycan engineering and production of ‘humanized’ glycoprotein in yeast cells. *Biol Pharm Bull* 32:786–795
- Chiba Y, Jigami Y (2007) Production of humanized glycoproteins in bacteria and yeasts. *Curr Opin Chem Biol* 11:670–676
- Clarke DM, Loo TW, Inesi G, MacLennan DH (1989) Location of high affinity Ca^{2+} -binding sites within the predicted transmembrane domain of the sarcoplasmic reticulum Ca^{2+} -ATPase. *Nature* 339:476–478
- Coleman JA, Kwok MC, Molday RS (2009) Localization, purification, and functional reconstitution of the P4-ATPase Atp8a2, a phosphatidylserine flippase in photoreceptor disc membranes. *J Biol Chem* 284:32670–32679
- Coleman JA, Quazi F, Molday RS (2012) Mammalian P4-ATPases and ABC transporters and their role in phospholipid transport. *Biochim Biophys Acta* 1831:555–574
- Conslar TG, Persson BL, Jung H, Zen KH, Jung K, Prive GG, Verner GE, Kaback HR (1993) Properties and purification of an active biotinylated lactose permease from *Escherichia coli*. *Proc Natl Acad Sci U S A* 90:6934–6938
- Cronan JE Jr (1990) Biotinylation of proteins in vivo. A post-translational modification to label, purify, and study proteins. *J Biol Chem* 265:10327–10333
- Daleke DL (2007) Phospholipid flippases. *J Biol Chem* 282:821–825
- Daum G, Tuller G, Nemeč T, Hrastnik C, Balliano G, Cattel L, Milla P, Rocco F, Conzelmann A, Vionnet C, Kelly DE, Kelly S, Schweizer E, Schuller HJ, Hojad U, Greiner E, Finger K (1999) Systematic analysis of yeast strains with possible defects in lipid metabolism. *Yeast* 15:601–614
- David-Bosne S, Florent I, Lund-Winther AM, Hansen JB, Buch-Pedersen M, Machillot P, le Maire M, Jaxel C (2013) Antimalarial screening via large-scale purification of *Plasmodium falciparum* Ca^{2+} -ATPase 6 and in vitro studies. *FEBS J* 280(21):5419–5429
- Decottignies A, Grant AM, Nichols JW, de Wet H, McIntosh DB, Goffeau A (1998) ATPase and multidrug transport activities of the overexpressed yeast ABC protein Yor1p. *J Biol Chem* 273:12612–12622
- Degani C, Boyer PD (1973) A borohydride reduction method for characterization of the acyl phosphate linkage in proteins and its application to sarcoplasmic reticulum adenosine triphosphatase. *J Biol Chem* 248:8222–8226
- Diers IV, Rasmussen E, Larsen PH, Kjaersig IL (1991) Yeast fermentation processes for insulin production. *Bioprocess Technol* 13:166–76.
- Devaux PF, Herrmann A, Ohlwein N, Kozlov MM (2008) How lipid flippases can modulate membrane structure. *Biochim Biophys Acta* 1778:1591–1600
- Dux L, Martonosi A (1983) Two-dimensional arrays of proteins in sarcoplasmic reticulum and purified Ca^{2+} -ATPase vesicles treated with vanadate. *J Biol Chem* 258:2599–2603
- Ebashi S, Lipmann F (1962) Adenosine triphosphate-linked concentration of calcium ions in a particulate fraction of rabbit muscle. *J Cell Biol* 14:389–400
- Fadok VA, Bratton DL, Rose DM, Pearson A, Ezekewitz RA, Henson PM (2000) A receptor for phosphatidylserine-specific clearance of apoptotic cells. *Nature* 405:85–90

- Furuta N, Fujimura-Kamada K, Saito K, Yamamoto T, Tanaka K (2007) Endocytic recycling in yeast is regulated by putative phospholipid translocases and the Ypt31p/32p-Rcy1p pathway. *Mol Biol Cell* 18:295–312
- Gall WE, Geething NC, Hua Z, Ingram MF, Liu K, Chen SI, Graham TR (2002) Drs2p-dependent formation of exocytic clathrin-coated vesicles in vivo. *Curr Biol* 12:1623–1627
- Gemmill TR, Trimble RB (1999) Overview of N- and O-linked oligosaccharide structures found in various yeast species. *Biochim Biophys Acta* 1426:227–237
- Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. *Science* 274(546):563–547
- Gourdon P, Liu XY, Skjorringe T, Morth JP, Moller LB, Pedersen BP, Nissen P (2011) Crystal structure of a copper-transporting PIB-type ATPase. *Nature* 475:59–64
- Graham TR (2004) Flippases and vesicle-mediated protein transport. *Trends Cell Biol* 14:670–677
- Griffith DA, Delipala C, Leadsham J, Jarvis SM, Oesterhelt D (2003) A novel yeast expression system for the overproduction of quality-controlled membrane proteins. *FEBS Lett* 553:45–50
- Grisshammer R, Tate CG (1995) Overexpression of integral membrane proteins for structural studies. *Q Rev Biophys* 28:315–422
- Guarente L, Yocum RR, Gifford P (1982) A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulatory region as an upstream site. *Proc Natl Acad Sci U S A* 79:7410–7414
- Gunge N (1983) Yeast DNA plasmids. *Annu Rev Microbiol* 37:253–276
- Hasselbach W, Makinose M (1961) [The calcium pump of the “relaxing granules” of muscle and its dependence on ATP-splitting]. *Biochem Z* 333:518–528
- Hauser H (2000) Short-chain phospholipids as detergents. *Biochim Biophys Acta* 1508:164–181
- Haviv H, Habeck M, Kanai R, Toyoshima C, Karlsh SJ (2013) Neutral phospholipids stimulate Na, K-ATPase activity: a specific lipid-protein interaction. *J Biol Chem* 288:10073–10081
- Helenius A, Aeby M (2001) Intracellular functions of N-linked glycans. *Science* 291:2364–2369
- Howard EM, Roepe PD (2003) Purified human MDR 1 modulates membrane potential in reconstituted proteoliposomes. *Biochemistry* 42:3544–3555
- Hua Z, Fatheddin P, Graham TR (2002) An essential subfamily of Drs2p-related P-type ATPases is required for protein trafficking between Golgi complex and endosomal/vacuolar system. *Mol Biol Cell* 13:3162–3177
- Jacquot A, Montigny C, Hennrich H, Barry R, le Maire M, Jaxel C, Holthuis J, Champeil P, Lenoir G (2012) Stimulation by phosphatidylserine of Drs2p/Cdc50p lipid translocase dephosphorylation is controlled by phosphatidylinositol-4-phosphate. *J Biol Chem* 287:13249–13261
- Jidenko M, Nielsen RC, Sorensen TL, Moller JV, le Maire M, Nissen P, Jaxel C (2005) Crystallization of a mammalian membrane protein overexpressed in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 102:11687–11691
- Jidenko M, Lenoir G, Fuentes JM, le Maire M, Jaxel C (2006) Expression in yeast and purification of a membrane protein, SERCA1a, using a biotinylated acceptor domain. *Protein Expr Purif* 48:32–42
- Kapri-Pardes E, Katz A, Haviv H, Mahmmoud Y, Ilan M, Khalfin-Penigel I, Carmeli S, Yarden O, Karlsh SJ (2011) Stabilization of the alpha2 isoform of Na, K-ATPase by mutations in a phospholipid binding pocket. *J Biol Chem* 286:42888–42899
- Kellosalo J, Kajander T, Kogan K, Pokharel K, Goldman A (2012a) The structure and catalytic cycle of a sodium-pumping pyrophosphatase. *Science* 337:473–476
- Kellosalo J, Kajander T, Honkanen R, Goldman A (2012b) Crystallization and preliminary X-ray analysis of membrane-bound pyrophosphatases. *Mol Membr Biol* 30:64–74
- Kinnunen PK, Holopainen JM (2000) Mechanisms of initiation of membrane fusion: role of lipids. *Biosci Rep* 20:465–482
- Kitson SM, Mullen W, Cogdell RJ, Bill RM, Fraser NJ (2011) GPCR production in a novel yeast strain that makes cholesterol-like sterols. *Methods* 55:287–292
- Lagane B, Gaibelet G, Meilhoc E, Masson JM, Cezanne L, Lopez A (2000) Role of sterols in modulating the human mu-opioid receptor function in *Saccharomyces cerevisiae*. *J Biol Chem* 275:33197–33200

- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921
- Lee KM, DaSilva NA (2005) Evaluation of the *Saccharomyces cerevisiae* ADH2 promoter for protein synthesis. *Yeast* 22:431–440
- Lee AG, East JM (1998) The effects of phospholipid structure on the function of a calcium pump. *Biochem Soc Trans* 26:359–365
- Lee JK, Stroud RM (2010) Unlocking the eukaryotic membrane protein structural proteome. *Curr Opin Struct Biol* 20:464–470
- Lenoir G, Menguy T, Corre F, Montigny C, Pedersen PA, Thines D, le Maire M, Falson P (2002) Overproduction in yeast and rapid and efficient purification of the rabbit SERCA1a Ca²⁺-ATPase. *Biochim Biophys Acta* 1560:67–83
- Lenoir G, Picard M, Moller JV, le Maire M, Champeil P, Falson P (2004) Involvement of the L6-7 loop in SERCA1a Ca²⁺-ATPase activation by Ca²⁺ (or Sr²⁺) and ATP. *J Biol Chem* 279:32125–32133
- Lenoir G, Jaxel C, Picard M, le Maire M, Champeil P, Falson P (2006) Conformational changes in sarcoplasmic reticulum Ca²⁺-ATPase mutants: effect of mutations either at Ca²⁺-binding site II or at tryptophan 552 in the cytosolic domain. *Biochemistry* 45:5261–5270
- Lenoir G, Williamson P, Holthuis JC (2007) On the origin of lipid asymmetry: the flip side of ion transport. *Curr Opin Chem Biol* 11:654–661
- Lenoir G, Williamson P, Puts CF, Holthuis JC (2009) Cdc50p plays a vital role in the ATPase reaction cycle of the putative aminophospholipid transporter drs2p. *J Biol Chem* 284:17956–17967
- Li WZ, Sherman F (1991) Two types of TATA elements for the *CYC1* gene of the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* 11:666–676
- Li M, Hays FA, Roe-Zurz Z, Vuong L, Kelly L, Ho CM, Robbins RM, Pieper U, O'Connell JD 3rd, Miercke LJ, Giacomini KM, Sali A, Stroud RM (2009) Selecting optimum eukaryotic integral membrane proteins for structure determination by rapid expression and solubilization screening. *J Mol Biol* 385:820–830
- Lin SM, Tsai JY, Hsiao CD, Huang YT, Chiu CL, Liu MH, Tung JY, Liu TH, Pan RL, Sun YJ (2012) Crystal structure of a membrane-embedded H⁺-translocating pyrophosphatase. *Nature* 484:399–403
- Long SB, Campbell EB, Mackinnon R (2005) Crystal structure of a mammalian voltage-dependent Shaker family K⁺ channel. *Science* 309:897–903
- Lopez-Marques RL, Holthuis JC, Pomorski TG (2011) Pumping lipids with P4-ATPases. *Biol Chem* 392:67–76
- Lund S, Orłowski S, de Foresta B, Champeil P, le Maire M, Moller JV (1989) Detergent structure and associated lipid as determinants in the stabilization of solubilized Ca²⁺-ATPase from sarcoplasmic reticulum. *J Biol Chem* 264:4907–4915
- Ma J, Ito A (2002) Tyrosine residues near the FAD binding site are critical for FAD binding and for the maintenance of the stable and active conformation of rat monoamine oxidase A. *J Biochem* 131:107–111
- Ma J, Yoshimura M, Yamashita E, Nakagawa A, Ito A, Tsukihara T (2004a) Structure of rat monoamine oxidase A and its specific recognitions for substrates and inhibitors. *J Mol Biol* 338:103–114
- Ma J, Kubota F, Yoshimura M, Yamashita E, Nakagawa A, Ito A, Tsukihara T (2004b) Crystallization and preliminary crystallographic analysis of rat monoamine oxidase A complexed with clorgyline. *Acta Crystallogr D Biol Crystallogr* 60:317–319
- Marchand A, Winther AM, Holm PJ, Olesen C, Montigny C, Arnou B, Champeil P, Clausen JD, Vilsen B, Andersen JP, Nissen P, Jaxel C, Moller JV, le Maire M (2008) Crystal structure of D351A and P312A mutant forms of the mammalian sarcoplasmic reticulum Ca²⁺-ATPase reveals key events in phosphorylation and Ca²⁺ release. *J Biol Chem* 283:14867–14882
- Maruyama K, MacLennan DH (1988) Mutation of aspartic acid-351, lysine-352, and lysine-515 alters the Ca²⁺ transport activity of the Ca²⁺-ATPase expressed in COS-1 cells. *Proc Natl Acad Sci U S A* 85:3314–3318

- McAleer WJ, Buynak EB, Maigetter RZ, Wampler DE, Miller WJ, Hilleman MR (1984) Human hepatitis B vaccine from recombinant yeast. *Nature* 307:178–180
- Mercer J, Helenius A (2008) Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells. *Science* 320:531–535
- Meusser B, Hirsch C, Jarosch E, Sommer T (2005) ERAD: the long road to destruction. *Nat Cell Biol* 7:766–772
- Midgett CR, Madden DR (2007) Breaking the bottleneck: eukaryotic membrane protein expression for high-resolution structural studies. *J Struct Biol* 160:265–274
- Miras R, Cuillel M, Catty P, Guillaïn F, Mintz E (2001) Purification of heterologous sarcoplasmic reticulum Ca²⁺-ATPase Serca1a allowing phosphoenzyme and Ca²⁺-affinity measurements. *Protein Expr Purif* 22:299–306
- Morth JP, Pedersen BP, Toustrup-Jensen MS, Sorensen TL, Petersen J, Andersen JP, Vilsen B, Nissen P (2007) Crystal structure of the sodium-potassium pump. *Nature* 450:1043–1049
- Mueller DM, Puri N, Kabaleswaran V, Terry C, Leslie AG, Walker JE (2004) Ni-chelate-affinity purification and crystallization of the yeast mitochondrial F1-ATPase. *Protein Expr Purif* 37:479–485
- Nagy Z, Montigny C, Leverrier P, Yeh S, Goffeau A, Garrigos M, Falson P (2006) Role of the yeast ABC transporter Yor1p in cadmium detoxification. *Biochimie* 88:1665–1671
- Natarajan P, Wang J, Hua Z, Graham TR (2004) Drs2p-coupled aminophospholipid translocase activity in yeast Golgi membranes and relationship to in vivo function. *Proc Natl Acad Sci U S A* 101:10614–10619
- Nury H, Dahout-Gonzalez C, Trezeguet V, Lauquin GJ, Brandolin G, Pebay-Peyroula E (2006) Relations between structure and function of the mitochondrial ADP/ATP carrier. *Annu Rev Biochem* 75:713–741
- Olesen C, Picard M, Winther AM, Gyrop C, Morth JP, Oxvig C, Moller JV, Nissen P (2007) The structural basis of calcium transport by the calcium pump. *Nature* 450:1036–1042
- Op den Kamp JA (1979) Lipid asymmetry in membranes. *Annu Rev Biochem* 48:47–71
- Osterberg M, Kim H, Warringer J, Melen K, Blomberg A, von Heijne G (2006) Phenotypic effects of membrane protein overexpression in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 103:11148–11153
- Pagadala V, Vistain L, Symersky J, Mueller DM (2011) Characterization of the mitochondrial ATP synthase from yeast *Saccharomyces cerevisiae*. *J Bioenerg Biomembr* 43:333–347
- Paulusma CC, Folmer DE, Ho-Mok KS, de Waart DR, Hilarius PM, Verhoeven AJ, Oude Elferink RP (2008) ATP8B1 requires an accessory protein for endoplasmic reticulum exit and plasma membrane lipid flippase activity. *Hepatology* 47:268–278
- Pedersen PA, Rasmussen JH, Joergensen PL (1996) Expression in high yield of pig alpha 1 beta 1 Na, K-ATPase and inactive mutants D369N and D807N in *Saccharomyces cerevisiae*. *J Biol Chem* 271:2514–2522
- Pedersen BP, Buch-Pedersen MJ, Morth JP, Palmgren MG, Nissen P (2007) Crystal structure of the plasma membrane proton pump. *Nature* 450:1111–1114
- Pedersen BP, Kumar H, Waight AB, Risenmay AJ, Roe-Zurz Z, Chau BH, Schlessinger A, Bonomi M, Harries W, Sali A, Johri AK, Stroud RM (2013) Crystal structure of a eukaryotic phosphate transporter. *Nature* 496:533–536
- Pomorski T, Lombardi R, Riezman H, Devaux PF, van Meer G, Holthuis JC (2003) Drs2p-related P-type ATPases Dnf1p and Dnf2p are required for phospholipid translocation across the yeast plasma membrane and serve a role in endocytosis. *Mol Biol Cell* 14:1240–1254
- Pompon D (1988) cDNA cloning and functional expression in yeast *Saccharomyces cerevisiae* of beta-naphthoflavone-induced rabbit liver P-450 LM4 and LM6. *Eur J Biochem* 177:285–293
- Pompon D, Louerat B, Bronine A, Urban P (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Method Enzymol* 272:51–64
- Poulsen LR, Lopez-Marques RL, McDowell SC, Okkeri J, Licht D, Schulz A, Pomorski T, Harper JF, Palmgren MG (2008a) The *Arabidopsis* P4-ATPase ALA3 localizes to the Golgi and requires a beta-subunit to function in lipid translocation and secretory vesicle formation. *Plant Cell* 20:658–676

- Poulsen LR, Lopez-Marques RL, Palmgren MG (2008b) Flippases: still more questions than answers. *Cell Mol Life Sci* 65:3119–3125
- Pouny Y, Weitzman C, Kaback HR (1998) In vitro biotinylation provides quantitative recovery of highly purified active lactose permease in a single step. *Biochemistry* 37:15713–15719
- Powl AM, East JM, Lee AG (2008) Importance of direct interactions with lipids for the function of the mechanosensitive channel MscL. *Biochemistry* 47:12175–12184
- Prive GG (2007) Detergents for the stabilization and crystallization of membrane proteins. *Methods* 41:388–397
- Pryor EE Jr, Horanyi PS, Clark KM, Fedoriw N, Connelly SM, Koszelak-Rosenblum M, Zhu G, Malkowski MG, Wiener MC, Dumont ME (2013) Structure of the integral membrane protein CAAX protease Ste24p. *Science* 339:1600–1604
- Putz CF, Panatala R, Hennrich H, Tsareva A, Williamson P, Holthuis JC (2012) Mapping functional interactions in a heterodimeric phospholipid pump. *J Biol Chem* 287:30529–30540
- Romanos MA, Scorer CA, Clare JJ (1992) Foreign gene expression in yeast: a review. *Yeast* 8:423–488
- Rosing J, Tans G, Govers-Riemslog JW, Zwaal RF, Hemker HC (1980) The role of phospholipids and factor Va in the prothrombinase complex. *J Biol Chem* 255:274–283
- Saito K, Fujimura-Kamada K, Furuta N, Kato U, Umeda M, Tanaka K (2004) Cdc50p, a protein required for polarized growth, associates with the Drs2p P-type ATPase implicated in phospholipid translocation in *Saccharomyces cerevisiae*. *Mol Biol Cell* 15:3418–3432
- Schultz LD, Hofmann KJ, Mylin LM, Montgomery DL, Ellis RW, Hopper JE (1987) Regulated overproduction of the *GAL4* gene product greatly increases expression from galactose-inducible promoters on multi-copy expression vectors in yeast. *Gene* 61:123–133
- Schwarz E, Oesterhelt D, Reinke H, Beyreuther K, Dimroth P (1988) The sodium ion translocating oxalacetate decarboxylase of *Klebsiella pneumoniae*. Sequence of the biotin-containing alpha-subunit and relationship to other biotin-containing enzymes. *J Biol Chem* 263:9640–9645
- Shinoda T, Ogawa H, Cornelius F, Toyoshima C (2009) Crystal structure of the sodium-potassium pump at 2.4 Å resolution. *Nature* 459:446–450
- Son SY, Ma J, Kondou Y, Yoshimura M, Yamashita E, Tsukihara T (2008) Structure of human monoamine oxidase A at 2.2-Å resolution: the control of opening the entry for substrates/inhibitors. *Proc Natl Acad Sci U S A* 105:5739–5744
- Sorensen TL, Moller JV, Nissen P (2004) Phosphoryl transfer and calcium ion occlusion in the calcium pump. *Science* 304:1672–1675
- Sorensen TL, Olesen C, Jensen AM, Moller JV, Nissen P (2006) Crystals of sarcoplasmic reticulum Ca²⁺-ATPase. *J Biotechnol* 124:704–716
- Stock D, Leslie AG, Walker JE (1999) Molecular architecture of the rotary motor in ATP synthase. *Science* 286:1700–1705
- Stolz J, Darnhofer-Demar B, Sauer N (1995) Rapid purification of a functionally active plant sucrose carrier from transgenic yeast using a bacterial biotin acceptor domain. *FEBS Lett* 377:167–171
- Strock C, Cavagna M, Peiffer WE, Sumbilla C, Lewis D, Inesi G (1998) Direct demonstration of Ca²⁺ binding defects in sarco-endoplasmic reticulum Ca²⁺ ATPase mutants overexpressed in COS-1 cells transfected with adenovirus vectors. *J Biol Chem* 273:15104–15109
- Toyoshima C, Nomura H (2002) Structural changes in the calcium pump accompanying the dissociation of calcium. *Nature* 418:605–611
- Toyoshima C, Sasabe H, Stokes DL (1993) Three-dimensional cryo-electron microscopy of the calcium ion pump in the sarcoplasmic reticulum membrane. *Nature* 362:467–471
- Toyoshima C, Nakasako M, Nomura H, Ogawa H (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature* 405:647–655
- Toyoshima C, Norimatsu Y, Iwasawa S, Tsuda T, Ogawa H (2007) How processing of aspartylphosphate is coupled to lumenal gating of the ion pathway in the calcium pump. *Proc Natl Acad Sci U S A* 104:19831–19836

- van der Velden LM, Wichers CG, van Breevoort AE, Coleman JA, Molday RS, Berger R, Klomp LW, van deGFSF (2010) Heteromeric interactions required for abundance and subcellular localization of human CDC50 proteins and class 1 P4-ATPases. *J Biol Chem* 285:40088–40096
- van Meer G, Voelker DR, Feigenson GW (2008) Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* 9:112–124
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD et al (2001) The sequence of the human genome. *Science* 291:1304–1351
- Vilsen B, Andersen JP, Clarke DM, MacLennan DH (1989) Functional consequences of proline mutations in the cytoplasmic and transmembrane sectors of the Ca²⁺-ATPase of sarcoplasmic reticulum. *J Biol Chem* 264:21024–21030
- Wach A (1996) PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *Saccharomyces cerevisiae*. *Yeast* 12:259–265
- Waight AB, Pedersen BP, Schlessinger A, Bonomi M, Chau BH, Roe-Zurz Z, Risenmay AJ, Sali A, Stroud RM (2013) Structural basis for alternating access of a eukaryotic calcium/proton exchanger. *Nature* 499:107–110
- Waugh DS (2005) Making the most of affinity tags. *Trends Biotechnol* 23:316–320
- Weis BL, Schleiff E, Zerges W (2013) Protein targeting to subcellular organelles via mRNA localization. *Biochim Biophys Acta* 1833:260–273
- West RW Jr, Yocum RR, Ptashne M (1984) *Saccharomyces cerevisiae GAL1-GAL10* divergent promoter region: location and function of the upstream activating sequence UASG. *Mol Cell Biol* 4:2467–2478
- White SH (1998–2014) Membrane protein of 3D known structure, <http://blanco.biomol.uci.edu/mpstruc/>
- Whorton MR, MacKinnon R (2011) Crystal structure of the mammalian GIRK2 K⁺ channel and gating regulation by G proteins, PIP2, and sodium. *Cell* 147:199–208
- Wildt S, Gerngross TU (2005) The humanization of N-glycosylation pathways in yeast. *Nat Rev Microbiol* 3:119–128
- Zhang P, Toyoshima C, Yonekura K, Green NM, Stokes DL (1998) Structure of the calcium pump from sarcoplasmic reticulum at 8-Å resolution. *Nature* 392:835–839
- Zhang Z, Lewis D, Strock C, Inesi G, Nakasako M, Nomura H, Toyoshima C (2000) Detailed characterization of the cooperative mechanism of Ca²⁺ binding and catalytic activation in the Ca²⁺ transport (SERCA) ATPase. *Biochemistry* 39:8758–8767
- Zimmermann R, Eyrisch S, Ahmad M, Helms V (2011) Protein translocation across the ER membrane. *Biochim Biophys Acta* 1808:912–924

Chapter 7

Amphipols: A General Introduction and Some Protocols

Manuela Zoonens, Francesca Zito, Karen L. Martinez and Jean-Luc Popot

7.1 General Introduction to the Properties and Uses of Amphipols

7.1.1 *Inactivation of Membrane Proteins in the Presence of Detergents and Strategies for Membrane Protein Stabilization*

In their native environment, i.e., the membrane, membrane proteins (MPs) are stabilized by various types of physical factors provided by the membrane architecture itself, such as membrane thickness, accessibility to water, the distribution of charges, lipid asymmetry, or other properties like transmembrane gradients, viscosity, etc. However, specific molecular interactions between proteins, lipids, and cofactors play a key role in MP stability (for discussions, see, e.g., Bowie 2001; Garavito and Ferguson-Miller 2001; Popot and Engelman 2000). Detergents compete with these interactions, inducing destabilization. In other words, most detergents can be considered as being too good a solvent, breaking more interactions than would be desirable. The extent of this problem varies from one detergent to another, leading to distinguishing “weak” and “strong” detergents. It also varies considerably depending on the nature of MPs and, in particular, tends to be more severe for α -helical than for β -barrel MPs. This variability creates a bias in our understanding of the structure and function of MPs, representing an important bottleneck.

M. Zoonens (✉) · F. Zito
Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099,
Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France
e-mail: manuela.zoonens@ibpc.fr

J.-L. Popot
Institute of Physico-Chemical Biology,
UMR-CNRS 7099, Université Paris Diderot, Paris, France

K. L. Martinez
Bio-Nanotechnology Laboratory, Department of Neuroscience and Pharmacology &
Nano-Science Center, University of Copenhagen, Copenhagen, Denmark

I. Mus-Veteau (ed.), *Membrane Proteins Production for Structural Analysis*,
DOI 10.1007/978-1-4939-0662-8_7, © Springer Science+Business Media New York 2014

To improve the stability of detergent-solubilized MPs, several approaches can be resorted to, such as: (1) transfer to a “weak” detergent like Tween or digitonin, (2) working close to the critical micellar concentration (CMC) of the detergent in order to limit the volume of the micellar phase, (3) supplementing the micelles with lipids or cofactors, (4) working fast enough so as to collect data or form crystals before the inactivation of the protein sets in, or (5) select or engineer more stable MPs. An alternative is to replace classical detergents by bilayer-like environments such as lipid vesicles, bicelles, nanodiscs, or cubic phases, or by novel, less aggressive surfactants such as detergents with multiple, branched, or cyclic hydrophobic moieties (see, e.g., Chae et al. 2010; Hong et al. 2011; Hovers et al. 2011) or stabilizing polar heads (Matar-Merheb et al. 2011), fluorinated surfactants (Breyton et al. 2004, 2009, 2010; Chabaud et al. 1998; Popot 2010; see also Chap. 8 in this volume), amphipathic peptides (Koutsopoulos et al. 2012; Schafmeister et al. 1993; Wang et al. 2011; Zhao et al. 2006), lipopeptides (McGregor et al. 2003; Privé 2009), or polymers such as styrene maleic acid lipid particles (SMALPs; Knowles et al. 2009; Long et al. 2013; Rajesh et al. 2011), and amphipols (Popot et al. 2011; Tribet et al. 1996; Zoonens and Popot 2014).

7.1.2 General Properties of APols in Aqueous Solution

Amphipols (APols) are short and flexible amphipathic polymers, designed so as to bind to the transmembrane domain of MPs by multiple hydrophobic contact points. MP/APol complexes, as a result, should present a low k_{off} and a small K_{D} . They should not dissociate even at extreme dilutions and, when they do, should do so extremely slowly (see below). This would make them radically different from MP/detergent complexes, in which the protein-bound detergent molecules are in rapid equilibrium with free monomers and micelles, and dissociate upon dilution below the CMC. The first APols to have been synthesized comprise a polyacrylic acid backbone onto which octylamine and isopropylamine side chains have been randomly grafted (Tribet et al. 1996; Fig. 7.1). The most widely used APol, called A8-35, features 35% of ungrafted carboxylic acid groups. Above pH 7, all of those are ionized (Gohon et al. 2004). Twenty-five percent of the carboxylic groups have been derivatized with octyl chains, giving A8-35 its amphipathy, and the last 40% with isopropyl groups, so as to reduce the charge density along the polymer. The average molecular weight (MW) of A8-35 is ~ 4.3 kDa (Giusti et al. 2014b).

The solution properties of A8-35 have been intensively studied and previously reviewed (Popot 2010; Popot et al. 2003, 2011; Zoonens and Popot 2014). A8-35 is highly soluble in water (> 200 g L⁻¹). Its concentration can reach up to 100 g L⁻¹ without affecting significantly the viscosity of the solution (L.J. Catoire, personal communication). In aqueous solutions, its molecules self-associate to form small, compact, hydrated particles with an average MW of ~ 40 kDa (Gohon et al. 2006). Based on an average MW of ~ 4.3 kDa per individual molecule, A8-35 particles therefore contain an average of ~ 9 – 10 molecules, corresponding to ~ 80 octyl chains.

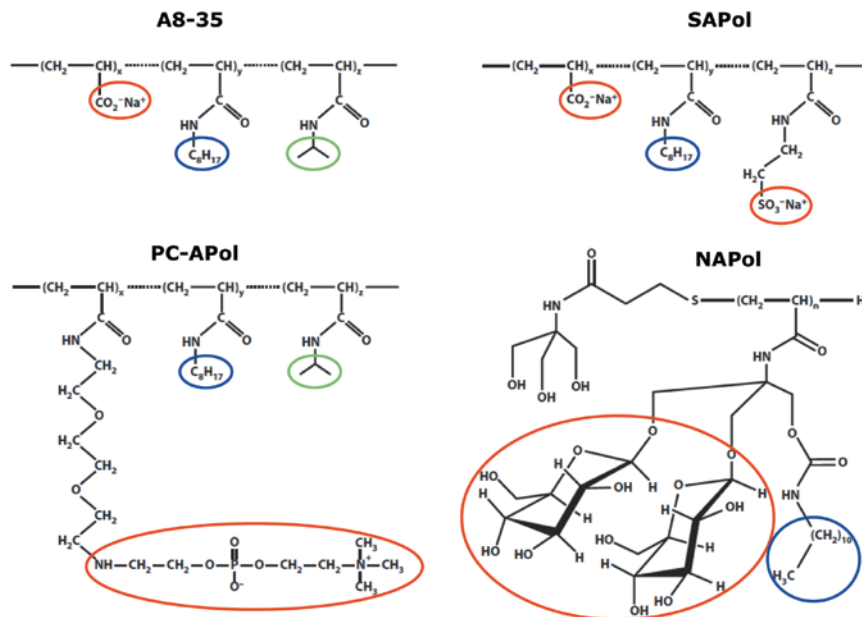


Fig. 7.1 Molecular structures of various amphipols: A8-35, SAPol, PC-APol, and NAPol. The groups conferring the aqueous solubility to the polymers are *circled in red*, the alkyl chains in *blue*, and the groups that modulate the charge density in *green*. The three types of groups are randomly distributed along the chain. For A8-35, the molar percentages of each group are $x=35\%$, $y=25\%$, and $z=40\%$. The cartoons are reprinted with permissions from Proc Natl Acad Sci USA 93:15047–15050, copyright 1996 National Academy of Sciences, USA (Tribet et al. 1996), Biopolymers 95:811–823, copyright 2011 Wiley Periodicals, Inc. (Dahmane et al. 2011), Langmuir 23:3025–3035, copyright 2007 American Chemical Society (Diab et al. 2007b) and Langmuir 28:4625–4639, copyright 2012 American Chemical Society (Sharma et al. 2012), respectively.

The minimal concentration at which A8-35 particles start to assemble—the critical aggregation concentration (CAC)—is $\sim 0.002 \text{ g L}^{-1}$ (Giusti et al. 2012). Whereas the general organization of A8-35 particles and detergent micelles resemble each other, with similar sizes and the presence of a hydrophobic core and a hydrophilic surface, differences include the smaller number of molecules per A8-35 particle, a much slower rate of exchange with the solution (expected from indirect evidence, but not measured directly yet), and a higher viscosity: According to molecular dynamics (MD) calculations, the backbone of A8-35 in a particle moves $\sim 10\times$ more slowly than the hydrophilic head group of micellar sodium dodecyl sulfate (SDS; Perlmutter et al. 2011). Upon size-exclusion chromatography (SEC), APol particles show a Stokes radius, R_s , of 3.15 nm (Gohon et al. 2006), with a highly homogeneous size distribution provided that the pH of the solution is ≥ 7 (Gohon et al. 2004, 2006) and divalent cations are absent (Picard et al. 2006). Indeed, the solubility of A8-35 particles being due to their negative charges, the protonation or complexation of carboxylate groups makes them less hydrophilic, leading them to aggregate

Table 7.1 Library of A8-35-related APols. (Adapted from Le Bon et al. 2014a, b)

APols	Type of modification with respect to A8-35
A8-35	None
A8-75	Charge density $\sim 2 \times$ higher
A34-35	Molecules $\sim 4 \times$ longer
A34-75	Molecules $\sim 4 \times$ longer, charge density $\sim 2 \times$ higher
Isotopically labeled APols	Isotopes: ^{14}C , ^3H , ^2H
Fluorescent APols (FAPols)	Fluorophores: naphthalene, NBD, fluorescein, Alexa Fluor 488, rhodamine, Atto 647, Alexa Fluor 647
Tagged APols	Tags: biotine (BAPol), polyhistidine (HistAPol), imidazole groups (ImidAPol), oligodeoxynucleotide (OligAPol, thiamorpholine (SulfidAPol))

References are given in the text

(isolated carboxylates have a pK_a of around 4.8; in A8-35, however, because of their close proximity, some of them start to protonate at or slightly below pH 7). Multivalent cations, such as Ca^{2+} , can bridge particles, leading, depending on the concentration of Ca^{2+} , to the formation of small oligomers or to massive precipitation (Diab et al. 2007a; Picard et al. 2006).

7.1.3 A Library of APols

APols with chemical structures different from that of A8-35 have been designed in order to make them insensitive to pH (Fig. 7.1). Replacing the isopropyl groups by sulfonate groups yields sulfonated APols (SAPols), which remain water soluble even at pH 2 (Dahmane et al. 2011). Zwitterionic APols with phosphocholine polar head groups (PC-APols; Diab et al. 2007a, b; Tribet et al. 2009) and nonionic APols (NAPols) carrying sugar groups (Bazzacco et al. 2012; Sharma et al. 2012) have also been synthesized and validated. This new generation of APols widens the field of possible applications.

A8-35, even if its pH sensitivity can create limitations in some specific applications, like nuclear magnetic resonance (NMR), remains one of the easiest APols to synthesize and to label, and by far the most thoroughly studied one in terms both of its properties and of its applications. This makes it very attractive to diversify, label, and/or functionalize it, and has led to the creation of a library of A8-35 variants specially designed for specific applications (Table 7.1). Variations around the A8-35 structure concern, for instance, the percentage of charges (A8-75), the length of the backbone (A34-35), or both (A34-75; Tribet et al. 1996). Isotopically labeled or functionalized versions of A8-35 have also been synthesized and validated. For

example, A8-35 and A8-75 have been labeled with such isotopes as ^{14}C (Tribet et al. 1997), ^3H (Gohon et al. 2008), and ^2H (Giusti et al. 2014b; Gohon et al. 2004, 2006). ^{14}C - and ^3H -labeled APols are useful as tracers for the detection of APols in solution, and have been used to quantify the amount of MP-bound A8-35 (Gohon et al. 2008; Tribet et al. 1997), whereas the deuterated versions have been used in NMR, analytical ultracentrifugation (AUC), and small-angle neutron scattering (SANS) experiments (see e.g., Catoire et al. 2009 2010a, ; Gohon et al. 2008; Zoonens et al. 2005).

Grafting a small percentage of a tag or label onto A8-35 does not affect its solution properties (reviewed in Le Bon et al. 2014b), while functionalizing it for specific purposes. Fluorescent APols (FAPols) bearing different fluorophores, which cover a broad range of excitation and emission wavelengths, have thus been synthesized (Fernandez et al. 2014; Giusti et al. 2012; Opačić et al. 2014; Zoonens et al. 2007). Affinity tags have also been grafted onto A8-35, such as biotin (Charvolin et al. 2009), polyhistidine (Giusti et al. 2014a), randomly distributed imidazole groups F. Giusti, unpublished data, an oligonucleotide (Le Bon et al. 2014a), or a thiamorpholine (unpublished data), yielding tagged APols nicknamed BAPol, HistAPol, ImidAPol, OligAPol and SulfidAPol, respectively (Table 7.1). Biotin has also been grafted onto a PC-APol (Basit et al. 2012) and a NAPol (Ferrandez et al. 2014). Tagged APols can be used to immobilize MPs onto solid supports (see below).

7.1.4 General Properties of MP/APol Complexes

The activity, stability, and physical–chemical features of MP/APol complexes have been investigated in some details (for reviews, see Popot 2010; Popot et al. 2003, 2011; Zoonens and Popot 2014). Briefly, APols adsorb exclusively onto the hydrophobic transmembrane region of MPs (Althoff et al. 2011; Cao et al. 2013; Catoire et al. 2009; Liao et al. 2013; Perlmutter et al. 2014; Zoonens et al. 2005), forming a compact layer ca. 1.5–2 nm thick (Althoff et al. 2011; Gohon et al. 2008; Perlmutter et al. 2014). Their association is very stable as long as no competing surfactant is present in the medium (Zoonens et al. 2007). On the other hand, because they are freely miscible with other surfactants, MP-bound APols can be easily displaced by detergents or other APols (Tribet et al. 1997, 2009; Zoonens et al. 2007). Similarly, MPs trapped in APols can be delivered to lipid bilayers (Nagy et al. 2001; Pocanschi et al. 2006) or three-dimensional (3D) lipid phases (Polovinkin et al. 2014).

After trapping, MP/A8-35 complexes are almost—although not completely—as homogeneous as MP/detergent ones, and they present the same sensitivity as pure A8-35 particles to low pH and divalent cations (see, e.g., Gohon et al. 2008; Picard et al. 2006; Zoonens et al. 2007). The size of MP/APol complexes is slightly larger than that of MP/detergent ones, and, as a result, their rotational correlation time, τ_c , is slightly longer (Catoire et al. 2010b). A major difference between APol-trapped and detergent-solubilized MPs is their stability, which is usually much higher in

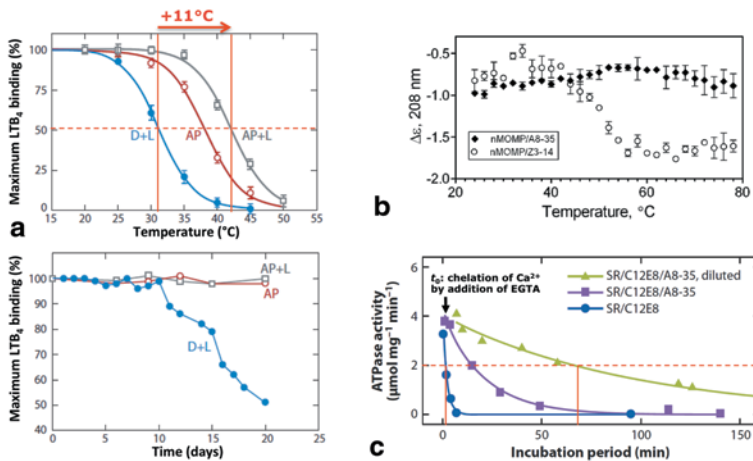


Fig. 7.2 Comparison of membrane protein stability in detergent and after trapping in A8-35. **a** Thermostability curves of a GPCR, the BLT1 leukotriene receptor, in the detergent Fos-choline 16 supplemented with lipids (asolectin) in a 2:1 mass ratio (*blue curve*), in pure A8-35 (*red*), and in a mixture of A8-35 and lipids in a 5:1 mass ratio (*gray*). The stability was monitored by following, during storage at 4°C, the ability of the GPCR to specifically bind LTB₄. Adapted with permission from *Biochemistry* 48:6516–6521, copyright 2009 American Chemical Society (Dahmane et al. 2009). **b** Thermostability of a trimeric porin, MOMP from *Chlamydia trachomatis*, monitored by CD in the detergent Z3-14 (*open symbols*) and after trapping with A8-35 (*solid symbols*). Reprinted from *Vaccine* 29:4623–46231, copyright 2011, with permission from Elsevier (Tifrea et al. 2011). **c** Evolution of the ATPase activity of the calcium pump (Serca-1a) over time after removal of Ca²⁺ from preparations. Sarcoplasmic reticulum vesicles were solubilized in the detergent C₁₂E₈ (*blue curve*), then supplemented with A8-35 (*magenta*), and finally diluted under the CMC of the detergent, so that the protein's environment became predominantly the APol (*green*; adapted from Champeil et al. 2000)

APols (Fig. 7.2). For example, the denaturation temperature of A8-35-trapped BLT1, a G protein-coupled receptor (GPCR) of leukotriene LTB₄, is increased by ~11°C as compared to that in Fos-choline-16 (Fig. 7.2a; Dahmane et al. 2009). Similar effects can be observed for β-barrel MPs, such as the monomeric outer membrane protein OmpA from *Escherichia coli* (Pocanschi et al. 2013), or a trimeric porin, MOMP, the major outer membrane protein from *Chlamydia trachomatis* (Fig. 7.2b; Tifrea et al. 2011).

In some cases, APols may affect the activity of the MPs they bind to. No or very moderate effects have been seen on the functional and pharmacological properties of the nicotinic acetylcholine receptor (Martinez et al. 2002), bacteriorhodopsin (Bazzacco et al. 2012; Dahmane et al. 2013; Gohon et al. 2008), or GPCRs (Banères et al. 2011; Bazzacco et al. 2012; Catoire et al. 2010a; Dahmane et al. 2009; Damian et al. 2012; Rahmeh et al. 2012). APol-trapped MPs can be recognized by antibodies (Charvolin et al. 2009; Giusti et al. 2014a; Le Bon et al. 2014b), by toxins (Charvolin et al. 2009), as well as by other soluble proteins (Basit et al. 2012). However,

presumably because of repulsive electrostatic interactions, A8-35 slows down the binding to GPCRs of G proteins and of arrestin, whereas NAPols do not (Bazzacco et al. 2012). The sarcoplasmic calcium pump (Ca^{2+} -ATPase) is reversibly inhibited after trapping in APols Champeil et al. 2000; Picard et al. 2006). At the same time, its half-life following calcium removal increases by $\sim 60\times$ over that in detergent solution (Fig. 7.2c). Our current working hypothesis is that both effects originate from the viscosity of the APol environment (Perlmutter et al. 2011), which would slow down the large conformational transitions of the transmembrane helix bundle that take place both during the functional cycle of the Ca^{2+} -ATPase and at the onset of denaturation (for discussions, see Picard et al. 2006; Popot et al. 2003, 2011). Other stabilization mechanisms also come into play, among which a less efficient competition of APols, as compared to detergents, with stabilizing protein/protein and protein/lipid interactions, as well as the reduction of the hydrophobic sink that is made possible by the high affinity of APols for MPs and their very low CAC (for a discussion, see Popot et al. 2011). The extent and origin of the stabilizing effect may vary from protein to protein, and different APols stabilize MPs to different extents. It seems, for instance, that the lesser the charge density along the APol chain, the better the stabilization (see e.g., Bazzacco et al. 2012; Picard et al. 2006).

7.1.5 Overview of APol Applications

APols were initially developed to handle MPs in aqueous solutions under less destabilizing conditions than can be achieved with detergents. This expectation has been largely vindicated (Popot 2010; Popot et al. 2011), despite a few exceptions (such as cytochrome b_6f , see Tribet et al. 1996). Their field of applications covers essentially all of those that are classically implemented in detergent solutions, with the benefit of improved stability. In addition, a few specific uses have been validated, based on the specific properties of MP/APol complexes and the rich chemistry of APols. Table 7.2 summarizes those applications that have been validated so far (for more details, see Popot 2010; Popot et al. 2011; Zoonens and Popot 2014). Two applications that remain problematical to date are infrared studies in the amide band absorption region—because all existing APols absorb in this region—and crystallization. Extensive tests have shown that a model MP, cytochrome bc_1 , does not crystallize when trapped in pure A8-35, whereas it does in a mixture of APols and detergent (Charvolin et al. 2014). Two phenomena can probably be incriminated, electrostatic repulsion between the complexes, which is reduced upon diluting A8-35 with a nonionic detergent, and the relative heterogeneity of MP/APol complexes, probably due to imperfect relaxation of the APol belt to its free energy minimum, which disappears in MP/APol/detergent ternary complexes (Zoonens et al. 2007). While not a good medium for crystallization, APols can be used to *produce*, by folding them from inclusion bodies or by cell-free expression (CFE), the MPs to be crystallized. APol-trapped MPs can then be transferred to a more favorable crystallization medium, such as lipid 3D phases (Polovinkin et al. 2014).

Table 7.2 Applications of amphipols. (Adapted from Zoonens and Popot 2014)

Applications	Implementability	References
Stabilization	+	Bazzacco et al. (2012); Champeil et al. (2000); Dahmane et al. (2011, 2013); Etzkorn et al. (2013); Feinstein et al. (2014); Gohon et al. (2008); Picard et al. (2006); Pocanschi et al. (2013); Popot et al. (2003); Tifrea et al. (2011) Tribet et al. (1996)
Functional studies	+/-	Bazzacco et al. (2012); Champeil et al. (2000); Charvolin et al. (2009); Dahmane et al. (2009, 2013); Gohon et al. (2008); Gorzelle et al. (2002); Martinez et al. (2002); Picard et al. (2006); Popot et al. (2003); Rahmeh et al. (2012)
Folding full-length MPs to their native state	+	Banères et al. (2011); Bazzacco et al. (2012); Catoire et al. (2010a); Dahmane et al. (2009, 2011, 2013); Etzkorn et al. (2013); Gohon et al. (2011); Leney et al. (2012); Pocanschi et al. (2006, 2013)
Cell-free expression	+	Bazzacco et al. (2012); Park et al. (2011)
Optical spectroscopy (UV-visible absorption spectrum, intrinsic MP fluorescence, UV CD)	+	Dahmane et al. (2013); Gohon et al. (2008); Pocanschi et al. (2006); Popot et al. (2011); Tifrea et al. (2011); Zoonens et al. (2007)
Infrared spectroscopy (peptide bond)	-	Popot et al. (2011)
MP solution studies by AUC, SEC, SANS, SAXS, affinity chromatography	+	Althoff et al. (2011); Bazzacco et al. (2009, 2012); Champeil et al. (2000); Charvolin et al. (2014); Dahmane et al. (2011, 2013); Diab et al. (2007a); Etzkorn et al. (2013); Gohon et al. (2008, 2011); Le Bon et al. (2013); Martinez et al. (2002); Picard et al. (2006); Prata et al. (2001); Sharma et al. (2012); Tribet et al. (1996, 1997); Zoonens et al. (2007)
Solution NMR	+	Bazzacco et al. (2012); Catoire et al. (2009, 2010a, b, 2011); Dahmane et al. (2011); Elter et al. 2014; Etzkorn et al. (2013); Planchard et al. (2014); Raschle et al. (2010); Zoonens et al. (2005)
Three-dimensional crystallization	-	Charvolin et al. (2014); Polovinkin et al. (2014); Popot et al. (2011)
Trapping MP supercomplexes	+	Althoff et al. (2011)
EM, STEM, AFM (single particles)	+	Althoff et al. (2011); Cao et al. (2013); Cvetkov et al. (2011); Flötenmeyer et al. (2007); Gohon et al. (2008); Liao et al. (2013); Tribet et al. (1998)
Transferring MPs to lipid bilayers (cubic phase, black film, cell)	+/-	Nagy et al. (2001); Pocanschi et al. (2006); Polovinkin et al. (2014);
Mediating MP immobilization for ligand-binding measurements	+	Basit et al. (2012); Charvolin et al. (2009); Fernandez et al. (2014); Giusti et al. (2013); Le Bon et al. (2013)

Table 7.2 (continued)

Applications	Implementability	References
Mass spectrometry	+	Bechara et al. (2012); Catoire et al. (2009); Leney et al. (2012); Ning et al. (2013, 2014)
Isoelectrofocusing and two-dimensional gels	+	unpublished data
Vaccination	+	Tifrea et al. (2011, 2014)

Signs in the second column indicate whether the method is easy to implement and general, or whether it may present difficulties or be protein dependent

7.1.6 Conclusion

In summary, APols are promising new surfactants for biochemical and biophysical studies of MPs in aqueous solution, because they form with MPs small and compact water-soluble complexes while improving the stability of the MPs they interact with. The APol layer surrounding the transmembrane domain of MPs is very stable, but it can be exchanged for other surfactants. The chemistry of APols allows modifications and labeling, generating a library of molecules with different molecular structures and more than a dozen types of labeled or functionalized versions of APol A8-35, expanding the scope of applications in both basic and applied research. Among the many novel surfactants developed to replace classical detergents, APols present the advantage of being remarkably easy to use. Their implementation is facilitated by the vast corpus of works describing in detail their properties and those of MP/APol complexes, as well as procedures for most conceivable applications, and by the commercial availability of APol A8-35.

In the next section, five protocols are presented in detail, describing: (1) how to trap MPs in APols so as to improve their stability, (2) how to quantify the amount of APol bound per MP, (3) how to fold an MP to its native state using APols, (4) how to produce MPs by CFE in presence of NAPols, and finally (5) how to immobilize MPs onto a solid surface for surface plasmon resonance (SPR) measurements.

7.2 Detailed Protocols for Some Major Applications

7.2.1 Protein Trapping in APols

APols are not (or very weak) detergents. As a consequence, they are poorly efficient for direct extraction of MPs from biological membranes, even if this has been observed in very few cases (Popot et al. 2003). Because of this feature, detergents must be used for solubilization, unless MPs are produced directly in the

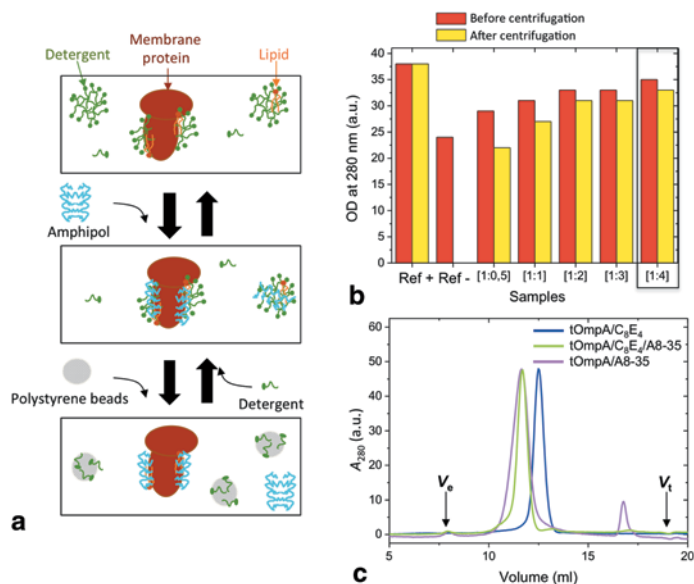


Fig. 7.3 Trapping membrane proteins in amphipols. **a** Schematic representation of MP trapping in APols. Adapted with permission from Proc Natl Acad Sci USA 102:8893–8898, copyright 2005 National Academy of Sciences, USA (Zoonens et al. 2005). **b** Determination of the optimal MP/APol mass ratio by measuring the optical density at 280 nm of samples of tOmpA trapped at different mass ratios of A8-35 before and after ultracentrifugation at $60,000 \times g$ for 20 min at 4°C (from Zoonens 2004). **c** SEC profiles of tOmpA in the detergent C_8E_4 , in a mixture of detergent and A8-35, and after trapping in A8-35. Reprinted with permission from Biochemistry 46:10392–10404, copyright 2007 American Chemical Society (Zoonens et al. 2007)

presence of APols (*cf.* § 7.2.4). The trapping procedure consists, then, of replacing detergents by APols in a sample of MPs, which, in general, has already been purified (Fig. 7.3). The protocol is simple, easy, fast, and requires no important biochemical optimization.

7.2.1.1 Preparation of a Stock Solution of APols

A8-35 is supplied as a white powder, which can be stored at room temperature. Note that most of APols are very stable molecules, except for NAPols and FAPols, which carry sugar groups and fluorescent probes, respectively. Storage of NAPols at -20°C is advisable whatever their conditioning, *i.e.*, in powder or in solution, because sugars can be hydrolyzed. FAPols need to be protected from UV-visible light with aluminum foil. When needed, a stock solution of APols at 100 g L^{-1} , or 10% w/w, is prepared with MilliQ water (water purified on an A10 Advantage Millipore system):

- Weigh some powder, for instance 20 mg, with an analytical balance in an Eppendorf tube or, if possible, in a small glass vial (note: The powder is sometimes very electrostatic and caution is required).
- Add 180 μL of MilliQ water in order to reach a final mass of 200 mg.
- Homogenize the solution with a vortex or by magnetic stirring. Incubate at least a couple of hours before use for a good rehydration of the lyophilized powder. The solution is then kept at 4 °C or, if need be, frozen at $-20\text{ }^{\circ}\text{C}$.

7.2.1.2 Determination of the Protein Concentration

The exchange of detergent for APols is carried out by supplying APols pre-solubilized in water to the sample of MPs (Fig. 7.3a). The amount of APol to be added is calculated on the basis of the mass of MP initially present in the sample. The concentration of protein must be known, at least approximately. After purification, the concentration of protein in the detergent solution is determined by its optical density from UV-visible spectra. If the epsilon of the protein is unknown, its concentration can be assessed by colorimetric measurements such as bicinchoinic acid (BCA) assay. Alternatively, amino acid analysis after HCl hydrolysis can also be used.

7.2.1.3 Determination of the Optimal MP/APol Mass Ratio

The sole optimization to do is that of the mass ratio of APols required to keep soluble the MP well dispersed in aqueous solution after detergent removal. For that, the protein and detergent concentrations are kept unchanged while increasing concentrations of APols are tested:

- Determine the MP/APol mass ratios to be tested. Table 7.3 gives an example, considering the concentration of MP to be 1 g L^{-1} and the volume of aliquots to be 500 μL for each condition.
- Pipet seven aliquots of equal volume of MP present in detergent solution. Add the appropriate volume of APols calculated. Note that dilution effects of the final volume can be neglected up to 10% of variation after adding APols. Keep aside the two control samples.
- Mix and incubate for 15–20 min at either room temperature or 4 °C depending on the stability of the protein of interest.

When APols are supplied to the samples, they mix freely with detergent molecules in micelles and at the transmembrane surface of the protein, as shown by fluorescence and isothermal calorimetry studies (Tribet et al. 2009; Zoonens et al. 2007). This leads to the formation of MP/detergent/APol ternary complexes.

Table 7.3 Example of conditions to be tested to determine the optimal MP/APol mass ratio

Mass of MP (mg)	MP/APol mass ratio	Mass of APol (mg)	Volume of APol (μL)
0.5	1:0	0	0
0.5	1:0.5	0.25	2.5
0.5	1:1	0.5	5
0.5	1:2	1	10
0.5	1:5	2.5	25
0.5	1:10	5	50

Note that the mass of protein to trap can be smaller or higher than 0.5 mg and the interval between ratios can be narrower. The sample at ratio 1:0 will be used for both positive and negative controls and, thus, should be prepared twice

7.2.1.4 Detergent Removal

This step can be achieved in various ways. Most often, detergent removal is carried out by adsorption into polystyrene beads (Bio-Beads). Note that APols do not absorb onto Bio-Beads (Zoonens 2004; Zoonens et al. 2007). The mass of beads to be added is typically $20\times$ the mass of detergent present in the sample. For instance, if the concentration of detergent is 6 g L^{-1} in $500\text{ }\mu\text{L}$, the amount of Bio-Beads to be added is $\sim 120\text{ mg}$:

- Calculate the appropriate amount of beads according to the mass of the detergent present in each sample. Weigh the beads and add them in the five samples containing APols plus in one of the two control samples, which will become the negative control (note: Bio-Beads are usually washed out successively in ethanol and water prior usage and then stored in water. Before weighing, dry them on a tissue paper for maximal water removal). The last sample, without beads, represents the positive control.
- Incubate the samples for 2 h under gentle shaking at either room temperature or $4\text{ }^{\circ}\text{C}$.
- Let the beads sediment by gravity or proceed to a quick centrifugation. Pipet the samples using a microcapillary tip that excludes the beads and put them in new Eppendorf tubes.

Alternatively, it is possible to eliminate the detergent micelles by dilution under the CMC of the detergent. Note that this method is more suitable to detergents with a high CMC rather than detergents with a low one, such as *n*-dodecyl- β -D-maltoside (DDM), because even under CMC these detergents are still able to keep MPs soluble. If the dilution method is employed, dilute the five samples containing APols plus that of the negative control with a detergent-free buffer. Dilute the last sample, which becomes the positive control, with buffer containing detergent at the same concentration as initially in the sample.

Whatever the protocol used, some detergent monomers can still be present in the samples. Usually, they are not problematic, as long as the negative control shows that the monomers cannot keep the MP in solution in the absence of APols, but if

need be they can be eliminated by dialysis or by several cycles of dilution/concentration using ultrafiltration devices. Note that the presence of APols in the external dialysis buffer is not required, as APols do not cross standard dialysis membranes of 12–14-kDa MW cutoff. Indeed, the MW of the particles of A8-35 is ~ 40 kDa (Gohon et al. 2006), and, because of its low CAC (0.002 g L^{-1} ; Giusti et al. 2012), there are very few free molecules in solution. Another procedure for detergent removal, albeit seldom used, is to adsorb it onto cyclodextrins (Althoff et al. 2011).

7.2.1.5 Identification of the Best MP/APol Ratio

- Measure the UV-visible spectrum of each sample.
- Centrifuge the samples at $100,000 \times g$ for 20 min (Note: The speed and duration of the centrifugation step are given for a small protein of about 30 kDa. These parameters can be adjusted if the protein of interest is bigger).
- Take off the supernatants and measure again their UV-visible spectra.
- Calculate the percentage of protein kept in the supernatant for each condition. This experiment determines the minimal MP/APol mass ratio required to keep MPs soluble (Fig. 7.3b). However, to establish the minimal MP/APol mass ratio required to obtain homogeneous complexes, which is slightly higher, it is recommended to analyze the samples by SEC (Fig. 7.3c).

The optimal MP/APol mass ratios for two model MPs of small MW like bacteriorhodopsin of *H. salinarium* (BR, 27 kDa) and the transmembrane domain of OmpA of *E. coli* (tOmpA, 19 kDa) are 1:5 and 1:4, respectively (Gohon et al. 2008; Zoonens et al. 2007). These ratios exceed by $\geq 2 \times$ the amount of A8-35 that binds to these MPs (see next protocol). This is because APols, which have a weak dissociating power, cannot prevent protein/protein interaction if they are not present in excess in the sample. To keep MP/APol complexes homogeneously distributed, an excess of APols is thus required. There is no need, however, to increase the concentration of APols beyond the minimal concentration yielding an acceptable monodispersity because, due to the hydrophobic sink effect, this may compromise the stability of fragile MPs (Popot et al. 2011).

7.2.2 Measuring of the Amount of Bound APols

APols specifically adsorb onto the transmembrane region of MPs (Althoff et al. Cao et al. 2013; 2011; Catoire et al. 2009; Liao et al. 2013; Zoonens et al. 2005), where they form a compact layer *ca.* 1.5–2 nm thick (Althoff et al. 2011; Gohon et al. 2008; Perlmutter et al. 2014). The amount of polymers constituting the belt surrounding MPs was estimated in two studies using BR and tOmpA as model MPs. The first determination is based on extensive physical measurements carried out on BR complexed with either plain or deuterated A8-35, using primarily SANS and AUC (Gohon et al. 2008). It is thought to give a relatively accurate measurement of

the amount of A8-35 bound per BR monomer, but is extremely laborintensive. The second study relied on the use of a fluorescent APol (FAPol). Under the conditions used in Zoonens et al. (2007), it yielded what is thought to be a lower limit to the amount of A8-35 bound to the tOmpA monomer (see below). We describe below first, how to express the amount of APol bound per MP, next, how to estimate a priori the amount of APol a given MP is likely to bind, and then three protocols for measuring it using FAPols.

7.2.2.1 Why is it Preferable to Express the Amount of APols Bound per MP in Mass Rather than as a Number of Molecules?

APols being highly polydisperse polymers, the size of individual molecule varies considerably and their MW can be estimated only on average. The average MW of A8-35 molecules is ~ 4.3 kDa (note: This revised mass is twice smaller than had been initially estimated; for a discussion, see Giusti et al. 2014b). However, despite the variable MW of individual APol chains, the particles they form in solution migrate upon SEC with a size distribution as narrow as that of globular proteins. SANS and AUC analyses indicate that they feature a well-defined Stokes radius ($R_s = 3.15$ nm) and MW (40 kDa; Gohon et al. 2006). Interestingly, the size and homogeneity of A8-35 particles do not depend on the degree of heterogeneity of the molecules that constitute them. Indeed, a version of A8-35 with a restricted length polydispersity forms particles with the same apparent size and dispersity as standard A8-35 particles (F. Giusti and C. Tribet, unpublished results). The average MW of individual molecules being only a very rough estimate, the amount of APol bound per MP is much more meaningfully expressed in mass ratio rather than as a molar stoichiometry. Similarly, in the case of functionalized APols, to preserve accuracy and reliability, the number of fluorophores or tags is expressed as their number per 40-kDa APol particle, which can be used as a mass reference, rather than as their number per APol chain, which has no great significance and is inaccurate.

7.2.2.2 How to Estimate a Priori the Likely Amount of APols Bound per MP Based on Structural Data?

In the case of α -helical MPs, the most thoroughly studied MP/APol complexes are those of BR with A8-35. In the complexes, the protein/APol mass ratio is $\sim 1:2$, i.e., ~ 54 kDa of A8-35 per monomer of BR (27 kDa; Gohon et al. 2008). Lipids (~ 9 kDa) are also present in the complexes. In the case of β -barrel MPs, the best-characterized complexes, in terms of composition, are those of tOmpA with A8-35. The mass ratio that has been estimated is $\sim 1:1.3$, i.e., ~ 25 kDa of A8-35 per monomer of tOmpA (19 kDa; Zoonens et al. 2007). This value should be considered as a minimal value, however, because the conditions under which the measurements were done (see below) led to some aggregation and, very likely, to the loss of some APol. A more likely assumption of bound APol per tOmpA monomer would be

~45 kDa, which corresponds to the amount of APol needed, in MD calculations, to fully cover the transmembrane domain of OmpX, an eight-stranded β -barrel whose dimensions are very similar to those of tOmpA (Perlmutter et al. 2014).

Based on these values, and assuming that the volume of the A8-35 belt surrounding an MP is roughly proportional to the perimeter of the transmembrane domain to be covered, it is possible to estimate the amount of APols interacting with any other MP. The only information needed is the dimensions of the hydrophobic domain of the protein of interest, modeled as a cylinder filled out by the transmembrane helices or delimited by the number of β -strands, whose perimeter increases roughly as the square root of its volume (α -helical bundles), or linearly with the number of β -strands. For instance, the transmembrane domain of BR is a bundle of seven α -helices. If the MP of interest has a similar organization, like, for example, a GPCR, it can be expected to bind approximately the same amount of A8-35, i.e., ~54 kDa (note, however, that BR/A8-35 complexes comprise ~9 kDa of lipids (Gohon et al. 2008), which increase the transmembrane perimeter). On the other hand, if the MP contains twice more helices than BR and features a more or less homothetic shape, the volume of its transmembrane domain doubles while the transmembrane surface increases by ~40%, and one can expect in the ballpark of 75 kDa of bound A8-35. It is fair to say, however, that too few reliable measurements are available to date to gather how reliable such an approach is. The only other relatively precise estimate of bound A8-35 has been obtained with the cytochrome bc_1 dimer, which has 22 transmembrane helices. On the basis of the above calculations, one would expect it to bind ~96 kDa A8-35. The experimental estimate is only 49–63 kDa (Charvolin et al. 2014; Popot et al. 2003). Note also that it is not unreasonable to expect that the ionic strength may modulate the volume and mass of the belts of ionic APols because it modulates the repulsion between charged polar groups (see Popot et al. 2003). Despite these uncertainties, estimating a priori the probable mass ratio of MP to APol in complexes is useful to provide guidelines when planning trapping experiments, or when endeavoring to measure experimentally the amount of bound APol.

7.2.2.3 How to Experimentally Measure the Quantity of APols Bound per MP?

As previously mentioned, the mass of APol to add for trapping is in excess of that of APol that actually binds to the surface of the MP. After trapping, some APol remains present as free particles in the sample. Measuring the amount of bound APols can be carried out by several approaches. Whatever the method chosen, the analysis is considerably facilitated by using FAPols as tracers:

- Prepare a stock solution of APol/FAPol mixture.

FAPols carrying various fluorescent probes have been synthesized (see Table 7.1). The choice of FAPol depends on the absorption spectrum of the protein of interest. For instance, if the protein absorbs only at 280 nm, FAPol_{NBD}, which shows a maximum absorption at 490 nm, is suitable. On the other hand, if the protein absorbs also

visible light, as BR does, another FAPol, like FAPol_{Alexa647} must be chosen in order to avoid overlapping between protein and FAPol absorption bands.

Pure FAPols usually absorb too much at the peak of absorbance of fluorophore, and possibly also at 280 nm, interfering with protein determination. They are better used diluted with nonfluorescent APol (A8-35). Because labeled and unlabeled APols freely and rapidly mix in salty aqueous solutions (Zoonens et al. 2007), they can be mixed from two stock solutions prepared at 100 g L⁻¹. A convenient FAPol/A8-35 ratio is one at which the absorbance of the FAPol at its maximal absorption wavelength is, in the complexes with the MP, ~25% of that of the protein at 280 nm. This ratio can be estimated a priori based on the extinction coefficients of the protein and FAPol, the estimation of the amount of APols bound per MP (*cf.* § 7.2.2.2), and on the MP/APol mass ratio needed for trapping. If the protein possesses many tryptophan residues, its extinction coefficient may be high enough so that no dilution of the FAPol stock solution is necessary:

- Measure the spectral absorbance of the FAPol/A8-35 mixture (or pure FAPol if dilution with A8-35 is not necessary) and determine the relative contribution of APols at 280 nm and at the peak of absorbance (note: Even if neither APol nor FAPol absorb significantly at 280 nm, it is advisable to quantify it).

After the complexes with an MP have been formed, they must be separated from the excess of APol used at the trapping step.

Method 1: Size-exclusion chromatography (SEC). This approach is appropriate for large MPs—40 kDa or above—because MP/APol complexes and free APol particles are sufficiently resolved:

- Wash the gel filtration column with three column volumes of running buffer. Note that, for analytic analysis of the samples, APol is usually not required in the running buffer, in contrast to detergents, which must always be present above their CMC. However, for this particular experiment, it cannot be excluded that a small amount of APol leaches from the protein as the complexes migrate into APol-free buffer. It might therefore be preferable to saturate the solution with which the column is equilibrated with “some” free APol, such as the CAC or twice the CAC, so as to prevent desorption, and to subtract the corresponding background. However, this modification to the procedure has not been carefully investigated yet. The composition of standard buffer is 20 mM Tris-HCl, 150 mM NaCl, and pH 8.0 but it can be modified provided that pH is above 7 and divalent cations are absent.
- Inject an aliquot of FAPol/A8-35 mixture (or pure FAPol) at 10 g L⁻¹. The elution profile is monitored at two wavelengths, 280 nm and the maximum absorption wavelength of the fluorophore, for example, 490 nm for FAPol_{NBD}. Determine the elution volume of APol particles.
- After trapping the protein in the FAPol/A8-35 mixture, inject an aliquot of the sample at an appropriate concentration in order to get a good signal-to-noise ratio of the elution peak. Follow the elution of MP/A8-35/FAPol_{NBD} complexes at the two wavelengths 280 and 490 nm. If the separation from free APol par-

ticles is good, calculate the amount of bound APols per MP as follows: Integrate the peak area of MP/A8-35/FAPol_{NBD} complexes at 280 and 490 nm in order to determine, respectively, the mass of MP and that of FAPol_{NBD} which had co-migrated with the protein. A subtraction of the APol contribution to the absorption of the protein at 280 nm may have to be applied, based on the ratio of the surface of the peaks at 280 and 490 nm observed with the pure FAPol/A8-35 mixture or on the ratio of the optical densities at 280 and 490 nm measured from a UV-visible spectrum. The total mass of APol is then calculated taking into account the dilution of FAPol with A8-35. The ratio of APol and MP masses gives the amount of bound APols per MP (note: If the elution peaks of APol particles and MP/APol complexes overlap, use a more resolutive gel filtration column or try another separation procedure).

Method 2: Immobilized metal affinity chromatography (IMAC). The presence of a tag fused to the MP under study makes it possible to immobilize MP/APol complexes onto an affinity column and to eliminate the excess of APol particles. This procedure is particularly convenient when the protein is small and MP/APol complexes cannot be efficiently separated from free APol particles by SEC. Note that free APol particles are, however, required to keep homogeneous MP/APol complexes. Indeed, as previously mentioned, APols are poorly dissociating surfactants, and so, in the absence of free APol particles, small MP/APol oligomers start to form, which is likely to be accompanied by some desorption of the MP-bound APol (Zoonens et al. 2007). Because of this effect, the MP/APol ratio determined by this method must be taken as an estimate by default unless buffers are saturated with some free APol:

- After trapping the MP in the FAPol/A8-35 mixture, inject the sample on an affinity resin. For instance, if the protein has a polyhistidine tag, load the sample on a Ni:NTA column. The majority of the protein (~80%) will be retained on the resin (Giusti et al. 2014a).
- Rinse the resin with equilibration buffer to wash out free FAPol/A8-35 particles. Elute the MP/FAPol/A8-35 complexes with a buffer containing imidazole. Note that, as said before, the presence of APol in equilibration and elution buffers at the CAC or twice the CAC may be advisable, so as to prevent desorption.
- Desalt the sample to remove imidazole and measure the optical density of the sample at 280 nm and at the maximum wavelength of FAPol. The concentrations of MP and FAPol are calculated using their respective extinction coefficients. Subtract, if need be, the contribution of FAPol at 280 nm and calculate the total mass of APol if FAPol was initially mixed with A8-35 before trapping. The ratio of APol and MP masses gives the amount of bound APol per MP.

Method 3: Analytical ultracentrifugation (AUC). The MP/APol mass ratio in complexes can be precisely determined by sedimentation velocity (SV) measurements using AUC. AUC is a priori applicable to any MP, because the density of APol particles and that of MP/APol complexes are different enough for them to separate during the centrifugation run, even if their hydrodynamic radii are not very different. For example, the sedimentation coefficients (S) of A8-35 particles ($R_s \approx 3.15$ nm; Gohon et al.

2006) and BR/A8-35 complexes ($R_g \approx 5.0$ nm; Gohon et al. 2008) are 1.6 and 3.2 S , respectively, making them easily distinguishable (Gohon et al. 2006, 2008). The specific volume of the sodium salt of A8-35, \bar{v}_2 , is 0.809 L g^{-1} , its density, $\rho = 1/\bar{v}_2$, 1.236 L g^{-1} (Gohon et al. 2004, 2006). The MP/APol mass ratio can be determined by sophisticated AUC measurements involving the comparison of sedimentation properties of complexes formed between the protein and unlabeled or deuterated A8-35 and/or simultaneous measurements of the absorbance and refractive index of the complexes (Gohon et al. 2008). However, with the advent of FAPols, it is simpler to measure the respective absorbance of the protein and the APol in the complexes, as done above for the complexes separated by SEC or affinity chromatography:

- After MP trapping in FAPol/A8-35 mixture, adjust the sample concentration by dilution or concentration so that the protein absorbance at 280 nm, in the AUC cell, reaches ~ 0.5 .
- Define the parameters of the SV run, namely time and speed, according to the sedimentation coefficient of the protein under study. For instance, in the case of small MPs, like BR or tOmpA, the SV experiment is carried out at 42,000 rpm during 4 h. The migration of the particles and complexes is followed at two wavelengths, 280 nm and the maximum absorption wavelength of FAPol, and, if available, with interference optics, which gives a measure of the refractive index.
- Measure the solvent density and viscosity.
- Analyze the SV profiles with Sedfit or an equivalent program (for details, see Gohon et al. 2008). The distribution $c(S)$ of sedimentation coefficients (S) shows peaks reflecting the migration of MP/APol complexes and of free APol particles during the SV run. Integrate the peak areas at 280 nm and at the second wavelength. As noted above for SEC and IMAC experiments, the contribution of APols at 280 nm may have to be subtracted from the protein signal. MP and FAPol masses can be determined from their respective extinction coefficients. Calculate the total mass of APol from the FAPol/A8-35 ratio. The ratio of APol and MP masses gives the amount of bound APol per MP.

Whereas this procedure is technologically more demanding, it presents the advantage, over that by SEC and by IMAC, that the MP/APol complexes are never separated from the free APol, which eliminates the risks of desorption and/or aggregation.

7.2.3 APols-Assisted Folding of a MP

APols have proven to be very helpful to fold MPs expressed as inclusion bodies in *E. coli*, like class A GPCRs or porins (Banères et al. 2011; Bazzacco et al. 2012; Dahmane et al. 2009, 2011, 2013; Pocanschi et al. 2006; see also Chap. 3 by JL Banères in this volume; Fig. 7.4). The protocol used for α -helical MPs is derived from one initially developed to refold BR in lipids (see Popot et al. 1987, in which many useful practical

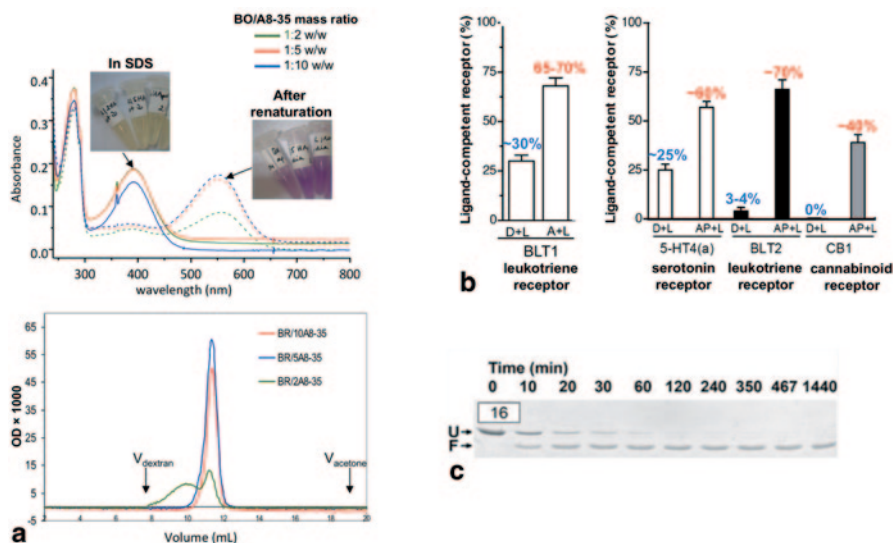


Fig. 7.4 Amphipol-assisted folding of membrane proteins to their native state. **a** Folding of BR from a denaturated state in SDS. *Top*, UV-visible spectra of samples before and after precipitation of SDS with KCl in the presence of A8-35 at various MP/APol mass ratios. In SDS, the sample is colored yellow because BR is denaturated in bacterio-opsin (BO) and the retinal is released in solution. After refolding, the color shifts to purple because the retinal has rebound to the protein, indicating that BO has recovered its native structure. *Bottom*, SEC profiles of the same three samples after folding. Reprinted with permission from Biochemistry 45:13954–13961, copyright 2006 American Chemical Society (Pocanschi et al. 2006). **b** Comparison of the yield of folding of four GPCRs in detergent + lipids (D+L) and in A8-35 + lipids (A+L). Adapted with permission from Biochemistry 48:6516–6521, copyright 2009 American Chemical Society (Dahmane et al. 2009). **c** Folding of a porin, full-length OmpA from *E. coli*, after 19× dilution from 8 M urea into an A8-35 solution. The unfolded (U) and folded (F) forms of OmpA were separated by SDS-PAGE. Reprinted from Eur Biophys J 42:103–118, copyright 2013, with kind permission from Springer Science and Business Media (Pocanschi et al. 2013)

details can be found). It is quite simple but requires some optimization regarding the quantity of APol to be added and the presence or absence of lipids. For variants and the effect of various modifications to this protocol, see Dahmane et al. 2013.

7.2.3.1 Solubilization and Purification of MPs in Denaturing Conditions

Inclusion bodies are clusters principally comprised of misfolded forms of the protein of interest, but they can also contain some DNA and other bacterial proteins. They need to be solubilized and purified in denaturing conditions. For α -helical MPs, the denaturing agent is usually SDS, whereas for β -barrel MPs it is urea. Purification is most often carried out by affinity chromatography. For instance, if the protein of interest is fused to a polyhistidine tag, purification can be carried out on Ni:NTA resin. Note that the concentration of SDS and urea tolerated by the resin is

given by the supplier. Also, it is essential to work at room temperature, because both urea and SDS crystallize at low temperature:

- Isolate the inclusion bodies by differential centrifugations and determine the concentration of protein by the BCA assay.
- Prepare the solubilization buffer containing the appropriate denaturing agent. For example, for α -helical MPs, the buffer contains 10 mM Tris-HCl, pH 7.5, 100 mM NaH₂PO₄, 6 M urea, 0.8% SDS, 10% glycerol, 4 mM β -mercaptoethanol (adapted from protocols described in Banères et al. 2005; Damian et al. 2006). For porins, the solubilization buffer contains 10 mM borate, pH 10.0, 8 M urea, 2 mM EDTA (Pocanschi et al. 2006, 2013; note: The presence of reducing agent is required only if cystein residues are present. The solubility of urea can be increased to 10 M by heating).
- Dissolve the inclusion bodies in the appropriate solubilization buffer at a final concentration of 10 g L⁻¹ and incubate overnight at room temperature (note: Sonication pulses can be applied to speed up solubilization).
- Centrifuge the sample for 20 min at 20,000 $\times g$ in order to remove insoluble material.
- Proceed to the purification step. Note that for a purification of α -helical MPs on Ni:NTA resin, the buffers are: (1) equilibration buffer: 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 0.8% SDS, 4 mM β -mercaptoethanol; (2) elution buffer: 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 0.8% SDS, 400 mM imidazole, 4 mM β -mercaptoethanol; and (3) desalting buffer: 50 mM Tris-HCl, pH 8.0, 0.8% SDS, 4 mM β -mercaptoethanol. For purifying β -barrel MPs, SDS in each buffer is replaced by 8 M urea.
- Determine the concentration of protein by UV-absorbance or by the BCA assay.

7.2.3.2 Renaturation in APols

This step consists of either exchanging SDS for APols (protocol 1) or diluting urea in the presence of APols (protocol 2). The optimal MP/APol mass ratio must be determined by carrying out folding tests with variable amounts of APols.

Protocol 1 (for α -helical MPs):

- Distribute 0.25 mg of the MP to be folded in three Eppendorf tubes. Add increasing volumes of APol—5, 12.5, 25 μ L—from a stock solution at 100 g L⁻¹ in order to obtain MP/APol mass ratios equal to 1:2, 1:5, and 1:10. Note that lipids generally help in the folding process, cf. (Dahmane et al. 2009, 2013). To test it, they can be supplied to the samples so that the APol/lipid mass ratio is 1:0.2, but this ratio can be optimized, as well as the nature of the lipids.
- Mix and incubate the samples for 30 min at room temperature.
- SDS is eliminated by precipitation with KCl added from a 4-M stock solution so that the final concentration of KCl in the samples is equal to 150 mM plus the concentration of SDS. For example, if the volume of the sample is 500 μ L and the concentration of SDS is 0.8% (28 mM), the final KCl concentration should be 178 mM. The volume of KCl to be added is thus 44.5 μ L.

- Incubate 30 min at room temperature under vigorous stirring.
- Centrifuge the samples 5 min at the maximum speed of a benchtop centrifuge at 20 °C.
- Collect the supernatant and repeat the centrifugation step.
- Measure the optical density of samples.

Protocol 2 (for β -barrel MPs):

- Fix, e.g., to 1 mg the mass of protein to fold. Dilute the sample by a 10 \times dilution factor into urea-free buffer containing 5 mg of APols, so that the final MP/APol mass ratio is 1:5. Test also ratios 1:2 and 1:10. Note that the dilution factor and speed of dilution can be optimized. If need be, incubate the samples at 40 °C for 24 hours
- Concentrate the samples using an ultrafiltration device and measure the optical density of supernatants.

7.2.3.3 Further Renaturation

To increase the yield of folding, urea or SDS traces can be further eliminated by a dialysis step:

- Dialyze the sample for 24 h at room temperature using a standard dialysis membrane of 12–14-kDa MW cutoff. Note that APols are not needed in the external bath, but the presence of 150 mM KCl is required to prevent redissolution of crystallites of potassium dodecyl sulfate that may not have been totally removed by centrifugation. The volume of the external bath is $\sim 500\times$ larger than the volume of the samples.
- Recover the sample and centrifuge it for 5 min at the maximum speed of a benchtop centrifuge.
- Measure the optical density of the samples.
- If the buffer needs to be exchanged, proceed to a second dialysis for 24 h at 4 °C.

The solubility of MP is not a criterion of folding. The simplest and most direct proof that the protein adopts its native conformation is to check its activity. If the activity assay is not easy to set up, the yield of folding can be assessed by other approaches such as ligand-binding experiments using equilibrium dialysis. In that case, ligand titration can be monitored by radioactivity or changes in the intensity of emission fluorescence or absorption. If the protein is naturally colored in its native conformation due to the binding of a cofactor, such as retinal for BR, the native state can be quantified by spectral absorption changes. It is also possible to check the homogeneity and size of the protein by SEC, its secondary structure by CD, the local environment of tryptophan residues by CD and fluorescence measurements, and the melting temperature by differential scanning calorimetry or by fluorescence thermal shift. In the case of porins, the folded state of the protein can usually be assessed by SDS-PAGE, upon which, as a rule, folded and unfolded forms exhibit different electrophoretic mobilities, or by dot blot if an antibody recognizing the native state of the protein is available, or by protease digestion, etc.

7.2.4 APols-Assisted CFE of a MP

Overexpression of MPs in vivo under their native form is extremely tricky because, unlike soluble proteins, these proteins, to be functional, have to be targeted and inserted into the membrane. Because this insertion process can be inefficient and is frequently toxic, it represents a major limiting factor for protein production. One strategy to overcome this problem is to resort to CFE (Zubay 1973; see also Chap. 2 by F. Bernhard in this volume). CFE presents many attractive features, among which to do away with the toxicity issue and to allow labeling using limited amounts of isotopically labeled or unnatural amino acids (Kigawa et al. 1999). Unfortunately, probably due to their anionic character, charged APols such as A8-35 or SAPols inhibit the CFE of MPs (Park et al. 2011). Such is not the case, however, for nonionic APols (NAPols; Bazzacco et al. 2012; Fig. 7.5). Excellent results have been obtained when BR was expressed in vitro in the presence of NAPols: The majority of BR (~90%) was properly folded and remained stable over several months, whereas in DDM, in spite of similar production and folding yields, it tends to precipitate rapidly (Bazzacco et al. 2012; Park et al. 2011). Because APols are so much milder than detergents, developing APol-assisted CFE of MPs appears as a very promising approach.

CFE is carried out using a commercial system (5prime) or a homemade lysate (see Chap. 2 by F. Bernhard, in this volume), in which an *E. coli* lysate provides the machinery to drive coupled transcription and translation in the presence of a DNA template (note: Prokaryotic lysate from *E. coli* is commonly employed to produce both pro- and eukaryotic MPs. The protocol below was developed using this lysate, but it is certainly possible to use eukaryotic lysates such as that from wheat germ). For large-scale production, a semipermeable membrane allows for a continuous supply of substrates and the removal of inhibitory by-products, thus extending the duration of expression and the protein synthesis yield.

7.2.4.1 CFE Small-Scale Reaction

To optimize the concentration of NAPols to be used, small-scale syntheses are carried out in the presence of a concentration range of NAPols:

- Prepare a stock solution of NAPols at 100 g L⁻¹, or 10% w/w, in MilliQ water as previously described in protocol 7.2.1.1.
- Small-scale syntheses are carried out in the presence of 0.5 µg of plasmid and 3, 5, 8, and 10 g L⁻¹ of NAPols in 50 µL of lysate. Note that the quantity of NAPol tends to be larger than is actually necessary. Indeed, because the amount of synthesized proteins is hardly predictable, it is better to add an excess of it. The reaction is performed for 6 h in a thermomixer (Eppendorf) at 700 rpm and 25 °C (note: It is possible to optimize also the temperature, Mg²⁺

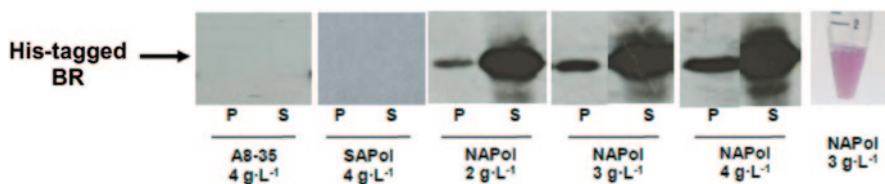


Fig. 7.5 Cell-free expression of bacteriorhodopsin in a nonionic amphipol SDS-PAGE followed by Western blot using an anti-His-tag antibody. The purple color of the sample indicates that BR has bound retinal, a proof that it has reached its native structure. (From Popot et al. 2011; see Bazzacco et al. 2012)

concentration, and other parameters, as exhaustively described in the same volume by Bernhard et al.).

- At the end of the small-scale tests, 10 μL samples are diluted with the same volume of 10 mM Tris-HCl, pH 8.0 buffer, and then centrifuged for 10 min at $16,000 \times g$ at room temperature to check for protein solubility by Western blot.

7.2.4.2 CFE Large-Scale Reaction

Once the concentration of NAPols to be used has been optimized, the reaction can be scaled up to 1 mL of lysate:

- A new solution of NAPols is prepared directly in the CFE reaction buffer used to dissolve all the components of the lysate. This allows reaching the desired final volume (1 mL) in the presence of the adequate quantity of NAPol. The solution of NAPols is stirred overnight to insure complete solubilization. Note that it is recommended to add 1% of sodium azide in the solution.
- 15 μg of plasmid, in CFE reaction buffer, is incubated for 20 h in the thermomixer in the presence of NAPols at the optimized temperature (note: It is not necessary to add NAPols in the feeding chamber).

7.2.4.3 Sample Collection

Plasmids for CFE (pIVEX) usually contain a hexa-His fusion tag to allow detection and purification of the expressed protein. After the lysate has been collected, the protein can be purified by IMAC, either in batch or using a pre-packed column. As previously mentioned, it is not necessary to add NAPols in the purification buffers.

7.2.5 APols-Mediated Immobilization of MPs

The development of sensors carrying MPs as recognition motifs has multiple applications in diagnostics and/or fundamental understanding of molecular interactions

necessary for the design of new drug-discovery strategies. The development of such biosensors requires the isolation and immobilization of MPs onto solid supports without alteration of their native conformation and function and the use of appropriate bionanotechnological platforms for high-sensitivity detection.

Surface plasmon resonance (SPR; see Rich and Myszka 2005) is a label-free technique based on the propagation of an evanescent wave along a gold-coated surface. It permits the detection of changes of refractive index in the vicinity of the surface that result from the binding of partners. This technique requires the immobilization of one of the interacting partners onto the surface. Only a few studies have been published on isolated MPs due to the difficulty to maintain them functional in solution, and because direct interaction of the proteins with the surface tends to impair their function. Specific immobilization strategies are usually a good alternative (Jonkheijm et al. 2008). Various affinity tags can be used for protein immobilization, including polyhistidine tags, which bind to Ni:NTA motifs, affinity tags recognized by antibodies, and biotin or streptags binding to streptavidin. Functionalized APols fulfill several criteria for a specific immobilization of functional MPs onto surfaces: (1) They stabilize MPs, providing a longer lifetime to the isolated protein; (2) they form with MPs permanent complexes, which do not dissociate in the absence of free APols; biosensors can thus be used in surfactant-free buffers, which simplifies their implementation and considerably limits the risk of missing weak or moderate interactions with ligands; and, finally, (3) the complexes formed between MPs and functionalized APols will bind onto the solid surface of the biosensor without the need for any genetic or chemical modification of the protein (Fig. 7.6; Charvolin et al. 2009).

The protocol described here is for MPs trapped in biotinylated A8-35 (BAPol), but it can be adapted to MPs trapped in other functionalized APols, such as HistAPol, ImidAPol, or OligAPol (*cf.* § 7.1.3). If the MP possesses a tag, its trapping with unfunctionalized APol, e.g., plain A8-35, will improve its stability and the immobilization of the complexes can be carried out following the same approach. This protocol is adapted for Biacore instruments (GE Healthcare), which are the most widely distributed commercial instruments. Specific conditions are indicated in the case of a Biacore X100 instrument. The signal can be optimized by adjusting experimental conditions (Karlsson and Fält 1997). For more details about data analysis, see Rich and Myszka 2005.

7.2.5.1 Experimental Setup

MP/BAPols complexes (called ligands in the Biacore terminology) are immobilized to the Biacore sensor chip before addition of the interacting molecules, for instance, antibodies (called analyte in the Biacore terminology), to reveal specific interactions.

Choose the adequate sensor chips. The binding capacity of most of Biacore chips is improved by pre-coupling of a dextran matrix. Existing CM5 chips, which are the most widely used, are already premodified for coupling of various proteins. If the immobilization is carried out via BAPols, the SA sensor chip (CM5 chips premodified with streptavidin) is used. Note that alternatives of SA sensor chips are the use of a standard CM5 chip followed by aminocoupling of streptavidin, or the

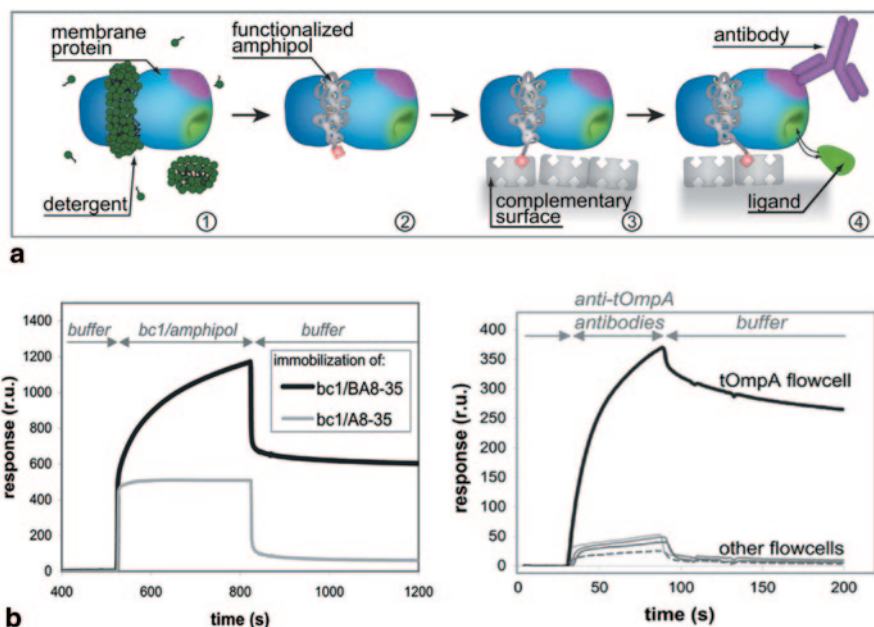


Fig. 7.6 Amphipol-mediated immobilization of membrane proteins onto solid supports **a** Principle of MPs immobilization using functionalized APols. **b** Measurement of the immobilization of an MP (*left*) and of the interaction of another with a specific antibody (*right*) by SPR. *Left*, application to streptavidin-coated chip lanes of cytochrome bc_1 complexed either by plain A8-35 (*gray*) or by biotinylated A8-35 (BAPol; *black*). *Right*, application of an antibody raised against tOmpA to lanes carrying BAPol-immobilized tOmpA (*black*) or other MPs (*other traces*). Reprinted with permission from Proc Natl Acad Sci USA 106:405–410, copyright 2009 National Academy of Sciences, USA. (Charvolin et al. 2009)

SA capture chip, which is regenerable. In the case of charge repulsion between the sample to be immobilized and the surface, it is possible to improve the detection signal by using commercial chips with shorter dextran matrices. Note that two flow cells are required for the measurements: The sample flow cell, where the total signal will result from specific and nonspecific binding onto surfaces, and the reference flow cell, used to evaluate the nonspecific binding of proteins onto surfaces.

7.2.5.2 Preparation of Samples

Note that it is important to exclude compounds influencing the refractive index of the solutions (e.g., glycerol). Moreover, all samples and buffers should be filtered on 0.22- μm filters or centrifuged 10 min at $10,000 \times g$:

- Trap the protein in BAPol as described in protocol 7.2.1. The final protein concentration to fix for a good signal depends on the protein under study. For example, in the case of BR, the protein concentration should be 0.03 g L^{-1} . Typical quantities required for the experiments are 0.06–0.1 μg of total protein.

- Prepare the solutions of antibodies (a nonspecific antibody and a specific one for negative and positive controls, respectively), both solutions at two different concentrations (0.055 and 0.01 g L^{-1}). Typical quantities required for the experiments are $6\text{--}10 \text{ }\mu\text{g}$ of total protein. Note that these quantities are indicative and should be optimized depending on sample purity.
- Prepare the running buffer, HBS-N, which contains 10 mM HEPES, $\text{pH } 7.4$, and 150 mM NaCl. Note that the dextran matrix and the APols are both negatively charged and so, the ionic strength of buffer can be optimized for an optimal signal with minimal repulsion.
- Prepare the surface regeneration solution (1 M NaCl, 50 mM NaOH).

7.2.5.3 Measurements

The signal observed during an SPR experiment is proportional to the refractive index, which is proportional to the mass for molecules with the same refractive index. Proteins have approximately the same refractive index and the empirical conversion factor ($1 \text{ RU} \approx 1 \text{ pg mm}^{-2}$) can be used for estimating the mass bound to the surface (Stenberg et al. 1991):

- Follow the instructions of the manufacturer for conditioning the sensor surface.
- After obtaining a stable baseline, immobilize the proteins onto the sensor surface, which becomes the sample flow cell. For this step, start a new cycle using only the second flow cell for its functionalization and inject the sample. In the case of qualitative analysis, it is recommended to try to saturate the surface with the MP/BAPol complexes for the highest possible response from antibody binding in the next step. This can be achieved by repeating consecutive injections until stabilization of the signal (note: For a flow rate of $10 \text{ }\mu\text{L min}^{-1}$, use consecutive 60-s injections of MP/BAPol complexes at 0.03 g L^{-1} to saturate the surface). On the other hand, in the case of kinetic studies, it is usually necessary to limit the density of immobilized ligand to avoid diffusion-limited data, analyte depletion, and analyte rebinding.
- Prepare the reference flow cell (the first flow cell), which can either be left unmodified, or be modified with either BAPol alone or a different MP trapped in BAPol.
- Stop the manual run after obtaining a stable signal in both cells.
- Measure the binding of the analyte, for instance, that of an antibody raised against the MP under study. It is recommended to use two different concentrations of antibodies (0.05 g L^{-1} for the first injection and 0.01 g L^{-1} for the second one) and one regeneration solution in each cycle. The solutions will be added according to the series of the following cycles:
 - Cycle 1: Blank cycle with running buffer
 - Cycle 2: Cycle with a nonspecific antibody to be used for negative control
 - Cycle 3: Cycle with a specific antibody raised against the MP under study

The sample parameters are: 30 $\mu\text{L min}^{-1}$ flow rate, 180 s contact time, and 120 s dissociation time. The regeneration parameters are: 30 $\mu\text{L min}^{-1}$ flow rate, 60 s contact time, and wash in running buffer.

Acknowledgments We would like to thank L. J. Catoire for reading the protocol of APol-assisted MP folding. The development of amphipols has been mainly supported by the Centre National de la Recherche Scientifique, the Human Frontier Science Program Organization (RG00223/2000-M), and the European Community (BIO4-CT98-0269 and STREP LSHG-CT-2005-513770 *Innovative Tools for Membrane Protein Structural Proteomics*).

References

- Althoff T, Mills DJ, Popot J-L, Kühlbrandt W (2011) Assembly of electron transport chain components in bovine mitochondrial supercomplex I₁III₂IV₁. *EMBO J* 30:4652–4664
- Banères J-L, Mesnier D, Martin A, Joubert L, Dumuis A, Bockaert J (2005) Molecular characterization of a purified 5-HT₄ receptor. A structural basis for drug efficacy. *J Biol Chem* 280:20253–20260
- Banères J-L, Popot J-L, Mouillac B (2011) New advances in production and functional folding of G protein-coupled receptors. *Trends Biotechnol* 29:314–322
- Basit H, Sharma S, Van der Heyden A, Gondran C, Breyton C, Dumy P, Winnik FM, Labbé P (2012) Amphipol mediated surface immobilization of FhuA: a platform for label-free detection of the bacteriophage protein pb5. *Chem Commun* 48:6037–6039
- Bazzacco P, Sharma KS, Durand G, Giusti F, Ebel C, Popot J-L, Pucci B (2009) Trapping and stabilization of integral membrane proteins by hydrophobically grafted glucose-based telomers. *Biomacromolecules* 10:3317–3326
- Bazzacco P, Billon-Denis E, Sharma KS, Catoire LJ, Mary S, Le Bon C, Point E, Banères J-L, Durand G, Zito F, Pucci B, Popot J-L (2012) Non-ionic homopolymeric amphipols: application to membrane protein folding, cell-free synthesis, and solution NMR. *Biochemistry* 51:1416–1430
- Bechara C, Bolbach G, Bazzacco P, Sharma SK, Durand G, Popot J-L, Zito F, Sagan S (2012) MALDI mass spectrometry analysis of membrane protein/amphipol complexes. *Anal Chem* 84:6128–6135
- Bowie JU (2001) Stabilizing membrane proteins. *Curr Opin Struct Biol* 11:397–402
- Breyton C, Chabaud E, Chaudier Y, Pucci B, Popot J-L (2004) Hemifluorinated surfactants: a non-dissociating environment for handling membrane proteins in aqueous solutions? *FEBS Lett* 564:312–318
- Breyton C, Gabel F, Abla M, Pierre Y, Lebaupain F, Durand G, Popot J-L, Ebel C, Pucci B (2009) Micellar and biochemical properties of (hemi)fluorinated surfactants are controlled by the size of the polar head. *Biophys J* 97:1077–1086
- Breyton C, Pucci B, Popot J-L (2010) Amphipols and fluorinated surfactants: two alternatives to detergents for studying membrane proteins *in vitro*. In: Mus-Veteau I (ed) *Heterologous expression of membrane proteins: methods and protocols*, vol 601. The Humana Press, Totowa, pp 219–245
- Cao E, Liao M, Cheng Y, Julius D (2013) TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* 504:113–118
- Catoire LJ, Zoonens M, van Heijenoort C, Giusti F, Popot J-L, Guittet E (2009) Inter- and intramolecular contacts in a membrane protein/surfactant complex observed by heteronuclear dipole-to-dipole cross-relaxation. *J Magn Res* 197:91–95
- Catoire LJ, Damian M, Giusti F, Martin A, van Heijenoort C, Popot J-L, Guittet E, Banères J-L (2010a) Structure of a GPCR ligand in its receptor-bound state: leukotriene B₄ adopts a highly constrained conformation when associated to human BLT2. *J Am Chem Soc* 132:9049–9057

- Catoire LJ, Zoonens M, van Heijenoort C, Giusti F, Guittet E, Popot J-L (2010b) Solution NMR mapping of water-accessible residues in the transmembrane β -barrel of OmpX. *Eur Biophys J* 39:623–630
- Catoire LJ, Damian M, Baaden M, Guittet E, Banères J-L (2011) Electrostatically-driven fast association and perdeuteration allow detection of transferred cross-relaxation for G protein-coupled receptor ligands with equilibrium dissociation constants in the high-to-low nanomolar range. *J Biomol NMR* 50:191–195
- Chabaud E, Barthélémy P, Mora N, Popot J-L, Pucci B (1998) Stabilization of integral membrane proteins in aqueous solution using fluorinated surfactants. *Biochimie* 80:515–530
- Chae PS, Rasmussen SGF, Rana R, Gotfryd K, Chandra R, Goren MA, Kruse AC, Nurva S, Loland CJ, Pierre Y, Drew D, Popot J-L, Picot D, Fox BG, Guan L, Gether U, Byrne B, Kobilka BK, Gellman SH (2010) Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* 7:1003–1008
- Champeil P, Menguy T, Tribet C, Popot J-L, le Maire M (2000) Interaction of amphipols with the sarcoplasmic reticulum Ca^{2+} -ATPase. *J Biol Chem* 275:18623–18637
- Charvolin D, Perez J-B, Rouvière F, Giusti F, Bazzacco P, Abdine A, Rappaport F, Martinez KL, Popot J-L (2009) The use of amphipols as universal molecular adapters to immobilize membrane proteins onto solid supports. *Proc Natl Acad Sci U S A* 106:405–410
- Charvolin D, Picard M, Huang L-S, Berry EA, Popot J-L (2014) Solution behavior and crystallization of cytochrome bc_1 in the presence of amphipols *J Membr Biol*, in the press
- Cvetkov TL, Huynh KW, Cohen MR, Moiseenkova-Bell VY (2011) Molecular architecture and subunit organization of TRPA1 ion channel revealed by electron microscopy. *J Biol Chem* 286:38168–38176
- Dahmane T, Damian M, Mary S, Popot J-L, Banères J-L (2009) Amphipol-assisted *in vitro* folding of G protein-coupled receptors. *Biochemistry* 48:6516–6521
- Dahmane T, Giusti F, Catoire LJ, Popot J-L (2011) Sulfonated amphipols: synthesis, properties and applications. *Biopolymers* 95:811–823
- Dahmane T, Rappaport F, Popot J-L (2013) Amphipol-assisted folding of bacteriorhodopsin in the presence and absence of lipids. Functional consequences. *Eur Biophys J* 42:85–101
- Damian M, Martin A, Mesnier D, Pin J-P, Banères J-L (2006) Asymmetric conformational changes in a GPCR dimer controlled by G-proteins. *EMBO J* 13:5693–5702
- Damian M, Marie J, Leyris J-P, Fehrentz J-A, Verdié P, Martinez J, Banères J-L, Mary S (2012) High constitutive activity is an intrinsic feature of ghrelin receptor protein: a study with a functional monomeric GHS-R1a receptor reconstituted in lipid discs. *J Biol Chem* 287:3630–3641
- Diab C, Tribet C, Gohon Y, Popot J-L, Winnik FM (2007a) Complexation of integral membrane proteins by phosphorylcholine-based amphipols. *Biochim Biophys Acta* 1768:2737–2747
- Diab C, Winnik FM, Tribet C (2007b) Enthalpy of interaction and binding isotherms of non-ionic surfactants onto micellar amphiphilic polymers (amphipols). *Langmuir* 23:3025–3035
- Elter S, Raschle T, Arens S, Viegas A, Gelev V, Eitzkorn M, Wagner G (2014) The use of amphipols for NMR structural characterization of 7-TM proteins. *J Membr Biol*, in the press
- Eitzkorn M, Raschle T, Hagn F, Gelev V, Rice AJ, Walz T, Wagner G (2013) Cell-free expressed bacteriorhodopsin in different soluble membrane mimetics: biophysical properties and NMR accessibility. *Structure* 21:394–401
- Feinstein HE, Tifrea D, Popot J-L, de la MLM, Cocco MJ (2014) Amphipols stabilize the *Chlamydia* major outer membrane protein vaccine formulation *J Membr Biol*, in the press
- Fernandez A, Le Bon C, Baumlin N, Giusti F, Crémel G, Popot J-L, Bagnard D (2014) In vivo characterization of the biodistribution profile of amphipols *J Membr Biol*, in the press
- Ferrandez Y, Dezi M, Bosco M, Urvoas A, Valério M, Le Bon C, Giusti F, Broutin I, Durand G, Polidori A, Popot J-L, Picard M, Minard P (2014) Amphipol-mediated screening of molecular orthoses specific for membrane protein targets *J Membr Biol*, in the press
- Flötenmeyer M, Weiss H, Tribet C, Popot J-L, Leonard K (2007) The use of amphipathic polymers for cryo-electron microscopy of NADH: ubiquinone oxidoreductase (complex I). *J Microsc* 227:229–235
- Garavito RM, Ferguson-Miller S (2001) Detergents as tools in membrane biochemistry. *J Biol Chem* 276:32403–32406

- Giusti F, Popot J-L, Tribet C (2012) Well-defined critical association concentration and rapid adsorption at the air/water interface of a short amphiphilic polymer, amphipol A8-35: a study by Förster resonance energy transfer and dynamic surface tension measurements. *Langmuir* 28:10372–10380
- Giusti F, Kessler P, Westh Hansen R, Lloret N, Le Bon C, Mourier G, Popot J-L, Martinez KL, Zoonens M (2014a) Synthesis of polyhistidine-bearing amphipols and its use for immobilization of membrane proteins. In submission
- Giusti F, Rieger J, Catoire L, Qian S, Calabrese AN, Watkinson TG, Casiraghi M, Radford SE, Ashcroft AE, Popot J-L (2014b) Synthesis, characterization and applications of a perdeuterated amphipol. *J Membr Biol*, DOI 10.1007/s00232-014-9656-x
- Gohon Y, Pavlov G, Timmins P, Tribet C, Popot J-L, Ebel C (2004) Partial specific volume and solvent interactions of amphipol A8-35. *Anal Biochem* 334:318–334
- Gohon Y, Giusti F, Prata C, Charvolin D, Timmins P, Ebel C, Tribet C, Popot J-L (2006) Well-defined nanoparticles formed by hydrophobic assembly of a short and polydisperse random terpolymer, amphipol A8-35. *Langmuir* 22:1281–1290
- Gohon Y, Dahmane T, Ruigrok R, Schuck P, Charvolin D, Rappaport F, Timmins P, Engelman DM, Tribet C, Popot J-L, Ebel C (2008) Bacteriorhodopsin/amphipol complexes: structural and functional properties. *Biophys J* 94:3523–3537
- Gohon Y, Vindigni J-D, Pallier A, Wien F, Celia H, Giuliani A, Tribet C, Chardot T, Briozzo P (2011) High water solubility and fold in amphipols of proteins with large hydrophobic regions: oleosins and caleosin from seed lipid bodies. *Biochim Biophys Acta* 1808:706–716
- Goetzle BM, Hoffman AK, Keyes MH, Gray DN, Ray DG, Sanders CR II (2002) Amphipols can support the activity of a membrane enzyme. *J Am Chem Soc* 124:11594–11595
- Hong W-X, Baker KA, Ma X, Stevens RC, Yeager M, Zhang Q (2011) Design, synthesis and properties of branch-chained maltoside detergents for stabilization and crystallization of integral membrane proteins: human connexin 26. *Langmuir* 26:8690–8696
- Hovers J, Potschies M, Polidori A, Pucci B, Raynal S, Bonneté F, Serrano-Vega M, Tate C, Picot D, Pierre Y, Popot J-L, Nehmé R, Bidet M, Mus-Veteau I, Bußkamp H, Jung K-H, Marx A, Timmins PA, Welte W (2011) A class of mild surfactants that keep integral membrane proteins water-soluble for functional studies and crystallization. *Mol Membr Biol* 28:171–181
- Jonkheijm P, Weinrich D, Schröder H, Niemeyer CM, Waldmann H (2008) Chemical strategies for generating protein biochips. *Angew Chem Int Ed Engl* 47:9618–9647
- Karlsson R, Fält A (1997) Experimental design for kinetic analysis of protein-protein interactions with surface plasmon resonance biosensors. *J Immunol Method* 200:121–133
- Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T, Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett* 442:15–19
- Knowles TJ, Finka R, Smith C, Lin Y-P, Dafforn T, Overduin M (2009) Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc* 131:7484–7485
- Koutsopoulos S, Kaiser L, Eriksson HM, Zhang S (2012) Designer peptide surfactants stabilize diverse functional membrane proteins. *Chem Soc Rev* 41:1721–1728
- Le Bon C, Della Pia EA, Giusti F, Lloret N, Zoonens M, Martinez KL, Popot J-L (2014a) Synthesis of an oligonucleotide-derivatized amphipol and its use to trap and immobilize membrane proteins *Nucleic Acids Res*, DOI: 10.1093/nar/gku250.
- Le Bon C, Popot J-L, Giusti F (2014b) Labeling and functionalizing amphipols for biological applications. *J Membr Biol*, DOI 10.1007/s00232-014-9655-y
- Leney AC, McMorran LM, Radford SE, Ashcroft AE (2012) Amphipathic polymers enable the study of functional membrane proteins in the gas phase. *Anal Chem* 84:9841–9847
- Liao M, Cao E, Julius D, Cheng Y (2013) Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 504:107–112
- Long AR, O'Brien CC, Malhotra K, Schwall CT, Albert AD, Watts A, Alder NN (2013) A detergent-free strategy for the reconstitution of active enzyme complexes from native biological membranes into nanoscale discs. *BMC Biotechnol* 13:41. doi:10.1186/1472-6750-1113-1141

- Martinez KL, Gohon Y, Corringier P-J, Tribet C, Mérola F, Changeux J-P, Popot J-L (2002) Allosteric transitions of *Torpedo* acetylcholine receptor in lipids, detergent and amphipols: molecular interactions vs. physical constraints. *FEBS Lett* 528:251–256
- Matar-Merheb R, Rhimi M, Leydier A, Huché F, Galián C, Desuzinges-Mandon E, Ficheux D, Flot D, Aghajari H, Kahn R, Di Pietro A, Jault J-M, Coleman AW, Falson P (2011) Structuring detergents for extracting and stabilizing functional membrane proteins. *PLoS ONE* 6:e18036
- McGregor C-L, Chen L, Pomroy NC, Hwang P, Go S, Chakrabarty A, Privé GG (2003) Lipopeptide detergents designed for the structural study of membrane proteins. *Nat Biotechnol* 21:171–176
- Nagy JK, Kuhn Hoffmann A, Keyes MH, Gray DN, Oxenoid K, Sanders CR (2001) Use of amphipathic polymers to deliver a membrane protein to lipid bilayers. *FEBS Lett* 501:115–120
- Ning Z, Hawley B, Seebun D, Figeys D (2014) APols aided protein precipitation: a rapid method for protein concentrating for proteomic analysis. *J Membr Biol*, in the press
- Opačić M, Giusti F, Broos J, Popot J-L (2014) Amphipol A8-35 preserves the activity of detergent-sensitive mutants of *Escherichia coli* mannitol permease EIMtl. *J Membr Biol*, in the press
- Park K-H, Billon-Denis E, Dahmane T, Lebaupain F, Pucci B, Breyton C, Zito F (2011) In the cauldron of cell-free synthesis of membrane proteins: playing with new surfactants. *New Biotech* 28:255–261
- Perlmutter JD, Drasler WJ, Xie W, Gao J, Popot J-L, Sachs JN (2011) All-atom and coarse-grained molecular dynamics simulations of a membrane protein stabilizing polymer. *Langmuir* 27:10523–10537
- Perlmutter JD, Popot J-L, Sachs JN (2014) Molecular dynamics simulations of a membrane protein/amphipol complex. *J Membr Biol*, in the press
- Picard M, Dahmane T, Garrigos M, Gauron C, Giusti F, le Maire M, Popot J-L, Champeil P (2006) Protective and inhibitory effects of various types of amphipols on the Ca²⁺-ATPase from sarcoplasmic reticulum: a comparative study. *Biochemistry* 45:1861–1869
- Plancharid N, Point E, Dahmane T, Giusti F, Renault M, Le Bon C, Durand G, Milon A, Guittet E, Zoonens M, Popot J-L, Catoire LJ (2014) The use of amphipols for solution NMR studies of membrane proteins: advantages and limitations as compared to other solubilizing media. *J Membr Biol*, DOI 10.1007/s00232-014-9654-z
- Pocanschi CL, Dahmane T, Gohon Y, Rappaport F, Apell H-J, Kleinschmidt JH, Popot J-L (2006) Amphipathic polymers: tools to fold integral membrane proteins to their active form. *Biochemistry* 45:13954–13961
- Pocanschi C, Popot J-L, Kleinschmidt JH (2013) Folding and stability of outer membrane protein A (OmpA) from *Escherichia coli* in an amphipathic polymer, amphipol A8-35. *Eur Biophys J* 42:103–118
- Polovinkin V, Gushchin I, Balandin T, Chervakov P, Round E, Schevchenko V, Popov A, Borshchevskiy V, Popot J-L, Gordeliy V (2014) High-resolution structure of a membrane protein transferred from amphipol to a lipidic mesophase. *J Membr Biol*, in the press
- Popot J-L (2010) Amphipols, nanodiscs, and fluorinated surfactants: three non-conventional approaches to studying membrane proteins in aqueous solutions. *Annu Rev Biochem* 79:737–775
- Popot J-L, Engelman DM (2000) Helical membrane protein folding, stability and evolution. *Annu Rev Biochem* 69:881–923
- Popot J-L, Gerchman S-E, Engelman DM (1987) Refolding of bacteriorhodopsin in lipid bilayers: a thermodynamically controlled two-stage process. *J Mol Biol* 198:655–676
- Popot J-L, Berry EA, Charvolin D, Creuzenet C, Ebel C, Engelman DM, Flötenmeyer M, Giusti F, Gohon Y, Hervé P, Hong Q, Lakey JH, Leonard K, Shuman HA, Timmins P, Warschawski DE, Zito F, Zoonens M, Pucci B, Tribet C (2003) Amphipols: polymeric surfactants for membrane biology research. *Cell Mol Life Sci* 60:1559–1574
- Popot J-L, Althoff T, Bagnard D, Banères J-L, Bazzacco P, Billon-Denis E, Catoire LJ, Champeil P, Charvolin D, Cocco MJ, Crémel G, Dahmane T, de la MLM, Ebel C, Gabel F, Giusti F, Gohon Y, Goormaghtigh E, Guittet E, Kleinschmidt JH, Kühlbrandt W, Le Bon C, Martinez KL, Picard M, Pucci B, Rappaport F, Sachs JN, Tribet C, van Heijenoort C, Wien F, Zito F, Zoonens M (2011) Amphipols from A to Z. *Annu Rev Biophys* 40:379–408

- Prata C, Giusti F, Gohon Y, Pucci B, Popot J-L, Tribet C (2001) Non-ionic amphiphilic polymers derived from *tris*(hydroxymethyl)-acrylamidomethane keep membrane proteins soluble and native in the absence of detergent. *Biopolymers* 56:77–84
- Privé G (2009) Lipopeptide detergents for membrane protein studies. *Curr Opin Struct Biol* 19:1–7
- Rahmeh R, Damian M, Cottet M, Orcel H, Mendre C, Durroux T, Sharma KS, Durand G, Pucci B, Trinquet E, Zwier JM, Deupi X, Bron P J-LB, Mouillac B, Granier S (2012) Structural insights into biased G protein-coupled receptor signaling revealed by fluorescence spectroscopy. *Proc Natl Acad Sci U S A* 109:6733–6738
- Rajesh S, Knowles TJ, Overduin M (2011) Production of membrane proteins without cells or detergents. *N Biotech* 28:250–254
- Raschle T, Hiller S, Etzkorn M, Wagner G (2010) Nonmicellar systems for solution NMR spectroscopy of membrane proteins. *Curr Opin Struct Biol* 20:471–479
- Rich RL, Myszka DG (2005) Survey of the year 2004 commercial optical biosensor literature. *J Mol Recognit* 18:431–478
- Schafmeister CE, Miercke LJW, Stroud RA (1993) Structure at 2.5 Å of a designed peptide that maintains solubility of membrane proteins. *Science* 262:734–738
- Sharma KS, Durand G, Gabel F, Bazzacco P, Le Bon C, Billon-Denis E, Catoire LJ, Popot J-L, Ebel C, Pucci B (2012) Non-ionic amphiphilic homopolymers: synthesis, solution properties, and biochemical validation. *Langmuir* 28:4625–4639
- Stenberg E, Persson B, Roos H, Urbaniczky C (1991) Quantitative determination of surface concentration of protein with surface plasmon resonance using radio-labeled proteins. *J Colloid Interface Sci* 143:513–526
- Tifrea D, Pal S, Cocco MJ, Popot J-L, de la Maza LM (2014) Increased immuno accessibility of MOMP epitopes in a vaccine formulated with amphipols may account for the very robust protection elicited against a vaginal challenge with *C. muridarum*. *J Immunol*, in the press
- Tifrea DF, Sun G, Pal S, Zardeneta G, Cocco MJ, Popot J-L, de la MLM (2011) Amphipols stabilize the *Chlamydia* major outer membrane protein and enhance its protective ability as a vaccine. *Vaccine* 29:4623–4631
- Tribet C, Audebert R, Popot J-L (1996) Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci U S A* 93:15047–15050
- Tribet C, Audebert R, Popot J-L (1997) Stabilisation of hydrophobic colloidal dispersions in water with amphiphilic polymers: application to integral membrane proteins. *Langmuir* 13:5570–5576
- Tribet C, Mills D, Haider M, Popot J-L (1998) Scanning transmission electron microscopy study of the molecular mass of amphipol/cytochrome b_6f complexes. *Biochimie* 80:475–482
- Tribet C, Diab C, Dahmane T, Zoonens M, Popot J-L, Winnik FM (2009) Thermodynamic characterization of the exchange of detergents and amphipols at the surfaces of integral membrane proteins. *Langmuir* 25:12623–12634
- Wang X, Corin K, Baaske P, Wienken CJ, Jerabek-Willemsen M, Dühr S, Braun D, Zhang S (2011) Peptide surfactants for cell-free production of functional G protein-coupled receptors. *Proc Natl Acad Sci U S A* 108:9049–9054
- Zhao X, Nagai Y, Reeves PJ, Kiley P, Khorana HG, Zhang S (2006) Designer short peptide surfactants stabilize G protein-coupled receptor bovine rhodopsin. *Proc Natl Acad Sci U S A* 103:17707–17712
- Zoonens M (2004) Caractérisation des complexes formés entre le domaine transmembranaire de la protéine OmpA et des polymères amphiphiles, les amphipols. Application à l'étude structurale des protéines membranaires par RMN à haute résolution. Thèse de Doctorat, Université Paris-6, Paris
- Zoonens M, Popot J-L (2014) Amphipols for each season. *J Membr Biol*, in the press
- Zoonens M, Catoire LJ, Giusti F, Popot J-L (2005) NMR study of a membrane protein in detergent-free aqueous solution. *Proc Natl Acad Sci U S A* 102:8893–8898
- Zoonens M, Giusti F, Zito F, Popot J-L (2007) Dynamics of membrane protein/amphipol association studied by Förster resonance energy transfer. Implications for *in vitro* studies of amphipol-stabilized membrane proteins. *Biochemistry* 46:10392–10404
- Zubay G (1973) *In vitro* synthesis of protein in microbial systems. *Annu Rev Genet* 7:267–287

Chapter 8

New Amphiphiles to Handle Membrane Proteins: “Ménage à Trois” Between Chemistry, Physical Chemistry, and Biochemistry

Grégory Durand, Maher Abla, Christine Ebel and Cécile Breyton

8.1 Introduction

Membrane proteins (MPs) are the front door of the cell in that they allow the communication between the cell and the outside world: They are receptors to a large variety of signals, the transporter of metabolites and catabolites in and out of the cell and between organelles within the cell, and also ensure energy conversion in specialized membranes. As such, they are the target of a large proportion of drugs and poisons (Yildirim et al. 2007). From a structural point of view, they are embedded in a lipid bilayer: The domain in contact with the acyl chains of lipids is hydrophobic and needs to be shielded from water when the protein is manipulated out of the membrane for biochemical, biophysical, and structural studies. This is done thanks to amphiphilic molecules, detergents, which are also used to solubilize the membrane and extract the MP from it.

Lipids and detergents share the same amphipathic character: They both belong to the larger class of surfactants, surface active agents, which group all amphiphilic molecules, composed of a hydrophilic (anionic, neutral, or zwitterionic) head

G. Durand (✉) · M. Abla
Avignon University, Avignon, France
e-mail: gregory.durand@univ-avignon.fr

G. Durand
Institute of Biomolecules Max Mousseron, UMR-CNRS 5247, Montpellier, France

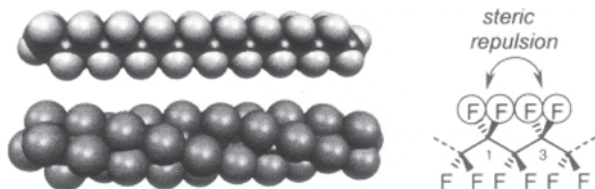
C. Breyton (✉) · C. Ebel
Institute of Structural Biology, French National Centre for Scientific Research,
Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France
e-mail: cecile.breyton@ibs.fr

and a hydrophobic part (aliphatic or cyclic carbon chains or aromatic groups), that partition at the air–water interface. Throughout this chapter, lipids will refer to the amphiphiles that compose biological membranes, whereas detergents refer to surfactants that are used to solubilize membranes and manipulate MPs. We will use the term surfactant for amphiphiles that are neither, but that can solubilize MPs in solution. Lipids and detergents differ in their geometry: Generally, lipids bear two hydrophobic acyl chains, and conventional detergents usually only one. This difference in geometry has a direct influence on the type of aggregates formed by these amphiphilic molecules. Lipids, which have a cylindrical shape, form planar aggregates, the bilayer, whereas the geometry of the detergent molecule, closer to a cone, induces the formation of small globular aggregates, the micelles. The solubility of lipids and detergents also differs noticeably. Detergents are soluble to a certain extent: At low concentration, they are soluble as monomers in the aqueous buffer, and micelles form only above the “critical micellar concentration” (cmc). The detergent concentration should always be above its cmc to allow the correct solubilization of a MP; thus, the solubilized protein coexists with free detergent micelles.

The MP is thus either embedded in the lipid bilayer, its native environment, or surrounded by a detergent, or surfactant, layer (le Maire et al. 2000) after extraction from the membrane, allowing biochemical, biophysical, and structural studies. MPs however are usually less stable in a detergent environment when compared to their native lipid bilayer. This hampers their biochemical, biophysical, and structural study, and even though the Protein Data Bank (PDB) certainly sees the number of MP structure growing exponentially, this rate is far behind that of soluble protein (<http://blanco.biomol.uci.edu/mpstruc/>). The main causes are MP’s difficult overexpression, their poor stability when extracted from the membrane, and the difficulty to obtain good-quality crystals. The first and third issues are dealt with in Chaps. 1, 2, 4, 5, and 11 of this book (Mus-Veteau et al., Proverbio et al., Attab et al., Bakari et al., and Hishchenko et al., respectively).

The hypotheses to explain MP destabilization by detergent include the dilution of stabilizing lipids and hydrophobic cofactors in the hydrophobic phase formed by the free micelles (Breyton et al. 1997), and their dissociative properties toward intramolecular protein–protein interactions, related to their shorter and more flexible hydrophobic tail when compared to lipids. In addition, the difference in lateral pressure exerted by the detergent molecules as compared to that exerted by the lipids of the bilayer would impact the highly dynamic nature of some MPs. Various strategies have been developed to address MP instability in detergent. They are reviewed in Chap. 1 of this book (see also Kang et al. 2013). They include stabilizing protein point mutations (Tate and Schertler 2009), embedding the protein in small membrane patches such as nanodiscs (Borch and Hamann 2009) and bicelles (Diller et al. 2009; Ujwal and Bowie 2011; see also Chap. 12 by Catoire et al. in this book), and the development of new amphiphiles, designed to be less denaturing than classical detergents. Whereas the choice of the “ideal detergent” is still highly empirical, purpose specific, and protein dependent, the development of new amphiphiles has followed some rational. A dedicated Chap. 7 by Zoonens et al. reviews the principle and applications of

Fig. 8.1 The zigzag conformation of octadecane (*above*) compared with the helical conformation of perfluorooctadecane. (From Kirsch 2004)



amphipols, which have multipoint attachment to the hydrophobic domain of protein and in principle, allow to avoid the hydrophobic phase formed by the excess detergent micelles.

In this chapter, we focus mostly on the design and use for MP study of new amphiphilic molecules whose polar head comprises sugar moieties and whose hydrophobic chain comprises fluorine atoms. Due to the peculiar properties of fluorinated chain, these molecules are not able to solubilize biological membranes and therefore do not belong to the class of detergents. The main advantage of resorting to fluorinated surfactants for MP study lies in their nonintrusive properties, which allow to retain the native structure of fragile proteins. In the first part, we introduce the concepts motivating the exploration of different *F*-surfactant families, we then present in the second part the design, and the physical-chemical and self-assembling properties, of recently developed—including new molecules—*F*-surfactants bearing branched polar heads whose molecular shape can be easily varied so as to form globular micelles. In the third part, the recent biochemical applications of these *F*-surfactants are discussed. Finally, in the last part, an overview of the recent literature on sugar-based amphiphiles for biochemical and structural applications is given.

8.2 Fluorinated Surfactants: General Concepts

8.2.1 General Properties of the Fluorinated Surfactants

F-surfactants resemble detergents in that they share the same general structure, but the hydrophobic tail instead of being hydrogenated is fully or partially fluorinated. The substitution of fluorine for hydrogen atoms is responsible for the peculiar properties and applications of *F*-surfactants. The low polarizability and larger size of the fluorine atom compared to that of the hydrogen atom have consequences on the structure and properties of fluorocarbon chains. Cross section of *F*-chains is larger than that of *H*-chains (30 \AA^2 vs. 20 \AA^2 , respectively) and CF_3 and CF_2 groups are bulkier than CH_3 and CH_2 (Riess 2005). Also, whereas *H*-chains possess a linear zigzag structure, a more reduced conformational freedom is observed with *F*-chains as a consequence of the helical structure of fluorocarbon chains (Fig. 8.1). Alkanes

Table 8.1 Comparison of surface tension and phobicity of water, alkanes, and fluoroalkanes. (Adapted from Riess 2005)

Phase	Surface tension (mN/m)	Phobicity
Water	72	Lipo- and fluorophobic
<i>H</i> -carbons	15–30	Hydro- and fluorophobic
<i>F</i> -carbons	10–20	Lipo- and hydrophobic

and perfluorinated alkanes, while both being hydrophobic, are poorly miscible, a property that motivated their use for MP biochemistry.

The larger fluorine atoms increase the hydrophobicity of the surfactant, and as a result, *F*-surfactants exhibit very low CMC. As a rule of thumb, the contribution of a CF₂ group to the CMC of a surfactant is 1.5 times higher than that of a CH₂ group (Sadtler et al. 1998). The tight packing of the *F*-chains at the air/water interface and the strong hydrophobic interactions among *F*-chains lower the surface energies and result in low surface tension values compared to that of their fully hydrogenated analogues (Table 8.1). Therefore, *F*-surfactants are commonly called “super surfactants.” The strong hydrophobic noncovalent interactions between *F*-chains are responsible for the formation of very stable self-assemblies in aqueous solution. *F*-surfactants are also known for their chemical and thermal stability. They have diversified uses in material science and in biomedical fields as well as in several other applications such as coating, firefighting, herbicide and insecticide formulations, cosmetics, paints, adhesives, and drug delivery (Kissa 2001).

8.2.2 *General Motivation for Using F-surfactants for the Study of MPs*

The original idea to substitute part of the hydrophobic chain by a lipophobic *F*-chain was the transfer of the protein in a hydrophobic phase that would be a poor solvent for lipids and hydrophobic cofactors, and composed of less denaturing molecules. *F*-surfactants are indeed expected to poorly solubilize lipids and cofactors that may be essential for the protein activity and stability. Moreover, because *F*-chains are bulky, *F*-surfactant hydrophobic chains are expected to intrude less easily than any *H*-surfactants into the structure of the MP.

However, due to the lipophobicity of *F*-chains, *F*-surfactants have a poor detergent character. They are unable to dissociate cell membranes and extract MPs. Therefore, from a practical point of view, detergents are generally needed for MP extraction before being exchanged to *F*-surfactants for further use. Alternative protocols without extraction steps are found when refolding the protein or when producing it in cell-free synthesis directly in *F*-surfactants. In addition, the low affinity of *F*-chains toward the hydrogenated transmembrane segments of the protein may also induce aggregation of the protein. Since it was demonstrated that it is

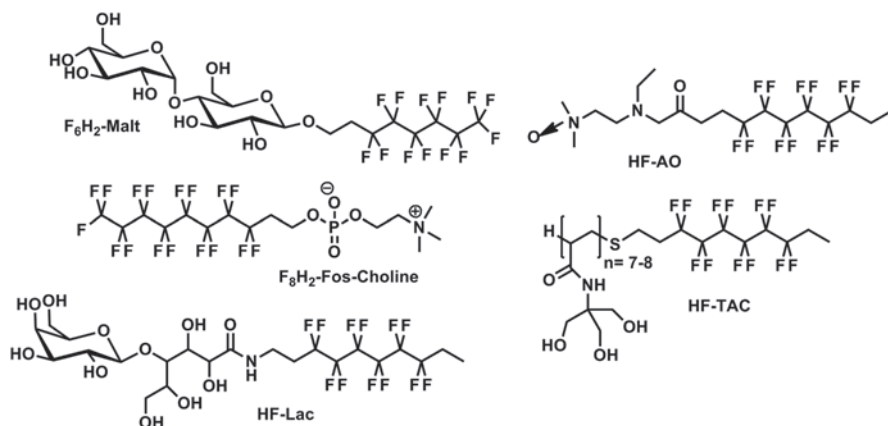


Fig. 8.2 Chemical structures of some *F*- and *HF*-surfactants. F₆H₂-Malt and F₈H₂-Fos-Choline are available from Anatrace. H₂F₆-AO, H₂F₆-Lac and H₂F₆-TAC were synthesized in the laboratory

mainly the extremity of the surfactant that interacts with the hydrophobic part of MPs (le Maire et al. 2000), a short hydrocarbon tip has been added to the fluorinated tail of *F*-surfactants leading to the so-called hemifluorinated surfactants (*HF*-surfactants). This addition aimed to preserve MP from long-term aggregation noticed in the seminal work by Chabaud et al. (1998). The addition of the hydrocarbon tip was expected to be a compromise between the lipophobic, nondetergent character of *F*-chains, and a necessity to confer affinity toward the lipophilic domains of the protein (Breyton et al. 2004).

8.2.3 Exploring the Previous Families of Fluorinated Surfactants

Several series of *F*-surfactants have been synthesized over the past 15 years and tested as tools for handling MPs in aqueous solutions. Figure 8.2 presents some of the surfactants produced in the laboratory of Avignon as well as some commercially available derivatives.

Different polar heads have been investigated. The TAC series (Fig. 8.2) bearing an oligomeric polar head group has been the most used (recent findings are summarized in Sect. 8.5). Among the other chemically defined ones, surfactants bearing aminoxide polar head (Fig. 8.2) were found to be destabilizing (Chaudier et al. 2002), whereas those bearing chemically well-defined sugar-based polar head groups, H₂F₆-Lac (Fig. 8.2; Lebaupain et al. 2006) and H₂F₆-Malt (Polidori et al. 2006) derivatives were found much more efficient in keeping MPs water soluble and active. This showed that a saccharide moiety could be an appropriate

polar head group for keeping the protein water soluble and stable. *HF*-surfactants have shown superiority over *F*-surfactants in preserving the stability of model proteins in some cases (Breyton et al. 2004, 2010). However, as their preparation requires complex and longer synthesis, their use is only preferred when *F*-surfactants failed to prevent aggregation. Interestingly, we also demonstrated that the nature of the sugar could affect the stability of the protein. Indeed, the maltoside polar head group with a glucose moiety was found more favorable than the lactobionamide group with a galactose moiety. However, none of these sugar-based surfactants formed small and homogeneous micelles, whatever the nature of the chain (Lebaupain et al. 2006; Polidori et al. 2006). Indeed, the heterogeneity of *F*-surfactant micelles was frequently noted to induce highly polydisperse protein/surfactant complexes, which is one of the major drawbacks for most in vitro studies.

Previous work has shown that the TAC series with a polymeric polar head group formed small micelles (Dupont et al. 2003), and was successfully used to stabilize fragile MPs (Breyton et al. 2004). Its availability in large quantities allowed the development of biochemical applications (Breyton et al. 2010), the most recent ones being discussed in Sect. 8.5. However, the heterogeneity of the polymeric head group of the TAC derivatives appeared to be the main limitation for their development. There was therefore a need to develop *F*-surfactants with chemically defined polar head groups that would form small micelles, and compact and homogeneous complexes with MPs.

8.2.4 *Relating the Shape of the Molecule to that of its Assemblies*

Tanford (1980) and Israelachvili, Mitchell and Ninham (1977) provided, more than 30 years ago, general insights on how molecular structure of amphiphiles can control the shape and size of the resulting aggregates. The concept of molecular packing parameter has been widely used to explain and rationalize the shape of molecular self-assembly of surfactants in aqueous solution. The critical packing parameter (CPP) links the geometry of the surfactant molecule to that of the aggregate formed (Fig. 8.3). It is defined as v_0/al_0 , where v_0 is the surfactant tail volume, l_0 its length, and a the area per molecule. The molecular packing parameter determines the shape of the aggregate: $0 \leq v_0/al_0 \leq 1/3$ for spherical aggregates, $1/3 \leq v_0/al_0 \leq 1/2$ for cylindrical aggregates and $1 \leq v_0/al_0 \leq 1/2$ for bilayer aggregates (Fig. 8.3). Therefore, by calculating the CPP value of a surfactant, one can predict if small globular compact or large—rods or vesicle—aggregates will be formed in aqueous solution.

While a can be precisely estimated using free energy models (leading to a_e , the equilibrium area per molecule; Nagarajan 2002), it is often identified as a simple geometrical property of the head group of the surfactant. We consider for a , the area A_{\min} determined from surface tension measurements.

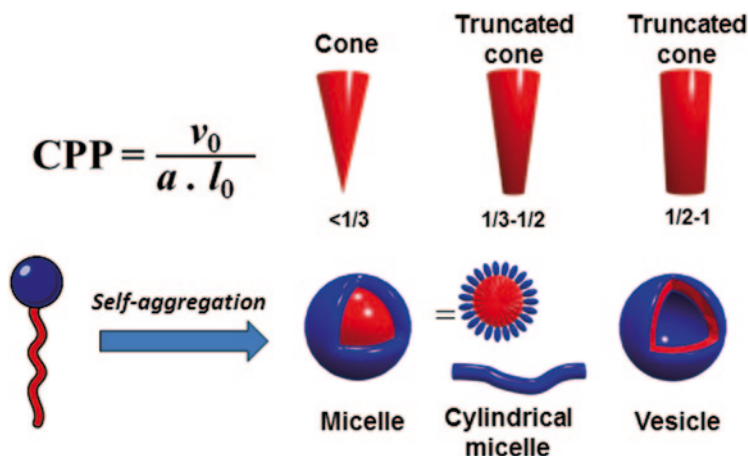


Fig. 8.3 Schematic illustration of the critical packing parameter

8.3 Physical–Chemical Properties of *F*-surfactants with Branched Polar Heads

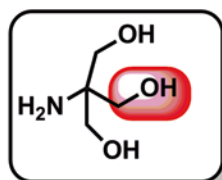
8.3.1 Design of Two Series of *F*-surfactants with Branched Polar Heads

Based on the biochemical results obtained with the first synthesized (*H*)*F*-surfactants, we have concluded that surfactants that form small, globular, and homogeneous micelles such as the TAC series (Dupont et al. 2003) could lead to monodisperse complexes with MPs (Breyton et al. 2004). Therefore, (*H*)*F*-surfactants with bulky polar head groups are expected to form smaller aggregates than those having small polar head groups, by increasing the interfacial curvature of the aggregates formed, which in turn could lead to monodisperse complexes with MPs.

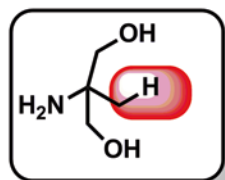
We next focused on designing *F*-surfactants whose volume of the polar head group could be easily tuned, while keeping the possibility of switching from the perfluorinated chain to hybrid ones if needed. This allowed to study in a systematic way the effect of the volumetric ratio of the surfactant on its physical-chemical and biochemical properties, and to choose thereby the most appropriate candidate surfactant for MP studies.

In that aim, two molecules were chosen as basic structures for the glycosylated polar head group: The tris(hydroxymethyl)aminomethane (TRIS) that possesses three hydroxyl groups on which we can graft one, two, or three sugar moieties, and the 2-amino-2-methyl propane-1,3-diol that has two hydroxyl groups (Fig. 8.4). From a synthesis point of view, 2-amino-2-methyl propane-1,3-diol allows an easy and direct grafting of two glucose moieties. Two series of (*H*)*F*-surfactants with glycosylated polar heads were thus synthesized and studied. In the first series,

Two molecules used as basic structures for the synthesis of poly glucose polar head groups

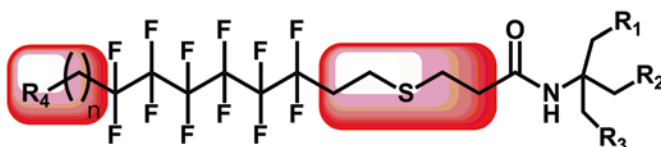


Tris(hydroxymethyl)aminomethane



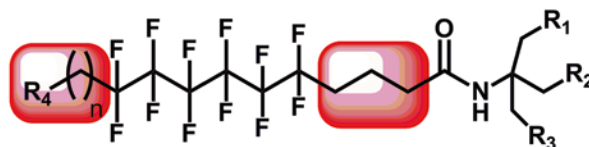
2-amino-2-methyl propane-1,3-diol

Two series of fluorinated surfactants



HF series : $R_4 = \text{CH}_3$, $n = 1$
 F series : $R_4 = \text{F}$, $n = 0$

First series



HF series : $R_4 = \text{CH}_3$, $n = 1, 2$
 F series : $R_4 = \text{F}$, $n = 0$

Second series

Fig. 8.4 Chemical structures of the two recent series of *HF*-surfactants. For R_1, R_2, R_3 see Table 8.2

the perfluorohexane central core is linked to the glycosylated polar head group via a thioether bond (CH_2SCH_2) and an ethyl-end group can be added (Abla et al. 2008, 2012). In the second series, the *HF*-chain comprises the same fluorocarbon central core and a shorter alkyl group is used as linker (Fig. 8.4; Abla et al. 2011; Abla et al. in preparation). The chain can be decorated at its extremity with a propyl tip. For the sake of comparison, perhydrogenated analogues of the first series were also prepared.

In the first series of *HF*-compounds bearing an ethyl tip, the following nomenclature was adopted: H_2F_6 , where H_2 stands for the ethyl tip and F_6 for the perfluorohexane core. Their fluorinated and hydrogenated analogues are labeled by F_6 and H_x (where “x” corresponds to the number of carbons located before the sulfur atom in the hydrophobic tail), respectively. In that nomenclature, for the fluorinated

and hemifluorinated derivatives, the linker between the fluorinated core and the polar head ($\text{CH}_2\text{CH}_2\text{SCH}_2\text{CH}_2$) was omitted. The second series of *HF*-surfactants prepared was labeled $\text{H}_3\text{F}_6\text{H}_3$, where H_3 , F_6 , and H_3 stand for the propyl tip, the perfluorohexane central core, and the hydrogenated linker between the polar and apolar groups, respectively. Once again, their fluorinated analogs were also prepared and were called F_6H_3 . In both series, the number and the nature of the sugar moieties grafted onto the polar head group was also labeled as Monoglu, Diglu, and Triglu for one, two, and three glucose moieties, respectively. In order to differentiate between the TRIS and 2-amino-2-methyl propane-1,3-diol derivatives, an M was added to the naming of the latter derivatives. For galactose-based compound, the label Gal was used.

In the next part, we will present the physical-chemical properties of these series of surfactant. Some of the compounds are original. We will present briefly the methods we use for their characterization. We will then discuss the surface properties and the size and shape of the micelles formed by these *F*-surfactants. We will show how, by fine-tuning the size of the polar head, we can vary the shape of the aggregates formed.

8.3.2 Measured Parameters and Methods

- Surface tension properties of surfactants can be determined by the Wilhelmy plate technique. This technique consists in measuring the surface tension of the solution at different concentrations. Two straight lines are typically obtained, one exhibiting a linear concentration-dependent negative slope and one being a plateau with a constant surface tension value (Fig. 8.5). The intersection of these two lines corresponds to the CMC value of the surfactant. From this technique, several parameters can be determined. However, in this chapter we will discuss only those we consider important for MP manipulations: the CMC, the limit surface tension (γ_{max}) attained at the CMC, and the minimum area occupied by the surfactant (A_{min}) at the air/water interface given by the slope of the first line. The γ_{max} and A_{min} provide information on the packing of the surfactant at the interface and the volume of its polar head.
- The hydrophobic character ($\log k'_{\text{w}}$) of surfactants is determined by reversed phase HPLC on C_{18} -hydrophobic columns. This parameter is closely related to the molecule water/octanol partition coefficient (Braumann 1986).
- Dynamic light scattering (DLS) characterizes the distribution of the hydrodynamic diameters of the aggregates formed by the surfactants.
- Analytical ultracentrifugation (AUC) sedimentation velocity, which combines particle separation and rigorous analysis, is a powerful technique for characterizing whether the size of surfactant micelle changes with concentration. This technique is discussed in detail in Chap. 10 by Le Roy et al. in this book. The CMC is determined by following the amounts of sedimenting micelles

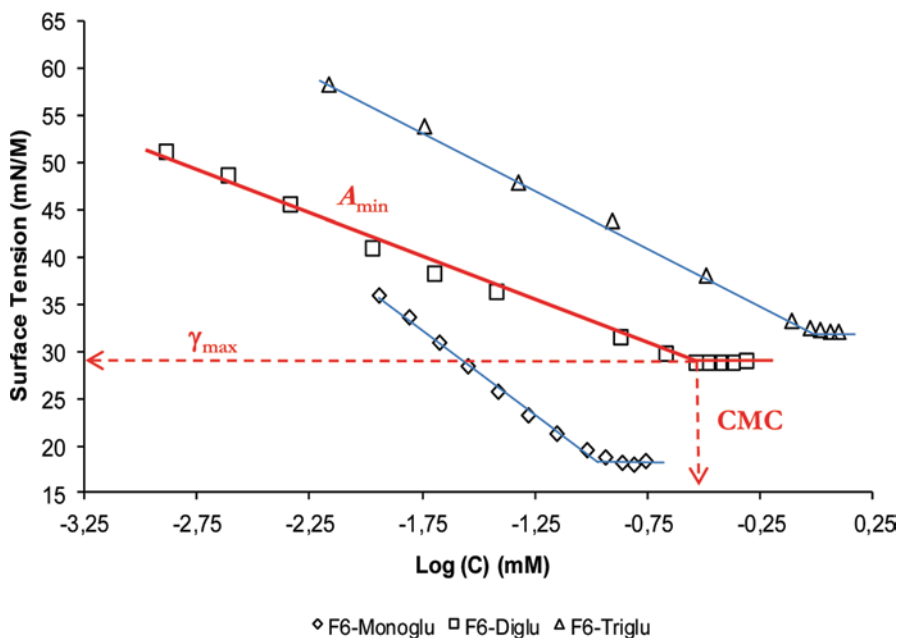


Fig. 8.5 Examples of critical micellar concentration curves for F_6 Monoglu (\diamond), F_6 Diglu (\square), and F_6 Triglu (\triangle)

in concentration series. From the measure of the sedimentation coefficient, combined with the translation diffusion coefficient (obtained in general from DLS), the mass and hydrodynamic radius of the micelles can be determined. A preliminary characterization of the inter-micelle interaction may also be done.

- Small-angle neutron or X-rays scattering (SANS or SAXS) complements AUC. From the measurement of the forward intensity, the mass (number of aggregation, N_{agg}) and the size of the micelles (radius of gyration, R_g) can be obtained. In addition, the shape can be determined by the analysis of the whole scattering curve (as an example, the rod shape of F_6 Monoglu described below; Breyton et al. 2009). It is also a key technique for characterizing inter-micelle interactions.

Table 8.2 reports the CMC, A_{min} , γ_{max} from surface tension and $\log k'_w$ of the recently investigated F -surfactants. It is used, with some other data provided in the text, to analyze the contribution of the alkyl tip and hydrocarbon spacer length as well as the polar head on the surface activity and hydrophobicity of the surfactants.

Table 8.2 Physical-chemical data for the two series of *F*-surfactants with branched polar heads

Chemical structure												Surface tension data		Hydrophobic character	Ref		
Series	N°	Surfactant Name	MW (g/mol)	Tail		Polar head group						CMC mM	A_{min} (Å ²)	γ_{max} (mN/m)	log k'_w		
				Name	R	Name	R ₁	R ₂	R ₃	R ₃							
First	1	H ₁₀ Diglu	673.3	H ₁₀	C ₁₀ H ₂₁ SC ₂ H ₄	Diglu	OGlu	OGlu	OGlu	OH	OH	0.555	373.7	92.0	36.0	4.4	Abla et al. 2008
	2	H ₁₀ Triglu	835.4	H ₁₀	C ₁₀ H ₂₁ SC ₂ H ₄	Triglu	OGlu	OGlu	OGlu	OGlu	OGlu	1.175	981.6	107.7	38.5	4.2	
	3	H ₁₂ Diglu	701.4	H ₁₂	C ₁₂ H ₂₅ SC ₃ H ₄	Diglu	OGlu	OGlu	OGlu	OH	OH	0.039	27.4	102.7	40.7	5.6	
	4	H ₁₂ Triglu	863.4	H ₁₂	C ₁₂ H ₂₅ SC ₃ H ₄	Triglu	OGlu	OGlu	OGlu	OGlu	OGlu	0.102	88.1	125.0	41.5	5.4	
	5	F ₆ Monoglu	717.1	F ₆	C ₆ F ₁₃ C ₂ H ₄ SC ₂ H ₄	Monoglu	OGlu	OH	OH	OH	OH	0.113	81.0	52.7	18.3	5.2	
	6	F ₆ Diglu	879.1	F ₆	C ₆ F ₁₃ C ₂ H ₄ SC ₂ H ₄	Diglu	OGlu	OGlu	OGlu	OH	OH	0.233	204.8	97.5	28.3	4.8	
	7	F ₆ DigluM	863.2	F ₆	C ₆ F ₁₃ C ₂ H ₄ SC ₂ H ₄	DigluM	OGlu	OGlu	OGlu	H	H	0.359	309.9	74.7	27.7	5.1	Abla et al. 2012
	8	F ₆ Triglu	1,041.6	F ₆	C ₆ F ₁₃ C ₂ H ₄ SC ₃ H ₄	Triglu	OGlu	OGlu	OGlu	OGlu	OGlu	0.947	986.4	114.0	32.3	4.7	Abla et al. 2008
	9	F ₆ Trigal	1,041.6	F ₆	C ₆ F ₁₃ C ₂ H ₄ SC ₂ H ₄	Trigal	OGal	OGal	OGal	OGal	OGal	0.730	760.4	114.5	35.5	4.4	This work
	10	H ₂ F ₆ Diglu	889.2	H ₂ F ₆	C ₂ H ₅ C ₆ F ₁₂ C ₂ H ₄ S C ₂ H ₄	Diglu	OGlu	OGlu	OGlu	OH	OH	0.347	308.6	104.5	36.0	4.9	Abla et al. 2008
Second	11	H ₂ F ₆ Triglu	1,051.8	H ₂ F ₆	C ₂ H ₅ C ₆ F ₁₂ C ₂ H ₄ S C ₂ H ₄	Triglu	OGlu	OGlu	OGlu	OGlu	OGlu	0.747	785.7	131.5	36.0	4.7	Abla et al. 2008
	12	H ₂ F ₆ Trigal	1,051.8	H ₂ F ₆	C ₂ H ₅ C ₆ F ₁₂ C ₂ H ₄ S C ₂ H ₄	Trigal	OGal	OGal	OGal	OGal	OGal	0.423	444.9	131.0	40.3	4.6	This work
	13	F ₆ H ₃ Diglu	833.2	F ₆ H ₃	C ₆ F ₁₃ C ₃ H ₆	Diglu	OGlu	OGlu	OGlu	OH	OH	1.027	855.7	126.3	24.7	4.5	Abla et al. 2011

8.3.3 Surface Activity

8.3.3.1 Contribution of the Alkyl Tip Length

Since the hydrophobic effect is the driving force for micelle formation, the hydrophobicity is usually inversely correlated with the CMC value (Tanford 1980). The addition of an ethyl group is indeed known to decrease ~ 5 – 15 times the CMC (Rosen 2004). This is observed for the *H*-surfactants presented in Table 8.2. For example, H_{12} Diglu (compound 3) has a CMC approximately 14 times lower than that of H_{10} Diglu (compound 1), and H_{12} Triglu (compound 4) has a CMC ~ 11 times lower than that of H_{10} Triglu (compound 2). However, when adding an ethyl tip at the end of the *F*-chain this general trend was not observed. The ethyl-ended *HF*-surfactants with branched polar heads present the unexpected behavior previously observed with the lactobionamide (Lebaupain et al. 2006), the maltoside (Polidori et al. 2006), and the TAC series (Barthelemy et al. 1999), in view of their CMCs. H_2F_6 Triglu (compound 11) exhibits very close CMC to F_6 Triglu (compound 8; ~ 0.7 and ~ 0.9 mM, respectively). The same trend was observed for H_2F_6 Diglu (compound 10) and F_6 Diglu (compound 6; ~ 0.3 and ~ 0.2 mM, respectively). The nonclassical behavior of the ethyl-ended *F*-surfactants was also observed on the second series of surfactants as the addition of an ethyl tip to F_6H_3 Diglu was found to decrease by only 2–3 times the CMC (~ 1.0 and ~ 0.4 mM for F_6H_3 Diglu and $H_2F_6H_3$ Diglu, compounds 13 and 16, respectively). The same trend was observed with the DigluM polar head (compounds 14 and 17).

On the contrary, the addition of a propyl tip was found to decrease the CMC by more than 10 times. Indeed, the CMC of $H_3F_6H_3$ Diglu (compound 18) does not reach 0.1 mM while F_6H_3 Diglu (compound 13) has a CMC of ~ 1.0 mM. The same order of magnitude was found when comparing $H_3F_6H_3$ DigluM and F_6H_3 DigluM bearing a slightly different polar head (compounds 19 and 14), and $H_3F_6H_3$ Triglu and F_6H_3 Triglu bearing three glucose groups (compounds 20 and 15).

The γ_{\max} of a given surfactant is mainly associated to the lateral interactions between the hydrophobic chains of the molecules adsorbed at the air/water interface (Adamson 1990), the greater the packing of surfactants at the interface, the lower the γ_{\max} . The high surface activity of *F*-surfactants is confirmed on the investigated surfactants, as most of them exhibit relatively low γ_{\max} , in the 20–30-mN/m range, when compared to their hydrogenated analogs in the 35–40-mN/m range. Both series of *HF*-surfactants exhibit a similar behavior when compared to their corresponding *F*-analogues. The addition of an ethyl or a propyl tip to the *F*-tail increases the γ_{\max} values by 6–8 for the Diglu compounds (compare compounds 6 and 10, 8 and 11, 9 and 12, 13 and 16) and 1–4 mN/m for the Triglu compounds (compounds 14 and 17) very likely a consequence of the unfavorable interactions between *F*- and *H*-segments that hamper the formation of a high compact film at the air/water interface.

A_{\min} gives information about the surface that each molecule needs to accommodate at the air/water interface. For the first series and for a given polar head, the

length and the nature of the tail have no significant effect on the A_{\min} , however with the following trend: $H_2F_6 > H_{12} > F_6 > H_{10}$ for the Diglu and Triglu derivatives. It is also interesting to note that the DigluM polar head leads to $\sim 20\%$ smaller area than Diglu derivatives.

8.3.3.2 Contribution of the Hydrocarbon Spacer

The hydrocarbon spacer between the polar head and the hydrophobic tail is also supposed to play an important role in the physical-chemical properties of a surfactant, because it affects the flexibility of the surfactant molecule. The sulfur atom inserted within the spacer of the first series could be considered as a hydrophobic unit. This assertion is based on the compared Hansch partition constant between octanol and water of SCH_3 and OCH_3 (Leo et al. 1971; i.e., $\pi = 0.45$ and -0.47 , respectively), and the poor H-bond acceptor properties of sulfur (Allen et al. 1997). Menger and colleagues have already discussed the role of a thioether bond and its effect on the CMC (Menger and Shi 2006; Lundberg et al. 2008), its magnitude depending essentially on its position within the hydrophobic tail, the farther to the polar head group, the stronger the increase of the CMC. For instance, the CMC of *n*-octyl- β -D-glucopyranoside is in the 20–25-mM range whereas that of its thio analog, *n*-octyl- β -D-thioglucopyranoside, is 9 mM, i.e., close to that of *n*-nonyl- β -D-glucopyranoside (6.5 mM; Wenk and Seelig 1997). This demonstrates that a sulfur atom, when grafted directly to the polar head, contributes almost like a CH_2 group.

The second series of *F*-surfactants has a smaller linker, with one out the sulfur atom and one methylene group less with respect to the first one. We found that the CMC of F_6H_3 Diglu (compound 13, second series) is about four times larger than that of F_6 Diglu (compound 6, first series) with CMC of ~ 1 and ~ 0.2 mM, respectively. Similarly, the CMC of F_6H_3 Triglu is about three times larger than that of F_6 Triglu (compounds 15 and 8, with CMC of ~ 3 and ~ 1 mM, respectively). The same trend is also observed for the DigluM compounds (compounds 14 and 7, with CMC of ~ 0.8 and ~ 0.4 mM, respectively), however with a smaller (~ 2) magnitude. Compared to the 12–14 times increase of the CMC observed when adding two methylene groups (see the CMCs of compounds 1–4), the addition of $S-CH_2$, with the sulfur atom at three atoms distance from the polar head group, increases the CMC only slightly. In addition, and as underlined above, H_2F_6 -surfactants behave atypically, as H_2F_6 Diglu (compound 10) and $H_2F_6H_3$ Diglu (compound 16) exhibit similar CMC values (0.35 and 0.37 mM, respectively).

As we discussed above, the nature of the chain has no strong effect on the A_{\min} values within the same series. However, when comparing the two series, higher molecular area was recorded for the second one with that of F_6H_3 Diglu (compound 13) being 25% higher than that of F_6 Diglu (compound 6). This difference might be explained by the kink induced by the presence of a sulfur atom within the chain of the first series, which would result in a better stacking of the molecules at the air–water interface.

8.3.3.3 Contribution of the Polar Head: General Considerations

As a general trend, the CMC value of a surfactant decreases when increasing its global hydrophobicity, a manner to minimize the unfavorable interactions with the aqueous phase, facilitating thereby micellization. The hydrophobicity of a surfactant is expected to decrease when increasing the volume of the polar head, such as the number of saccharides for instance. The presence of hydroxyl groups in the polar head, in general, favors the formation of H-bonding, which in turn may delay micelle formation (Venkatesan et al. 1994). However, it appears that the volume distribution as well as the stereochemistry of the saccharides have a strong influence on the micellization (Soderberg et al. 1995; Garofalakis et al. 2000). Indeed, the number of saccharide moieties may affect micellization in a rather limited way. For instance, dodecyl alkyl chain-based derivatives with maltose-based and polydextrane polar heads (2 and 9 linear glucose moieties, respectively) were found to exhibit very similar CMC values (Zhang and Marchant 1996).

8.3.3.4 Increasing the Number of Glucose Groups in the Polar Head

With the *F*-surfactants of the first series, however, an increase in the number of glucose groups on the polar head results in an increase of the CMC. For the given F_6 -chain (compounds 5, 6, 8), adding one and two glucose group increased 2 and 8 times the CMC, respectively. The branched arrangement of the polar head favors the formation of H-bonding with water, and as a consequence may result in higher water solubility. Clearly the $\log k'_w$ value of F_6 Monoglu was higher than that of F_6 Triglu (5.2 and 4.7, respectively), confirming the influence of the overall hydrophobicity on micellization. This same general behavior was found for H_{10^-} , H_{12^-} , and H_2F_6 - surfactants described in Table 8.2.

In contrast, the CMC of propyl-ended surfactants is not significantly affected by the polar head size: $H_3F_6H_3$ Diglu (compound 16) and $H_3F_6H_3$ Triglu (compound 20) exhibit similar CMCs of 0.08 and 0.07 mM, respectively, their hydrophobic character being comparable. A similar observation was made on a series of galactosylated surfactants having long fluorinated chain (C_8F_{17} ; Polidori et al. 1994), and this was further confirmed with a series of galactosylated ethyl-ended *HF*-surfactants (Barthelemy et al. 1999). This indicates that the effect of adding a sugar moiety is lower when the surfactant has a low CMC.

For a given hydrophobic chain, increasing the polar head size, demonstrated by higher A_{\min} values, is expected to induce higher γ_{\max} values, as a consequence of a greater gap between surfactant molecules and weaker interactions between them at the air/water interface. This rule is in a perfect agreement with the results obtained with *F*-surfactants and *H*-surfactants. However, no significant changes were observed with *HF*-surfactants (Table 8.2). This proves once again the peculiar behavior of these compounds at the air/water interface.

8.3.3.5 Difference Between Glucose and Galactose-Based Polar Heads

Although glucose and galactose are structural isomers, they exhibit different physical-chemical properties such as water solubility. The effect of the nature of the sugar moiety grafted onto the polar head is underlined in the difference of the CMC of the Triglu and Trigal surfactants (compounds 8 and 9, 11 and 12). The galactosylated derivatives exhibit CMCs that are half that of the glucosylated ones. This is reasonably explained by the lower solubility of the galactose compared to the glucose, a consequence of the stereochemistry of the hydroxyl groups. Compared to galactose where the hydroxyl group in position 4 is axial, that of glucose is equatorial which is more favorable to the formation of H-bonding and therefore delays the micelle formation.

8.3.3.6 Conclusion on the Surface Activity

F-surfactants exhibit a lower CMC, a decreased γ_{\max} , and a slightly larger A_{\min} than the hydrogenated homologs. The peculiar behavior of *HF*-surfactants with the ethyl tip—same CMC, similar $\log k'_w$, but increased γ_{\max} and slightly larger A_{\min} —is most likely related to: (1) a poor packing induced by unfavorable van der Waals interactions between the hydrocarbon and fluorocarbon segments of these hybrid surfactants (Lebaupain et al. 2006; Abla et al. 2008) and (2) the acidic character of the methylene adjacent to the *F*-core (Kirsch 2004), as a consequence of the high electron-withdrawing effect of the fluorine atoms. This may favor the formation of *H*-bonds with water molecules which in turn may increase the affinity of the surfactant for the aqueous phase and lessen at the same time the hydrophobicity of H_2F_6 -surfactants. As a result of the addition of these two effects, micellization is delayed. This is in agreement with the findings by Eastoe et al. (2001) who demonstrated that, upon replacing the terminal CF_3 group of a nonionic fluorinated surfactant by an $H-CF_2$ group, the CMC increases by a factor of ~ 4 .

We observed a different scenario with the propyl-ended derivatives exhibiting more pronounced hydrophobicity as demonstrated by higher $\log k'_w$ (e.g., 5.2 for $H_3F_6H_3$ Diglu and 4.5 for F_6H_3 Diglu), and favored micellization compared with their shorter perfluorinated analogs. This suggests that the acidity of the methylene group vicinal to the fluorinated core is largely compensated by the additional CH_3-CH_2 group, while in the case of the ethyl-ended compounds the additional CH_3 is not sufficient.

8.3.4 Shape and Size of the Micelles

8.3.4.1 Calculation of the Critical Packing Parameter

When calculating the CPP introduced in Sect. 2.4, the distinction between hydrophobic tail and polar head calls for caution. Indeed, the first methylene group of the tail, which is attached to the polar head group of surfactant, lies within the

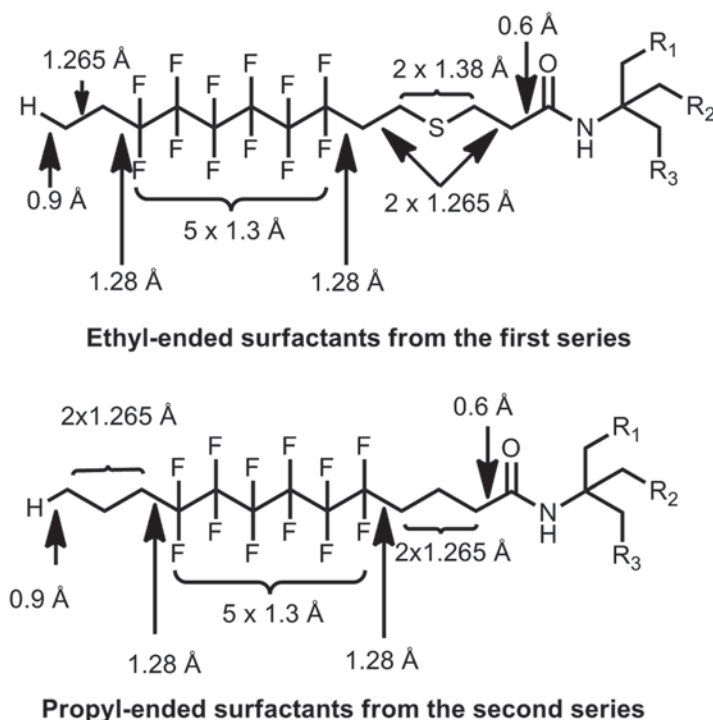


Fig. 8.6 Estimation of the maximum extended length of *HF*-surfactants

hydration sphere of the head (Rosen 2004). Thus, it does not possess any hydrophobic character and should not be considered part of the tail. In the present work, we assume that the hydration sphere of the glucose-based polar head includes the TRIS moiety, the amide bond being the frontier with the tail. Estimation of the length and volume of hydrophobic chains (Table 8.3) is done thanks to values from the literature. While the length of the $\text{CH}_2\text{-CH}_2$ bond is 1.265 Å (Tanford 1980), those of $\text{CF}_2\text{-CF}_2$ and $\text{CF}_2\text{-CH}_2$ are 1.3 and 1.28 Å, respectively (Lo Nostro and Chen 1993), and 1.1 Å and 0.9 Å are the respective contributions of the terminal C-F and C-H bonds (Fig. 8.6). According to the C-S bond length and C-S-C angle value observed in dimethyl sulfide, we derived a contribution of 1.38 Å for the C-S bond within the chain. Finally, we added half a bond length (0.6 Å) corresponding to the link between the tail and the head. Hydrophobic chains are not fully extended within the micelles, so that the effective length of the chain l_0 is shorter than the maximum extended length. We consider for the hydrocarbon tail the usually employed corrective factor (~ 0.75). Due to the greater stiffness (Srinivasan and Blankschtein 2005) and the reduced conformational freedom of the helical structure of fluorocarbon chains (Kirsch 2004), we can assume that the approximated extended lengths of *F*- and *HF*- chains give a good approximation of their effective lengths. Steric volume are calculated with the following values:

54.3 and 26.9 Å³, respectively, for a CH₃ and a CH₂ unit (Tanford 1980), and 86 and 41 Å³, respectively, for a CF₃ and a CF₂ (Lo Nostro and Chen 1993). That of a sulfur atom was taken as 30 Å³.

The CPP values were calculated according to the geometrical parameters listed above and reported, as the results, in Table 8.3.

8.3.4.2 Experimentally determined Size and Shape of the Micelles

The experimental hydrodynamic diameters determined by DLS, and aggregation numbers and shapes determined from AUC and SANS, are reported in Table 8.3. The DLS values for the unpublished derivatives were determined at concentrations several times above their CMC. The results clearly show the impact of the volume of the surfactant polar head on the size of the micelles formed in aqueous solutions. All surfactants bearing Diglu or Triglu (or Trigl) polar head groups self-organize into small and well-defined globular micelles, with apparent hydrodynamic diameter (D_H) of 5–7 nm and aggregation numbers below 70. In addition, for both series, surfactants with three sugar groups in the polar head formed slightly smaller micelles than those with two sugar groups (compounds 4, 8, 11, 15, and 20 vs. 3, 6, 10, 13, and 18). The surfactant bearing a single glucose moiety forms large rod-like micelles with size increasing with concentration, and large D_H and aggregation number (above 25 nm and 1,000, respectively, for concentration of some mg/mL).

8.3.4.3 Validity of the CPP Model for Branched Fluorinated Surfactants

Within each family and for a given tail, the D_H of micelles decreases when the molecular area (A_{\min}) increases, which is in perfect agreement with an increase in the volume of the polar head. This is obvious for the *F*-surfactants of the first series, where the small A_{\min} value for F₆Monoglu corresponds to a D_H distribution significantly shifted to large values, as compared to F₆Diglu and F₆Triglu (Fig. 8.7). Interestingly, when looking at the two series, a good correlation between A_{\min} and the CPP values was observed (Fig. 8.8). Among the 17 derivatives plotted in Fig. 8.7, 15 have A_{\min} values comprised between ~90 and 140 Å² and D_H comprised between 5 and 7 nm, which is in good agreement with their CPP values below or close to 1/3, indicative of globular micelles. F₆Monoglu, for which the CPP of 0.54 would correspond to the formation bilayers or cylinders, self-assembles into rod-like micelles. Only the F₆DigluM—forming globular compact micelles—is out of the frame as its CPP value of 0.38 would predict elongated micelles in relation to its relative small polar head (~75 Å²).

Table 8.3 Aggregate geometry of *F*-surfactants with branched polar heads

Series	No.	Surfactant Name	MW (g/mol)	l_0 (Å)	v_0 (Å ³)	A_{min} (Å ²)	CPP	Predicted aggregate geometry ^a	D_H (nm) ^b	N_{agg}	Experimental aggregate geometry ^c	Ref
First	1	H ₁₀ Diglu	673.3	16.9	377.1	92	0.24	Globular	5.4	ND	ND	Abla et al. 2008
	2	H ₁₀ Triglu	835.4			108	0.21	Globular	4.5	ND	ND	
	3	H ₁₂ Diglu	701.4	19.5	430.9	103	0.21	Globular	5.9	ND	ND	
	4	H ₁₂ Triglu	863.4			125	0.18	Globular	5.2	ND	ND	
	5	F ₆ Monoglu	717.1	14.8	425.5	53	0.54	Bilayer	25.3	3600 (4 g/L) ^d 3200 (3 g/L) ^e	Cylindrical	Abla et al. 2008 Breyton et al. 2009
	6	F ₆ Diglu	879.1			98	0.29	Globular	5.8	68 ^{d,e}	Globular	
	7	F ₆ DigluM	863.2			75	0.38	Cylindrical	5.5	51 ^d 74 ^e	Globular	Abla et al. 2012 and this work
	8	F ₆ Triglu	1,041.6			114	0.25	Globular	5.1	35 ^{d,e}	Globular	Breyton et al. 2013b Abla et al. 2008 Breyton et al. 2009
	9	F ₆ Trigal	1,041.6			115	0.25	Globular	5.2	ND	ND	This work
	10	H ₃ F ₆ Diglu	889.2	17.1	461.7	105	0.26	Globular	6.5	62 ^{d,e}	Globular	Abla et al. 2008
	11	H ₂ F ₆ Triglu	1,051.8			132	0.20	Globular	5.1	ND	ND	Breyton et al. 2009
	12	H ₂ F ₆ Trigal	1,051.8			140	0.19	Globular	5.1	ND	ND	This work

Table 8.3 (continued)

Series	No.	Surfactant Name	MW (g/mol)	l_0 (Å)	v_0 (Å ³)	A_{\min} (Å ²)	CPP	Predicted aggregate geometry ^a	D_H (nm) ^b	N_{agg}	Experimental aggregate geometry ^c	Ref
Second	13	F ₆ H ₃ Di glu	833.2	12.0	371.1	126	0.25	Globular	6.6	ND	ND	Abla et al. 2011
	14	F ₆ H ₃ Di gluM	817.2			98	0.31	Globular	5.1	48 ^d	Globular	Abla et al. in preparation
	15	F ₆ H ₃ Tri glu	995.2			ND	ND	ND	5.7	ND	ND	Abla et al. 2011
	16	H ₂ F ₆ H ₃ Di glu	843.2	14.3	408	ND	ND	ND	5.8	ND	ND	
	17	H ₂ F ₆ H ₃ Di- gluM	827.6			1227	0.22	Globular	5.9	ND	ND	This work
	18	H ₃ F ₆ H ₃ Di glu	857.3	15.6	434.8	122	0.23	Globular	6	ND	ND	Abla et al. 2011
	19	H ₃ F ₆ H ₃ Di- gluM	841.3			93	0.29	Globular	6	70 ^d	Globular	Abla et al. in preparation
	20	H ₃ F ₆ H ₃ Tri glu	1019.3			ND	ND	Globular	5.5	ND	ND	Abla et al. 2011

ND not determined

^a Predicted geometry is from CPP as described in Sect. 1.4

^b Distribution by volume

^c From the combination of DL-S, AUC, and SANS or SAXS data

^d From AUC

^e From SANS

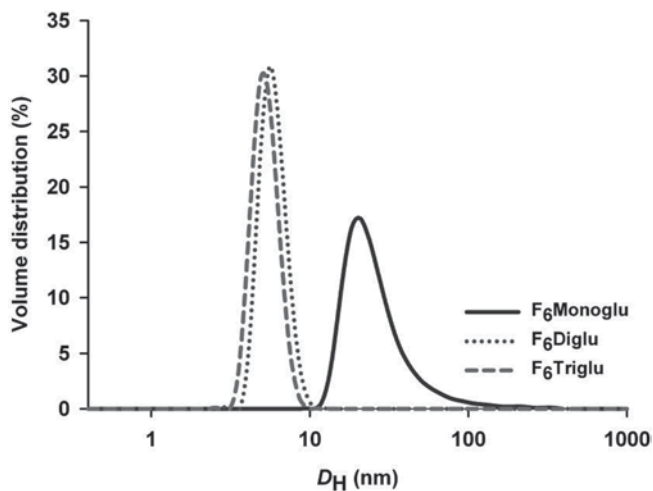
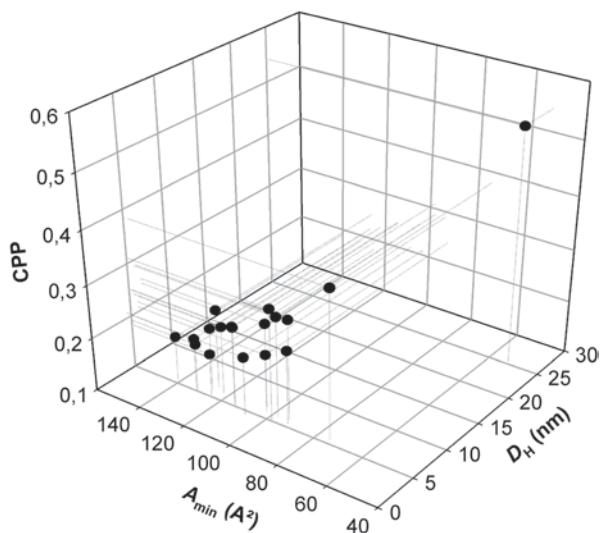


Fig. 8.7 Hydrodynamic diameter distribution statistical plot by volume from DLS for F₆Monoglu, F₆Diglu, and F₆Triglu at 4 mM in pure water

Fig. 8.8 3D correlation between CPP, A_{min} , and D_H for the two series of *F*-surfactants



8.4 Biochemical Applications of Fluorinated Surfactants

A detailed description of the application of *F*-surfactants has already been reviewed in Breyton et al. 2010 and Popot 2010. These included stabilizing MPs in solution (Breyton et al. 2004; Talbot et al. 2009), in vitro synthesis (Park et al. 2007), inserting proteins into preformed membranes (Posokhov et al. 2008 and references therein), and MP refolding (Lebaupain 2007). The comparison of F₆ Monoglu, F₆ Diglu, and F₆ Triglu showed clearly the inability of the surfactant bearing three sugar groups to stabilize the *b₆f*. The same inability was noted when using H₂F₆ Triglu, F₆ Trigal, and H₂F₆ Trigal. We will here review what has been published since 2010, with the emergence of a particularly interesting application in SANS of membrane complexes, and with the use of *F*-surfactants for single particle analysis of MPs by electron microscopy.

8.4.1 MP Stabilization

The stabilization effect of *F*-surfactants for biochemical and biophysical characterization of MPs was demonstrated on the two human membrane receptors of the Sonic hedgehog signaling pathway, the 12-transmembrane-helix plasma membrane Patched (hPtc; Joubert et al. 2010), and the G protein-coupled receptor (GPCR) Smoothed (hSmo; Nehmé et al. 2010). The two proteins were overexpressed independently in the plasma membrane of the yeast *Saccharomyces cerevisiae*, purified and characterized. The stability of the proteins was monitored after membrane solubilization in *n*-dodecyl- β -D-maltopyranoside (DDM) and surfactant exchanged during the purification, by the ability of the membrane receptor to bind to its ligand by surface plasmon resonance technology. hPtc and hSmo showed increased stability over time when transferred in F₆- and F₈-TAC, respectively. In particular, whereas the melting temperature of hSmo in F₆-TAC was decreased by 7 °C when compared to that in DDM, it was increased by 5 °C when in F₈-TAC. Whereas the stabilizing effect of F₈-TAC is very encouraging for further structural studies, the opposite effect of F₆-TAC is more surprising, as it proved stabilizing for the ATP-synthase dimer (Talbot et al. 2009) and for other proteins. It was explained by the authors by a better ability of the longer chain surfactant to shield the hydrophobic domain of the protein. This points to the necessity of developing different surfactant molecules, as there does not exist a universal one that would suit all MPs. Interestingly, synergic stabilizing effects are observed with a point mutation in the third helix of hSmo and solubilization of the protein in F₈-TAC, increasing the melting temperature by 13 °C with respect to that of the wild type (WT) in DDM. The authors note, however, the difficulty to compare sur-

face plasmon resonance signals for the protein solubilized in different surfactant because of difference (dn/dc) for DDM and *F*-surfactants (Nehmé et al. 2010). Another interesting study is that of Cho et al., who have grafted fluorine-containing benzene derivatives onto a maltose-neopentyl glycol (MNG, presented in Sect. 4) scaffold. This F4-MNG molecule seems to be easily synthesized in high yields, and confers enhanced stability on *Rhodobacter capsulatus* supercomplexes compared to conventional detergents and nonfluorinated MNG amphiphiles (Cho et al. 2013). Furthermore, we have shown the stabilizing properties of F₆DigluM (Abla et al. 2012) and derived molecules (Abla et al. in preparation) on our test protein, bacteriorhodopsin.

8.4.2 Cell-Free Synthesis

Cell-free synthesis is an alternative to the production of proteins. It couples, in the test tube, transcription and translation reactions, allowing to circumvent the toxicity and/or the crowding of the membrane insertion machinery *in vivo*, which often leads to poor MP overexpression yields (Bernhard and Tozawa 2013). The use of different *F*-surfactants to accommodate and solubilize the newly synthesized protein was further investigated on different protein systems. One advantage of this technique is that it does not include a surfactant-exchange step for further biochemical and biophysical characterization. (H₂)F₆-TAC, F₈-TAC, and (H₂)F₆Diglu proved compatible with the cell-free synthesis machinery, and allowed the production of soluble bacterial proteins as the leader peptidase and bacteriorhodopsin (Breyton et al. 2009; Park et al. 2010). Cell-free synthesis of the mitochondrial uncoupling protein, *i.e.*, a eukaryotic protein, was also investigated, screening a wide range of conventional and noncommercial surfactants. The best results, both in synthesis yield and solubilization of the protein, were obtained when the synthesis reaction was performed in the presence of the *F*-surfactant F₈-TAC combined with small amounts of cardiolipin. This latter additive is an inner mitochondrial membrane lipid, known to interact with and stabilize a structural homologue of this protein (Nury et al. 2005). The purified protein showed to be monomeric and correctly folded with a melting temperature as high as 60 °C, in yields up to 0.8 mg/mL of reaction mixture, *i.e.*, entirely compatible with structural studies (Blesneac et al. 2012). Success was also achieved with the adenosine diphosphate (ADP)/adenosine triphosphate (ATP) carrier, in particular to obtain deuterated protein (Breyton et al. 2013b).

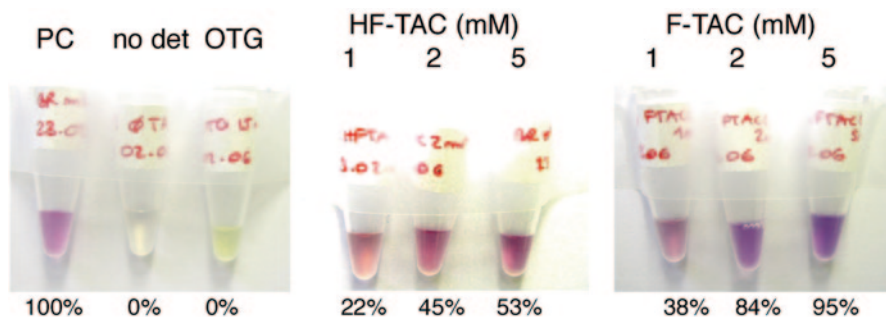


Fig. 8.9 Refolding of bacteriorhodopsin in the presence of the F_6 -TAC or H_2F_6 -TAC. Purple membrane was denatured by addition of 5% SDS, releasing retinal (yellow). The preparation was then supplemented with either lipids (PC), 15-mM octylthioglucoiside (OTG), no surfactant (no det), or F_6 - or H_2F_6 -TAC at the indicated concentration. Refolding is initiated by precipitating the dodecyl-sulfate with K^+ ions, followed by dialysis (Popot et al. 1987). It is indicated by the reappearance of the purple color. Spectrum analysis of the samples confirms the results and allows quantification of the extent of renaturation (% below the tubes). (Lebaupain 2007)

8.4.3 Refolding of MPs

A very promising application of F -surfactants is the refolding of MPs. We have demonstrated the proof of concept on the refolding of the outer MP A (OmpA), a β -barrel protein, and bacteriorhodopsin, a seven-helix bundle protein that covalently binds a retinal molecule. Retinal is a convenient reporter of protein folding, as it displays a purple color when it is correctly inserted in the protein (maximum absorption spectrum 555–570 nm), and yellow when it is free in the surfactant micelle (maximum 380 nm). Refolding yields of up to 95% were obtained with F_6 -TAC (Fig. 8.9). The protein however is not monodisperse, as a sucrose gradient analysis shows that monomers, dimers, and higher molecular weight oligomers are present as well. We suspect that purple membrane patches are reconstituted. It is interesting to note that no refolding at all could be observed with the F_8 -TAC. Hemifluorinated anionic surfactants were also used to promote the refolding of the *soluble* protein bovine carbonic anhydrase. The rationale to use partially fluorinated surfactants was that they would have high enough affinity for hydrophobic residues exposed in an unfolded state to prevent protein aggregation, yet have a low enough affinity that they can be readily displaced by dialysis. Results show that up to 95% refolding, after thermal or chemical denaturation, can be obtained with hemifluorinated anionic surfactants (Singh and Flowers 2010).

8.4.4 *F*-surfactants as “chaperones” for Protein Insertion in Membrane

Ladokhin’s group has been investigating the thermodynamics of membrane insertion of model proteins with *F*-surfactants (F_6 - and F_8 - and H_2F_6 -TAC) as “chaperones.” Indeed, their poor miscibility with membranes makes them particularly useful for thermodynamic studies of membrane insertion under equilibrium conditions.

The first studies were performed using diphtheria toxin T-domain, which at neutral pH exists as a soluble globular protein, and upon acidification is converted into a membrane-inserting form (see Posokhov et al. 2008 and references therein). The effects of sub-CMC concentrations of several surfactants on the thermal denaturation of the T-domain was recently investigated, in order to identify compounds that would show the lowest perturbation on the unfolding of the soluble form of the T-domain. *F*-surfactants, especially F_6 -TAC, stand out as the best candidate for various thermodynamic studies of MPs (Kyrychenko et al. 2012b).

This general “chaperone” approach was further applied to measure, using F_6 -TAC, the free energy of insertion of a single transmembrane helix, the fluorescently labeled WALP peptides. Fluorescence correlation spectroscopy was used to measure very precisely the fraction of peptide bound to the membrane, and thus to measure the free energy of membrane insertion of different peptides, WALP23 and WALP27, which differ only in the insertion of LALA in the hydrophobic core of the latter. Their comparison allowed to determine the partitioning free energy of this small sequence from the aqueous environment into the hydrocarbon core in the context of a continuous helical segment. This, in turn, allowed to calculate the unfavorable contribution of the backbone ($\Delta G = +2.2$ kcal/mol), which very closely coincides with previous estimates (see the discussion in Kyrychenko et al. 2012a). Thus, the use of *F*-surfactants allowed a very thorough thermodynamical study of membrane insertion of proteins under equilibrium conditions, that were otherwise forbidden by the insolubility of hydrophobic peptides/protein and the solubilization of the membrane when detergents were used to solubilized the peptides.

Quite different conclusions were found by Raychaudhury et al. (2011). Indeed, they observed that *F*-surfactants (F_6 -Foscholine, F_6H_2 maltoiside, and F_6 -TAC presented in Fig. 8.2) prevent protein insertion into preformed membranes rather than promoting it. Three pore-forming proteins were tested: staphylococcal α -hemolysin, which forms a pore when it heptamerizes in the membrane; MspA, an octameric porin from *Mycobacterium smegmatis*; and Kcv, a tetrameric potassium channel from the chlorella virus, all three proteins purified in sodium dodecyl sulfate (SDS). The tools to quantify insertion of the membrane were the leakage of a self-quenching dye sequestered in the liposome (macroscopic obser-

vation), and patch-clamp measurements in planar lipid bilayers (single-channel recording). This “membrane insertion inhibitory ability” of *F*-surfactants was taken as an advantage in view of patch-clamp measurements in planar lipid bilayer, which needs very few insertion events. The authors propose that the proteins are sequestered in the *F*-surfactant phase when these latter are added at concentrations higher than their CMC. Transmission electron microscopy (TEM) images of F₆-Foscholine are presented that show vesicle-like aggregates, in which the protein would be inserted and “sequestered.” Indeed, we have shown that F₆H₃ maltoside forms large, probably cylindrical micelles and that protein–surfactant complexes are heterogeneous (Polidori et al. 2006). However, this is not true for F₆-TAC, which forms small aggregates (Dupont et al. 2003), making this sequestering hypothesis weaker. The experimental setup between the experiments of Raychaudhuri et al. (2011) and the Ladokhin’s team differs: In the latter, the proteins are solubilized in the *F*-surfactant prior to being mixed with the membrane, whereas the proteins are solubilized in SDS and the *F*-surfactant added after in the former case.

This work of Raychaudhuri et al. (2011) confirms previous observations (Rodnin et al. 2008 for F₆-TAC) by showing that the three investigated *F*-surfactants did not permeate membranes, neither by hemolytic assay, liposome leakage, nor patch-clamp measurements. F₆Diglu however has been shown to induce leakage on liposomes in conditions where F₆-TAC does not (Kryuchenko et al. 2012b).

8.4.5 Structure of MPs from Small-Angle Scattering

A very promising application of *F*-surfactants, demonstrated for F₆DigluM, is their use for structural investigation of proteins within a MP complex using SANS. SANS combined to deuterium labeling of specific proteins within a multi-subunit/protein complex and with variation of the D₂O % in the experimental buffer, is a particularly powerful technique to monitor conformational changes undergone by each protein within the complex. Indeed, at a defined D₂O % of the buffer (~45 % D₂O), which is defined as the contrast match point, the scattering of hydrogenated proteins is cancelled, whereas that of deuterated proteins is highlighted. In a MP sample, an additional component needs to be considered: free micelles and bound detergent (Breyton et al. 2013b). We have shown that the contrast match point of micelles of F₆-Diglu (Abla et al. 2008; Breyton et al. 2009) and F₆DigluM (Abla et al. 2012; Breyton et al. 2013b) is very close to that of hydrogenated proteins. Thus, when an MP complex is solubilized in these compounds, both the surfactant contribution and that of hydrogenated proteins can be cancelled in the same buffer conditions, allowing structural investigation of the deuterated protein within the complex. The feasibility of this approach has been experimentally validated

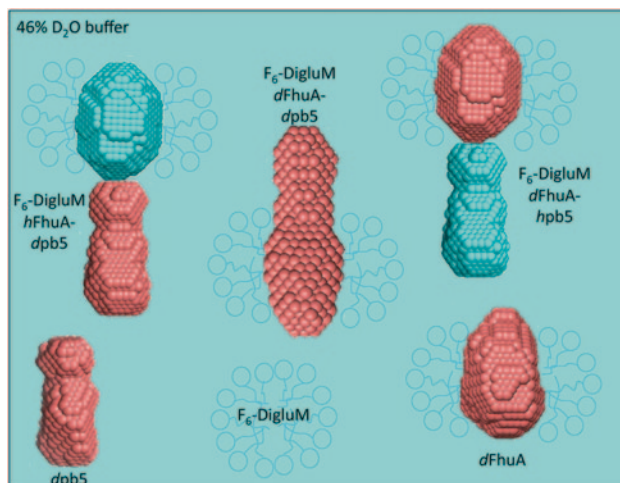


Fig. 8.10 SANS contrast strategy to highlight specific proteins of the FhuA-pb5-detergent complexes: deuterated proteins (*d* protein) are represented in *salmon pink*. Hydrogenated proteins (*h* protein) and the F₆DigluM micelles are represented in *cyan*, as is the 46% D₂O buffer in which they are contrast matched

in the case of conformational changes occurring upon formation of the complex between pb5, the receptor-binding protein of phage T5, and its receptor FhuA, an *E. coli* outer-membrane protein. The structure of each partner in the context of the complex was determined and compared to its structure in solution (Fig. 8.10). We could show that upon formation of the complex, neither FhuA nor pb5 undergoes large conformational changes (Breyton et al. 2013a). This was an unexpected result that constrains the mode of transmission of the binding information to the rest of the phage.

8.4.6 Structure of MPs from Electron Microscopy

Another very interesting application of *F*-surfactants, based on their inability to mix with detergents, is their use to promote the concentration and orientation, and possibly the two-dimensional (2D) crystallization of MPs for imaging by electron microscopy. The proof of concept was made using partially fluorinated lipids (Lebeau et al. 2001): The air–water interface is saturated with an *F*-surfactant, whose polar head bears a nitrilotriacetic (NTA) Ni²⁺-chelating group. The detergent-solubilized His-tagged protein is then injected in the subphase. The *F*-surfactant monolayer at the air–water interface is not perturbed by the presence of the de-

tergent, even at concentrations over its CMC. His-tag binds to the Ni-NTA head group, and the protein is thus concentrated and oriented beneath the *F*-surfactant monolayer. The addition of detergent-solubilized lipids followed by that of Bio-Beads in the subphase allows the protein to be further reconstituted in a bilayer that can form 2D crystals. The bilayer tethered to the monolayer is then transferred onto an electron microscopy grid and imaged. This technique was recently applied using a new *F*-surfactant (Dauvergne et al. 2008) for the investigation of a megadalton MP complex, the ATP-sensitive potassium channel regulatory sulfonylurea receptor (Fotinou et al. 2013). Single-particle electron microscopy of the purified subunit tethered to the *F*-surfactant monolayer reveals that it assembles as a tetramer of four subunits surrounding a central hole. Homology modeling showed that this latter can accommodate the potassium channel tetramer (Fotinou et al. 2013).

8.4.7 Conclusion on the Biochemical Properties of Fluorinated Surfactants

F-surfactants stand up to their expectations, and have proven very useful for different types of applications, for which classical detergents fail. One application that we have not seriously investigated is the crystallization of MPs. Two options are available: Lipidic mesophase phase crystallization, where the surfactant-solubilized protein is added to a lipidic cubic or sponge phase, is the first one. The protein inserts in the membrane-like environment of the lipidic mesophase, which stabilizes sensitive proteins and leads to type I crystal packing (Caffrey 2009). In this crystal packing, crystal contacts are formed between both polar and nonpolar parts of the protein (Caffrey 2009). As the protein inserts in the mesophase, the surfactant is not present in the crystal. We have seen above that *F*-surfactants are stabilizing toward fragile MPs, and a very convenient medium for MP insertion into a lipid bilayer. Lipid mesophase crystallization using *F*-surfactants-solubilized proteins would therefore be quite appropriate. The other—and more commonly used—alternative is the crystallization of the surfactant-protein complex, using the vapor-diffusion method. In this case, type II crystal packing, where crystal contacts are formed exclusively between the polar parts of the protein, itself being surrounded by the surfactant belt, is generally favored (Ostermeier and Michel 1997). Thus, small micelle-forming detergents are considered as more favorable to promote crystal contacts (see Sect. 8.5). Unfortunately, these are also usually the more denaturing ones. The use of *F*-surfactants could be very interesting to investigate: because protein-protein and protein-lipid interactions are favored over protein-*F*-surfactants ones; in vapor-diffusion crystallization conditions, type I crystals could be favored. Having more crystal contacts, type I crystals are usually better ordered and diffracts at higher resolution. The addition of lipids to MP-fluorinated surfac-

tant complex could be necessary, or not, given that lipids sometimes co-purify with proteins (Palsdottir and Hunte 2004).

8.5 Sugar-Based Nonconventional Amphiphiles for MP Study: An Overview

8.5.1 Introduction to Table 8.4

To keep this section reasonable in length, we will focus on the molecules developed in the past 5 years and that bear a sugar head group. Indeed, these past years have witnessed an explosion in the development of new amphiphiles for the manipulation of MPs, and the large majority bears a sugar-based head group. This mainly stems from the large popularity of DDM among biochemists, as it is indeed one of the most stabilizing detergents in MP biochemistry and the most used one for MP crystallization. However, DDM is often not good enough for proteins that are highly dynamic, in particular GPCR that are in equilibrium between multiple conformations. These latter proteins are unstable, denaturing, and precipitate rapidly when extracted from the membrane and manipulated in detergent. Thus emerged the need to develop alternative detergents/amphiphiles to manipulate more fragile proteins (see also reviews, Tate and Schertler 2009; Bill et al. 2011; Sonoda et al. 2011; Zhang et al. 2011; Kang et al. 2011). Tables 8.4 and 8.5 presents the different classes of new amphiphiles, the rationale for their development and the structure of the lead compound. The main physical-chemical characterization of the micelle is also listed when available (CMC , N_{agg} , R_h , and R_g), together with the capacity of the amphiphile to solubilize biological membranes and stabilize MPs. When explored, success for specific applications is mentioned, in particular crystallization. In the following, please refer to Tables 8.4 and 8.5 for references and chemical formulas.

8.5.2 Rational for Developing Different Amphiphile Families

As regards the rationale of the development of new amphiphiles, we can distinguish different motivations, which are not exclusive to one another:

1. One of the hypotheses for the instability of solubilized MPs is the highly flexible nature of the alkyl chain of the detergent. *F*-surfactants, cyclic-maltosides, steroid-based maltosides, and tripods addressed this issue by rendering the hydrophobic moiety of the molecule more rigid: This was attained either by substituting hydrogen to fluor atoms (*F*-surfactants, see earlier sections), incor-

porating cycles (phenyl and/or cyclohexyl, and steroids, producing cyclic-maltosides, facial amphiphiles, steroid-based di-maltosides, and tripods), branching the hydrophobic moiety with three short chains (tripods), or by anchoring the hydrophobic chain at both ends in the Bola-maltosides. Dodecyl-trehalose falls in this category if we consider the head group that has a greater rigidity than the maltoside group.

2. Another rationale is the mimic of lipids able to stabilize solubilised MPs. Maltoside-neopentyl glycols (MNG; and to a lesser extent glucoside derivatives) indeed bear two hydrophobic chains of equal length, and the branched maltosides bear a long main chain and a shorter branch, mimicking the departure of a second alkyl chain. These latter molecules were designed to prevent water penetration into the micelle, increasing the hydrophobicity in the interior of the micelle, thereby stabilizing the protein–detergent complex (PDC). The design of Bola-maltosides was inspired by the lipids found in the membranes of extremophiles. In the native membrane, cholesterol interacts with many eukaryotic proteins, including GPCRs, and greatly stabilizes them in mixed detergent micelles. However, cholesterol and its analogue cholesteryl hemisuccinate are poorly soluble and increase the size of the detergent-mixed micelles. Mimicking cholesterol was the rationale for the development of the steroid-based amphiphiles, which share with cholesterol the steroid moiety. This is the case of CHOBIMALT, the steroid di-maltosides and the facial-amphiphiles presented in Table 8.4.
3. When the ultimate goal is the crystallization of the PDC, the aim is to reduce the size of the micelle. Indeed, the smaller the detergent belt around the hydrophobic domain of the protein, the more the possibilities of crystal contacts between the hydrophilic parts of the protein. Shortening the alkyl chain length of the detergent can straightforwardly reduce the micelle size, but corresponds to more denaturing detergent. Branching multiple hydrophobic short tails or rigidification of the molecule are issues to obtain small, nondenaturing micelles. This motivated the development of the glucoside neopentyl glycol (GNG) with two short alkyl chains. Dodecyl-trehalose has, in addition to an alkyl chain, a rigid head group—and different possibilities of coupling them. The presence of a small, branched chain in branched maltosides allows the main chain to be shorter while keeping the overall hydrophobicity of the molecule. Tripods with two short alkyl chains (four carbons) and a third phenyl one are expected to form small micelles. Finally, the facial amphiphiles, which have much smaller aggregation number than conventional detergents, were also designed with the aim of forming smaller PDC. In all cases, the strategy proved successful, and fragile MPs were stabilized, as well as or, more often much more than in DDM or other conventional detergents. As for conventional detergents, it is noted however that proteins are in general more stable with longer alkyl chains: MNG is better than GNG, which is still as good as or better than DDM. When tested, these new amphiphiles also stabilize MPs when

Table 8.4 Sugar-based non-conventional amphiphiles for membrane protein study. *C: commercial, NC: non-commercial; A: Available upon request. **The molecular weight, MW, is in g/mole; the cmc in mM and (wt%); N_{agg} is the micelle aggregation number; the hydrodynamical radii, R_h , and the radius of gyration, R_g in nm.; PDC : protein-detergent complex; bR: bacteriorhodopsin; Cx26: human Connexin26; CMP-Sia: mouse cytidine-5'-monophosphate-sialic acid transporter; β_1 AR: human β_2 adrenergic receptor; β_2 AR: human β_2 adrenergic receptor-T4Lysosyme fusion; cyt $\beta\alpha_3$: bacterial cytochrome bo3 ubiquinol oxidase; cyt. b_6/f : cytochrome b_6/f from *Citlamydomonas reinhardtii*; δ OR: δ -opioid receptor-T4Lysosyme fusion; DAGK: diacylglycerol kinase; GlpG: rhomboidintramembrane serine protease; hKOR1: human kappa opioid receptor type 1; hSmo: human GPCR Smoothened; hPtc: human receptor Patched; LacY: bacterial Lactose permease; LeuT: bacterial Leucine transporter; MelB: Melibiose permease; M3AchrR: Muscarinic M3 acetylcholin receptor; RC-LHI: Reaction centre-Light Harvesting complex I supercomplex, *Rhodobacter capsulatus*; Rho: bovine Rhodopsin; SQR: Succinate:quinone oxydoreductase; TmPPase : Sodium pumping pyrophosphatase.

Family Compound (main ref/review)	Rational	Availability*	Size and shape**	Membrane solubilisation	Membrane protein stabilization	Other applications	Crystallization
F- and HF- surfactants (this work)	Bulkier and more rigid tail; lipophobic tail that poorly solubilises lipids & hydrophobic cofactors	NC	May form rods or globular micelles F ₆ -DigluM: MW = 863 cmc ~ 0.38 $N_{agg} \sim 54$ $R_h \sim 2.7$ $R_g \sim 1.9$	- Does not solubilise thylakoid membranes	- Stabilizes bR, OmpA, cyt. b_6/f , hSmo, hPtc ATPsynthase, UCPI, AAC, as well or better than conventional detergents (Breyton et al 2004, Breyton et al 2009, Talbot et al 2009, Joubert et al 2010, Nehmé et al 2010, Blesneac et al 2012, Breyton et al 2013b)	- Promotes refolding (Lebaupain 2007)) - Cell-free synthesis (Park et al 2007, Breyton et al 2009, Park et al 2010, Blesneac et al 2012) - Chaperone activity for membrane insertion (Posokhov et al 2008, Raychaudhuri et al (2011, Kyrtychenko et al 2012a) - SANS (Breyton et al 2009, Breyton et al 2009, Breyton et al 2013b)	- 2D crystallization and electron microscopy (Fotinou et al 2013)

Table 8.4 (continued)

Family Compound (main ref/review)	Rational	Avail-ability*	Size and shape**	Membrane solubilisation	Membrane protein stabilization	Other applications	Crystallization
Neopentyl-maltoside MNG (Chae et al 2010c)	Lipid-like detergents with a central quaternary carbon that restrains the conformational flexibility	C	MNG-3: MW = 1005 CMC ~ 0.01 (0.001) (Chae et al 2010c) CMC ~ 0.00001 (Selao et al 2011) R_h 7.2	- As efficient as DDM on insect cell and photo-synthetic membranes (Chae et al 2010c, Selao et al 2011, Jiang et al 2012)	- Stabilizes δ_2 AR, M ₃ AchR, MelB, SQR, LeuT, cyt b_{03} , CMP-Sia, δ OR as well or better than DDM (Chae et al 2010c) - Stabilizes δ_2 AR and LacY in their membrane conformation (Chung et al 2011, Jiang et al 2012)	- NMR of δ_2 AR (Chung et al 2011) - Cell-free synthesis (Chung et al 2011) - H-D exchange mass spec. (Chung et al 2011, Westfield et al 2011) - Neg. stain EM (Westfield et al 2011) - Blue Native gels better resolved / DDM (Selao et al 2011)	- b_6/f (3.4Å) (Chung et al 2011) - Transfer in Lipidic Cubic Phase, crystals of various GPCR (3.5Å) (Rasmussen et al 2011, Rosenbaum et al 2011) - Acetate transporter, (4.1Å) (Chae et al 2013b) and TmPPase (2.6Å) (Kell-losalo et al 2012)
Neopentyl-glucoside GNG (Chae et al 2013b)	Forms smaller micelles than MNG, thus smaller PDC, to promote membrane protein crystallization	C	MW = 541-629 CMC ~ 0.1-1.6 (0.0092-0.1) R_h = 2.6-10.6	- As efficient as DDM on and bacterial photosyn-thetic and plasma membranes (Chae et al 2013b)	- Stabilizes RC-LHI, SQR, GIpG as well or better than conventional detergent but not LeuT. - PDC smaller than in DDM (Chae et al 2013b)	-	

Table 8.4 (continued)

Family Compound (main ref/review)	Rational	Availability*	Size and shape**	Membrane solubilisation	Membrane protein stabilization	Other applications	Crystallization
Branched-maltosides (Hong et al 2010)	<ul style="list-style-type: none"> Prevents water penetration into micelle Forms smaller PDC 	A	<p>CMC = 6–0.02 (0.28–0.0011)</p> <p>R_h 2.4–3.6 (Hong et al 2010)</p>	–	<ul style="list-style-type: none"> Stabilizes LacY (Jiang et al 2012) in its membrane conformation Stabilizes Cx26 dodecamer (Hong et al 2010) 	–	<ul style="list-style-type: none"> Different crystal form of Cx26/DDM, anisotropic diffraction (Hong et al 2010) – b_6f (2.8Å) (Hovers et al 2011) and RC-LHI (Barret et al 2013)
Cyclic maltosides (Hovers et al 2011)	<ul style="list-style-type: none"> More rigid hydrophobic domain to increase van der Waals contacts in the bulk and decrease them with the protein surface 	C	<p>CMC ~ 0.036</p> <p>$N_{agg} \sim 165$</p> <p>R_g 3.3 (Barret et al 2013)</p> <p>Oblate micelles</p>	–	<ul style="list-style-type: none"> Stabilizes hSmO, hPT, δ_1AR and cyt. b_6f as well or better than DDM (Hovers et al 2011) 	–	–
Cholesterol-based detergent (Howell et al 2010)	<ul style="list-style-type: none"> Mimics cholesterol in mixed micelles 	C	<p>CMC ~ 0.003</p> <p>$N_{agg} \sim 200$</p> <ul style="list-style-type: none"> Does not increase detergent micelles size much compared to cholesterol 	<ul style="list-style-type: none"> As efficient as DDM on <i>E. coli</i> membranes, increased efficiency in mixed micelles 	<ul style="list-style-type: none"> Stabilizes the GPCR hKOR1 better than DDM 	<ul style="list-style-type: none"> Easier to manipulate than Cholesterol and Cholesteryl hemisuccinate 	–

Table 8.4 (continued)

Family Compound (main ref/review)	Rational	Availability*	Size and shape**	Membrane solubilisation	Membrane protein stabilization	Other applications	Crystallization
Steroid-based dimaltoside (Chae et al 2012)	Rigid and flat hydrophobic domain to promote association with complementary protein surface	C	GDN: MW = 1165 CMC ~ 0.018 (0.0021) R_h 3.9	– More efficient than DDM at solubilising bacterial, yeast, insect cells membranes	– Stabilizes bR, RC-LHI, CMP-Sia, GlpG, SQR, MelB, Leu I and the GPCRs δ OR, δ_2 AR as well or better than DDM	– Stabilizes proteins (<i>i.e.</i> GPCR) that are stabilized by cholesterol	–
Steroid-based facial maltoside and phoscholines (Lee et al 2013)	– Mimics cholesterol – Face to face packing when self-assembled – Increases interaction surface with hydrophobic domain of protein – Makes small PDC	C	CMC 0.18–0.19 (0.0219–0.028) N_{agg} 10–30 R_h 1.9–2.5 Forms smaller PDC/DDM	– Solubilises bacterial and insect cell membranes (Zhang et al 2007, Lee et al 2013) – Does not solubilise purple membrane.	– Stabilizes MsbA and bR better than conventional detergents, Cx26 as well (Zhang et al 2007) – Initial activity of MsbA greater than in DDM (Lee et al 2013)	– Tight detergent-protein detergent interactions evidenced by “native” Mass spectrometry – Neg. stain EM – Forms bicelles with lipids	– Open (3.3Å) and closed (3.8Å) Cx26, – MsbA, larger and better diffracting crystals/DDM, – bR in bicelles – More and better crystals of cyt. P450s, in mixed micelles

Table 8.4 (continued)

Family Compound (main ref/review)	Rational	Availability*	Size and shape**	Membrane solubilisation	Membrane protein stabilization	Other applications	Crystallization
Facial maltoside (Chae et al 2010a)	Same as above, covers the whole trans-membrane domain with one molecule	NC	MW = 2232–2148 CMC 0.007–0.013 (0.0016–0.0028) $N_{agg} \sim 6$ $R_h 1.9–3.3$	–	–Stabilizes bR, RC-LHI, LeuT, cyt <i>b₀₃</i> better than conventional detergents, not δ_2 AR	–	–
Bola-amphiphiles (Li et al 2009)	Mimics the lipids of extremophiles to stabilize membrane proteins	C	–	–	– Enhanced thermal stabilization of DAGK 19 times in mixed micelles – Plays a lipid-like role in the enzymatic activation of DAGK	– Compatible with NMR measurements	–
Tripod (Chae et al 2013a)	– Decreases the flexibility of the molecule – Forms small PDC	C	TPA-8 (Chae et al 2013a) MW = 716 CMC ~ 0.42 $R_h 4.1$	– Very efficiently solubilises purple, photosynthetic and Rod outer segment membranes (Chae et al 2010b, Chae et al 2013a)	– Stabilizes delipidated bR, Rho and RC-LHI better than conventional detergents (Chae et al 2010b, Chae et al 2013a)	–	– Vapour diffusion crystallisation of bR and K-channel (reviewed in (Chae et al 2010b))

Table 8.4 (continued)

Family Compound (main ref/review)	Rational	Availability*	Size and shape**	Membrane solubilisation	Membrane protein stabilization	Other applications	Crystallization
Dodecyl-trehalose (Tao et al 2012)	<ul style="list-style-type: none"> - Trehalose: high chemical stability and rigidity - Investigate different molecule geometry 	NC	CMC 0.14–0.47 (0.007–0.024) $R_h = 2.9$ –3.4	- Solubilises insect cell and bacterial membranes as well as DDM	- 4-DDTre stabilizes ORL1 and β_2 AR better than DDM, - 2-DDTre stabilizes MsbA better /DDM	-	-
DDM (for a comparison)	-	C	MW = 510 cmc = 0.017 (0.0087) $N_{agg} = 132$ $R_h = 3.4$ Oblate ellipsoid micelle (Dupuy et al 1997)	-	-	-	-

Table 8.5 Sugar-based non-conventional amphiphiles for membrane protein study: names and chemical formulae.

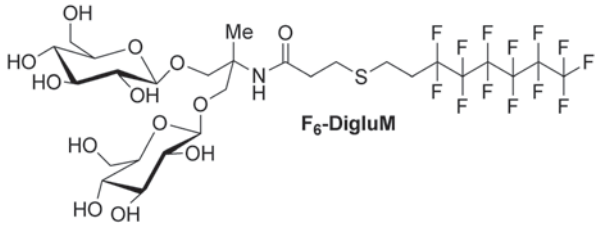
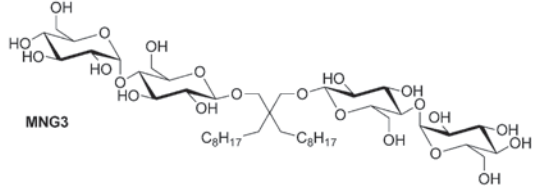
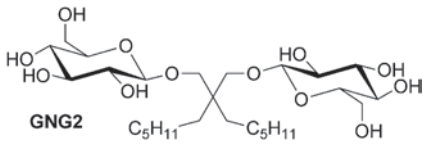
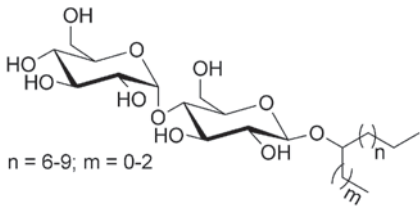
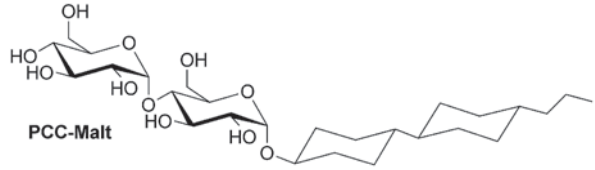
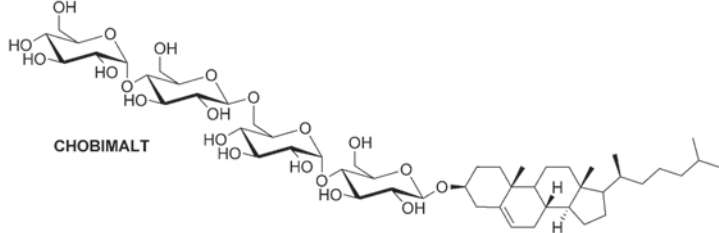
Family Compound	Example
F- and HF- surfactants	 <p>F₆-DigluM</p>
Neopentyl- maltosides MNG	 <p>MNG3</p>
Neopentyl- glucosides GNG	 <p>GNG2</p>
Branched- maltosides	 <p>$n = 6-9; m = 0-2$</p>
Cyclic maltosides	 <p>PCC-Malt</p>
Cholesterol- based detergent	 <p>CHOBIMALT</p>

Table 8.5 (continued)

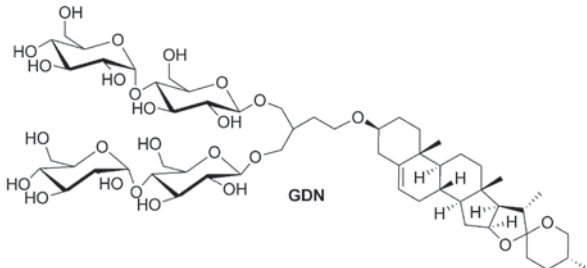
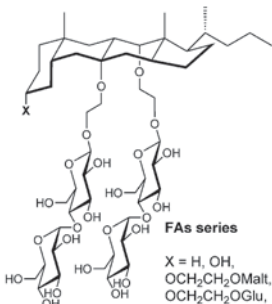
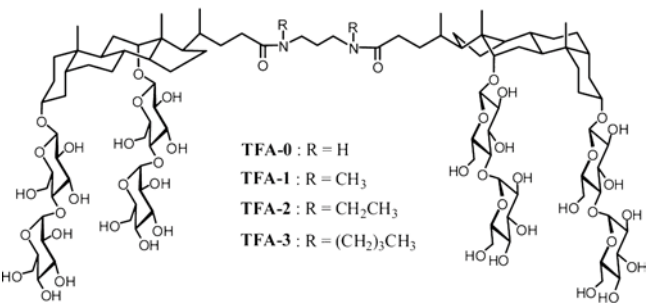
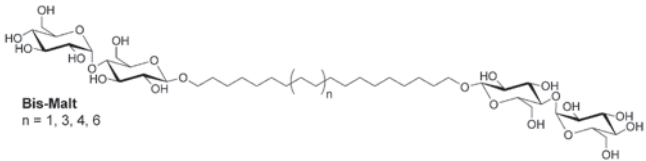
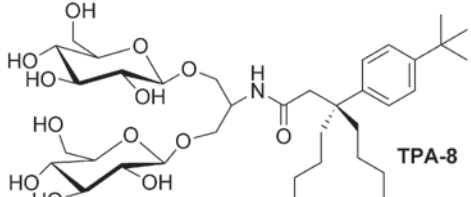
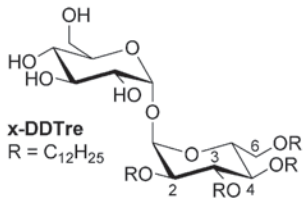
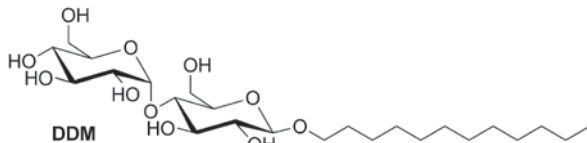
Family Compound	Example
Steroid-based di-maltoside	 <p style="text-align: center;">GDN</p>
Steroid-based facial maltosides and pho- scholines	 <p style="text-align: center;">FAs series</p> <p style="text-align: center;">X = H, OH, OCH₂CH₂OMalt, OCH₂CH₂OGlu,</p>
Facial maltosides	 <p style="text-align: center;">TFA-0 : R = H TFA-1 : R = CH₃ TFA-2 : R = CH₂CH₃ TFA-3 : R = (CH₂)₃CH₃</p>
Bola-amphi- philes	 <p style="text-align: center;">Bis-Malt n = 1, 3, 4, 6</p>
Tripods	 <p style="text-align: center;">TPA-8</p>

Table 8.5 (continued)

Family Compound	Example
Dodecyl-trehaloses	 <p>x-DDTre R = C₁₂H₂₅</p>
DDM (for a comparison)	 <p>DDM</p>

they are at concentrations well above their CMC (Breyton et al. 2004; Chae et al. 2010c; Hovers et al. 2011; Chae et al. 2012). This is not the case for conventional detergents (e.g., Breyton et al. 1997). In some cases, the efficiency of the amphiphile proved enhanced when used in mixed micelles with conventional detergents, as for CHOBIMALT and Bola-maltosides (Li et al. 2009; Howell et al. 2010).

8.5.3 The Importance of the Variation Around the Theme

For each strategy, several molecules are synthesized and tested. Indeed, even if one can have guesses at the general physical-chemical trend of a designed molecule, predicting accurately its behavior is more difficult. Furthermore, biochemistry of MPs is still a trial-and-error science. Subtle modifications in the amphiphile can affect protein stability and crystal quality: the sodium-proton antiporter NhaA forms well-diffracting crystals in α -DDM, when they are of poor quality in β -DDM (Screpanti et al. 2006). In addition, what is best for one protein is not necessarily what is best for the next, as each protein behaves its own way.

- In the *F*-surfactant series, a number of parameters were varied, that had interesting effects on the physical chemistry of the aggregates (see Sect. 8.3), as in biochemistry (see Sect. 8.4).
- In the neopentyl glycol series, the length of the alkyl chain (C₅-C₁₀) and of the head group (maltose or glucose) was varied, as the chemical nature of the linker between the hydrophilic and hydrophobic moieties (amine, MNG-1, ether,

MNG-2, or direct C–C linkage, MNG-3; Chae et al. 2010c). This latter parameter, for example, had an effect on the CMC of the different compounds, but also on their capacity to solubilize membranes and to stabilize MPs: MNG-1 was found less efficient at solubilizing membranes than DDM, whereas MNG-2 and 3 were as efficient (Chae et al. 2013b). This could correlate with the CMC of MNG-1 being twice that of MNG-2 and MNG-3. As regards the GNG series, the CMC of GNG-1 and 3 is one order of magnitude higher than that of MNG derivatives, while that of GNG-2 is in the same range. GNGs were generally stabilizing; however, the Leucin transporter was less stable in any of them when compared to either MNGs or DDM. As a rule, within each series, MNG-3 and GNG-2 proved more stabilizing toward proteins tested (see Table 8.4). It is interesting that both molecules form relatively large aggregates of $R_h \sim 7\text{--}10$ nm while for the others, smaller aggregates of $R_h \sim 2.5\text{--}3.5$ nm were observed.

- In the case of branched maltosides, different length of the main chain, with six to nine carbons, was investigated, in combination with the shorter chain with one to three carbons. Regarding the CMC, each carbon increment on the short branch is equivalent to a half-carbon increment on the main chain. The addition of a methyl branch reduces the micelle size, the presence of an ethyl branch increases it slightly, and the branching of a propyl group induces the formation of larger aggregates, the effect being more pronounced on the longer-chained maltosides. The effect of the addition of a methyl group either on the sixth or tenth carbon of the main chain was also investigated. In the latter, micelle size was not affected, whereas in the former, it increased from $R_h \sim 3$ to 4–6 nm. With regard to biochemistry, different members of this family indeed stabilized different proteins in their membrane conformation/oligomeric state (Hong et al. 2010).
- In the cyclic-maltoside family, different combinations of phenyl and cyclohexyl groups were grafted onto a maltoside head group. Whereas the number of carbon atoms was identical in each molecule, phenyl-containing molecules exhibits higher CMCs, and the bi-cyclohexyl group promotes a micelle of smaller curvature, with the radius of the hydrophobic core of the micelles estimated to 1.12 nm, i.e., shorter than that of DDM (1.37 nm). Interestingly, its CMC was also significantly lower than DDM, suggesting that cycloalkanes lead to more hydrophobic and shorter tails than aliphatic alkanes with the same number of carbon atoms. Whereas the bi-cyclohexyl molecule (PCC-Malt) seems to be the favorite of the majority of the proteins tested, one of them is more stable with cyclic maltosides having a combination of cyclohexyl and phenyl groups (Hovers et al. 2011).
- In the steroid-based di-maltosides, two steroid-based lipophilic groups were investigated, leading to glyco-diosgenin (GDN) and glyco-lithocholate (GLC), the latter being grafted to the di-maltoside head group via amine, ether, or C–C linkage. As for the MNG family, the nature of the linker in the GLC series influences slightly the hydrophobicity of the molecules, and their CMC, the amide derivative being the more polar compound. Both GLC and GDN form small

micelles of $R_h \sim 3\text{--}4$ nm, those of the GLC series being slightly smaller. However, for all proteins tested, GDN proved superior to GLD and DDM (Chae et al. 2012). CHOBIMALT is another steroid-based surfactant with two in-series maltosides as head group in contrast to the GDN and GLC series whose polar head is branched. The steroid moiety also differs. CHOBIMALT shows a cholesterol-like behavior, with preaggregation below its CMC.

- In the steroid-facial maltoside family, the cholate molecule was grafted with different sugars (glucosides or maltosides) or phosphocholine. The identity, number, and position of the polar groups significantly affect the detergent properties and their capacity to stabilize MPs (Lee et al. 2013).
- In the tripod family, the location of the branching point, whether in the hydrophilic or hydrophobic part of the molecule, and the nature of the phenyl tail, by introducing different alkyl substituents at the para position, were modulated. The addition of two carbons per alkyl chain results in a ~ 6 -times drop of the CMC and a doubling of the R_h , and the addition of a *tert*-butyl group to the aromatic ring also induced a 9-times drop of the CMC and the formation of larger aggregates (Chae et al. 2013a). Sugar-based tripods exhibit good water solubility of up to $\sim 20\%$ (w/v) which is significantly higher than that of the *N*-oxide derivatives.
- In the dodecyl-trehalose family, the grafting position of the dodecyl chain onto the trehalose head group was shown to influence the physical-chemical properties of the resulting detergent. The differences observed in the CMC of the derivatives were not correlated with the hydrophobicity parameter, indicating that the position of the free hydroxyl groups and thus interaction between polar head groups play also a role in micellization. This was further confirmed by the size of the micelles formed by these derivatives, which ranged from small micelles (2-DDT, $R_h = 2.9$ nm) to large and polydisperse aggregates (4-DDT, $R_h \sim 6$ nm). Interestingly, of the three proteins tested, two were more stable with one member of the family, while the third preferred another one (Tao et al. 2012).

8.6 Conclusion

The Graal of “the” universal amphiphile for stabilizing and crystallizing MPs is clearly out of reach. It is unfortunate that no universal amphiphile suits every MP. Rather, each new family of amphiphile is another tool in the toolbox of the biochemist. Depending on the question asked, on the technique tackled, and on the protein handled, one molecule will be more appropriate than another one. In this context, the availability of the molecules is an important issue for the biochemist. As indicated in Table 8.4, some of these amphiphiles are commercially available, whereas others are not, making their accessibility to the community more difficult.

As regards *F*-surfactants, we are currently working on the scale-up synthesis of the most promising agents to make them available to the community of biochemists and biophysicists. Among the particular applications for which resorting to *F*-surfactants looks quite promising, one should cite the study of fragile complexes, structural investigation using SANS, cell-free synthesis, as well as the delivery of MPs to preexisting membranes.

Acknowledgments This work was supported by grants from the E.U. (FP7 Grant agreements N226507-NMI3) and the “Agence Nationale pour la Recherche” (ANR 07 PCV—Promensurf—n 0010-02). Respectful thanks to Prof. Bernard Pucci (University of Avignon) and Dr. Jean-Luc Popot (IBPC, Paris) who pioneered the development of *F*-surfactants for the study of MPs. Thanks are also due to A. Polidori and F. Bonneté (University of Avignon) who are involved in the synthesis and physical-chemical analysis of various *F*-surfactants. We thank Simon Raynal (University of Avignon) for help with the synthesis of some *F*-surfactants, Aline Le Roy (IBS, Grenoble) for help with the Biochemistry and AUC measurements, and Dr. Pierre Guillet (University of Avignon) for help with some of the figures.

References

- Abla M, Durand G, Pucci B (2008) Glucose-based surfactants with hydrogenated, fluorinated or hemifluorinated tails: synthesis and comparative physical-chemical characterization. *J Org Chem* 73:8142–8153
- Abla M, Durand G, Pucci B (2011) Propyl-ended hemifluorinated surfactants: synthesis and self-assembling properties. *J Org Chem* 76:2084–2093
- Abla M, Durand G, Breyton C, Raynal S, Ebel C, Pucci B (2012) A diglucosylated fluorinated surfactant to handle integral membrane proteins in aqueous solution. *J Fluor Chem* 134:63–71
- Abla M, Unger S, Keller S, Bonneté F, Ebel C, Pucci B, Breyton C, Durand G Micellar and biochemical properties of a propyl-ended fluorinated surfactant designed for membrane-protein study. In preparation
- Adamson AW (1990) *Physical chemistry of surfaces*, 5th ed., Wiley Interscience, New York, pp 131–145
- Allen FH, Bird CM, Rowland RS, Raithby PR (1997) Hydrogen-bond acceptor and donor properties of divalent sulfur (Y-S-Z and R-S-H). *Acta Crystallogr Sect B Struct Sci* B53:696
- Barret L-A, Barrot-Ivolot C, Raynal S, Jungas C, Polidori A, Bonneté F (2013) Influence of hydrophobic micelle structure on crystallization of the photosynthetic RC-LH1-PufX complex from *rhodobacter blasticus*. *J Phys Chem B* 117:8770–8781
- Barthelemy P, Ameduri B, Chabaud E, Popot JL, Pucci B (1999) Synthesis and preliminary assessments of ethyl-terminated perfluoroalkyl nonionic surfactants derived from tris(hydroxymethyl) acrylamidomethane. *Org Lett* 1:1689–1692
- Bernhard F, Tozawa Y (2013) Cell-free expression—making a mark. *Curr Opin Struct Biol* 23:374–380
- Bill RM, Henderson PJF, Iwata S, Kunji ERS, Michel H, Neutze R, Newstead S, Poolman B, Tate CG, Vogel H (2011) Overcoming barriers to membrane protein structure determination. *Nat Biotech* 29:335–340
- Blesneac I, Ravaud S, Juillan-Binard C, Barret L-A, Zoonens M, Polidori A, Miroux B, Pucci B, Pebay-Peyroula E (2012) Production of UCP1 a membrane protein from the inner mitochondrial membrane using the cell free expression system in the presence of a fluorinated surfactant. *Biochim Biophys Acta* 1818:798–805

- Borch J, Hamann T (2009) The nanodisc: a novel tool for membrane protein studies. *Biol Chem* 390:805–814
- Braumann T (1986) Determination of hydrophobic parameters by reversed-phase liquid chromatography: theory, experimental techniques, and application in studies on quantitative structure-activity relationships. *J Chromatogr A* 373:191–225
- Breyton C, Tribet C, Olive J, Dubacq J-P, Popot J-L (1997) Dimer to monomer conversion of the cytochrome b_6/f complex. Causes and consequences. *J Biol Chem* 272:21892–21900
- Breyton C, Chabaud E, Chaudier Y, Pucci B, Popot JL (2004) Hemifluorinated surfactants: a non-dissociating environment for handling membrane proteins in aqueous solutions? *FEBS Lett* 564:312–318
- Breyton C, Gabel F, Abla M, Pierre Y, Lebaupain F, Durand G, Popot J-L, Ebel C, Pucci B (2009) Micellar and biochemical properties of (hemi)fluorinated surfactants are controlled by the size of the polar head. *Biophys J* 97:1–10
- Breyton C, Pucci B, Popot J-L (2010) Amphipols and fluorinated surfactants: two alternatives to detergents for studying membrane proteins in vitro. In: Mus-Veteau I (ed) *Heterologous expression of membrane proteins*, vol 601. Humana Press, New York, pp 219–245
- Breyton C, Flayhan A, Gabel F, Lethier M, Durand G, Boulanger P, Chamig M, Ebel C (2013a) Assessing the conformation changes of pb5, the receptor binding protein of phage T5, upon binding to its *E. coli* receptor FhuA. *J Biol Chem* 288(42):30763–30772
- Breyton C, Gabel F, Lethier M, Flayhan A, Durand G, Jault J-M, Juillan-Binard C, Imbert L, Moulin M, Ravaud S, Härtlein M, Ebel C (2013b) Small angle neutron scattering for the study of solubilised membrane proteins. *Eur Phys J E Soft Matter* 36:71–86
- Caffrey M (2009) Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu Rev Biophys* 38:29–51
- Chabaud E, Barthélémy P, Mora N, Popot JL, Pucci B (1998) Stabilization of integral membrane proteins in aqueous solution using fluorinated surfactants. *Biochimie* 80:515–530
- Chae PS, Gotfryd K, Pacyna J, Miercke LJW, Rasmussen SGF, Robbins RA, Rana RR, Loland CJ, Kobilka B, Stroud R, Byrne B, Gether U, Gellman SH (2010a) Tandem facial amphiphiles for membrane protein stabilization. *J Am Chem Soc* 132:16750–16752
- Chae PS, Laible PD, Gellman SH (2010b) Tripod amphiphiles for membrane protein manipulation. *Mol Biosyst* 6:89–94
- Chae PS, Rasmussen SGF, Rana RR, Gotfryd K, Chandra R, Goren MA, Kruse AC, Nurva S, Loland CJ, Pierre Y, Drew D, Popot J-L, Picot D, Fox BG, Guan L, Gether U, Byrne B, Kobilka B, Gellman SH (2010c) Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* 7:1003–1008
- Chae PS, Rasmussen SGF, Rana RR, Gotfryd K, Kruse AC, Manglik A, Cho KH, Nurva S, Gether U, Guan L, Loland CJ, Byrne B, Kobilka BK, Gellman SH (2012) A new class of amphiphiles bearing rigid hydrophobic groups for solubilization and stabilization of membrane proteins. *Chem Eur J* 18:9485–9490
- Chae PS, Cho KH, Wander MJ, Bae HE, Gellman SH, Laible PD (2013a) Hydrophobic variants of ganglio-tripod amphiphiles for membrane protein manipulation. *Biochim Biophys Acta (BBA)—Biomembr.* 1838:278–286
- Chae PS, Rana RR, Gotfryd K, Rasmussen SGF, Kruse AC, Cho KH, Capaldi S, Carlsson E, Kobilka B, Loland CJ, Gether U, Banerjee S, Byrne B, Lee JK, Gellman SH (2013b) Glucose-neopentyl glycol (GNG) amphiphiles for membrane protein study. *Chem Comm* 49:2287–2289
- Chaudier Y, Zito F, Barthelemy P, Stroebel D, Ameduri B, Popot J-L, Pucci B (2002) Synthesis and preliminary biochemical assessment of ethyl-terminated perfluoroalkylamine oxide surfactants. *Bioorg Med Chem Lett* 12:1587–1590
- Cho KH, Byrne B, Chae PS (2013) Hemifluorinated maltose-neopentyl glycol (HF-MNG) amphiphiles for membrane protein stabilisation. *Chembiochem* 14:452–455
- Chung KY, Rasmussen SGF, Liu T, Li S, DeVree BT, Chae PS, Calinski D, Kobilka BK, Woods VL Jr, Sunahara RK (2011) Conformational changes in the G protein Gs induced by the β_2 adrenergic receptor. *Nature* 477:611–615

- Dauvergne J, Polidori A, Vénien-Bryan C, Pucci B (2008) Synthesis of a hemifluorinated amphiphile designed for self-assembly and two-dimensional crystallization of membrane protein. *Tet Lett* 49:2247–2250
- Diller A, Loudet C, Aussenac F, Raffard G, Fournier S, Laguerre M, Grélard A, Opella SJ, Marassi FM, Dufourc EJ (2009) Bicelles: a natural “molecular goniometer” for structural, dynamical and topological studies of molecules in membranes. *Biochimie* 91:744–751
- Dupont A, Eastoe J, Barthélémy P, Pucci B, Heenan R, Penfold J, Steytler DC, Grillo I (2003) Neutron reflection and small-angle neutron scattering studies of a fluorocarbon telomer surfactant. *J Coll Interface Sci* 261:184–190
- Dupuy C, Auvray X, Petipas C, Rico-Lattes I, Lattes A (1997) Anomeric effects on the structure of micelles of alkyl maltosides in water. *Langmuir* 13:3965–3967
- Eastoe J, Paul A, Rankin A, Wat R, Penfold J, Webster JRP (2001) Fluorinated nonionic surfactants bearing either CF₃- or H-CF₂- terminal groups: adsorption at the surface of aqueous solutions. *Langmuir* 17:7873–7878
- Fotinou C, Aittoniemi J, de Wet H, Polidori A, Pucci B, Sansom MSP, Vénien-Bryan C, Ashcroft FM (2013) Tetrameric structure of SUR2B revealed by electron microscopy of oriented single particles. *FEBS J* 280:1051–1063
- Garofalakis G, Murray BS, Sarney DB (2000) Surface activity and critical aggregation concentration of pure sugar esters with different sugar headgroups. *J Coll Interface Sci* 229:391–398
- Hong W-X, Baker KA, Ma X, Stevens RC, Yeager M, Zhang Q (2010) Design, synthesis, and properties of branch-chained maltoside detergents for stabilization and crystallization of integral membrane proteins: human connexin 26. *Langmuir* 26:8690–8696
- Hovers J, Potschies M, Polidori A, Pucci B, Raynal S, Bonneté F, Serrano-Vega MJ, Tate CG, Picot D, Pierre Y, Popot JL, Nehmé R, Bidet M, Mus-Veteau I, Bußkamp H, Jung K-H, Marx A, Timmins PA, Welte W (2011) A class of mild surfactants that keep integral membrane proteins water-soluble for functional studies and crystallization. *Mol Membr Biol* 28:171–181
- Howell SC, Mittal R, Huang L, Travis B, Breyer RM, Sanders CR (2010) Chobimalt: a cholesterol-based detergent. *Biochemistry* 49:9572–9583
- Israelachvili JN, Mitchell DJ, Ninham BW (1977) Theory of self-assembly of lipid bilayers and vesicles. *Biochim Biophys Acta—Biomembr* 470:185–201
- Jiang X, Guan L, Zhou Y, Hong W-X, Zhang Q, Kaback HR (2012) Evidence for an intermediate conformational state of lacy. *Proc Natl Acad Sci U S A* 109:E698–E704
- Joubert O, Nehmé R, Bidet M, Mus-Veteau I (2010) Heterologous expression of human membrane receptors in the yeast *saccharomyces cerevisiae*. In: Mus-Veteau I (ed) Heterologous expression of membrane proteins, vol 601. Humana Press, New York, pp 87–103
- Kang HJ, Lee C, Drew D (2013) Breaking the barriers in membrane protein crystallography. *Int J Biochem Cell Biol* 45:636–644
- Kellosalo J, Kajander T, Kogan K, Pokharel K, Goldman A (2012) The structure and catalytic cycle of a sodium-pumping pyrophosphatase. *Science* 337:473–476
- Kirsch P (2004) Modern fluoroorganic chemistry: synthesis, reactivity, applications. Wiley-VCH, Weinheim
- Kissa E (2001) Fluorinated surfactants and repellents. Marcel Dekker, New York
- Kyrychenko A, Rodnin MV, Posokhov YO, Holt A, Pucci B, Killian JA, Ladokhin AS (2012a) Thermodynamic measurements of bilayer insertion of a single transmembrane helix chaperoned by fluorinated surfactants. *J Mol Biol* 416:328–334
- Kyrychenko A, Rodnin MV, Vargas-Urbe M, Sharma SK, Durand G, Pucci B, Popot J-L, Ladokhin AS (2012b) Folding of diphtheria toxin T-domain in the presence of amphipols and fluorinated surfactants: toward thermodynamic measurements of membrane protein folding. *Biochim Biophys Acta—Biomembr* 1818:1006–1012
- le Maire M, Champeil P, Møller JV (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta—Biomembr* 1508:86–111
- Lebaupain F (2007) Développement de l'utilisation des tensioactifs fluorés pour la biochimie des protéines membranaires. PhD Degree, Université Pierre et Marie Curie, Paris

- Lebaupain F, Salvay AG, Olivier B, Durand G, Fabiano A-S, Michel N, Popot J-L, Ebel C, Breyton C, Pucci B (2006) Lactobionamide surfactants with hydrogenated, perfluorinated or hemifluorinated tails: physical-chemical and biochemical characterization. *Langmuir* 22:8881–8890
- Lebeau L, Lach F, Venien-Bryan C, Renault A, Dietrich J, Jahn T, Palmgren MG, Kühlbrandt W, Mioskowski C (2001) Two-dimensional crystallization of a membrane protein on a detergent-resistant lipid monolayer. *J Mol Biol* 308:639–647
- Lee SC, Bennett BC, Hong W-X, Fu Y, Baker KA, Marcoux J, Robinson CV, Ward AB, Halpert JR, Stevens RC, Stout CD, Yeager MJ, Zhang Q (2013) Steroid-based facial amphiphiles for stabilization and crystallization of membrane proteins. *Proc Nat Acad Sci U S A* 110:1203–1211
- Leo A, Hansch C, Elkins D (1971) Partition coefficients and their uses. *Chem Rev* 71:525–616
- Li Q, Mittal R, Huang L, Travis B, Sanders CR (2009) Bolaamphiphile-class surfactants can stabilize and support the function of solubilized integral membrane proteins. *Biochemistry* 48:11606–11608
- Lo Nostro P, Chen S-W (1993) Aggregation of a semifluorinated n-alkane in perfluorooctane. *J Phys Chem* 97:6535–6540
- Lundberg D, Shi L, Menger FM (2008) Self-assembly of cationic surfactants that contain thioether groups in the hydrophobic tails. *Langmuir* 24:4530–4536
- Menger FM, Shi L (2006) Exposure of self-assembly interiors to external elements: a kinetic approach. *J Am Chem Soc* 128:9338–9339
- Nagarajan R (2002) Molecular packing parameter and surfactant self-assembly: the neglected role of the surfactant tail. *Langmuir* 18:31–38
- Nehmé R, Joubert O, Bidet M, Lacombe B, Polidori A, Pucci B, Mus-Veteau I (2010) Stability study of the human G-protein coupled receptor, smoothed. *Biochim Biophys Acta* 1798:1100–1110
- Nury H, Dahout-Gonzalez C, Trézéguet V, Lauquin G, Brandolin G, Pebay-Peyroula E (2005) Structural basis for lipid-mediated interactions between mitochondrial ADP/ATP carrier monomers. *FEBS Lett* 579:6031–6036
- Ostermeier C, Michel H (1997) Crystallization of membrane proteins. *Curr Opin Struct Biol* 7:697–701
- Palsdottir H, Hunte C (2004) Lipids in membrane protein structures. *Biochim Biophys Acta—Biomembr* 1666:2–18
- Park KH, Berrier C, Lebaupain F, Pucci B, Popot J-L, Ghazi A, Zito F (2007) Fluorinated and hemifluorinated surfactants as alternatives to detergents for membrane protein cell-free synthesis. *Biochem J* 403:183–187
- Park K-H, Billon-Denis E, Dahmane T, Lebaupain F, Pucci B, Breyton C, Zito F (2010) In the cauldron of cell-free synthesis of membrane proteins: playing with new surfactants. *N Biotechnol* 28:255–261
- Polidori A, Pucci B, Maurizis JC, Pavia AA (1994) Synthesis of nonionic glycosidic surfactants derived from tris(hydroxymethyl) aminomethane: preliminary assessment. *N J Chem* 18:839–848
- Polidori A, Passet M, Lebaupain F, Ameduri B, Popot J-L, Breyton C, Pucci B (2006) Fluorinated and hemifluorinated surfactants derived from maltose: synthesis and application to handling membrane proteins in aqueous solution. *Bioorg Med Chem Lett* 16:5827–5831
- Popot J-L (2010) Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annu Rev Biochem* 79:737–775
- Popot J-L, Gerchman S-E, Engelman DM (1987) Refolding of bacteriorhodopsin in lipid bilayers: a thermodynamically controlled two-stage process. *J Mol Biol* 198:655–676
- Posokhov YO, Rodnin MV, Das SK, Pucci B, Ladokhin AS (2008) FCS study of the thermodynamics of membrane protein insertion into the lipid bilayer chaperoned by fluorinated surfactants. *Biophys J* 95:L54–L56

- Rasmussen SGF, Choi H-J, Fung JJ, Pardon E, Casarosa P, Chae PS, Devree BT, Rosenbaum DM, Thian FS, Kobilka TS, Schnapp A, Konetzki I, Sunahara RK, Gellman SH, Pautsch A, Steyaert J, Weis WI, Kobilka BK (2011) Structure of a nanobody-stabilized active state of the $\beta(2)$ adrenoceptor. *Nature* 469:175–180
- Raychaudhuri P, Li Q, Mason A, Mikhailova E, Heron AJ, Bayley H (2011) Fluorinated amphiphiles control the insertion of α -hemolysin pores into lipid bilayers. *Biochemistry* 50:1599–1606
- Riess JG (2005) Fluorous materials for biomedical uses. In: Gladysz JA, Curran DP, Horvath IT (ed) *Handbook of fluororous chemistry*. Wiley-VCH, Weinheim, pp 521–573
- Rodnina MV, Posokhov YO, Contino-Pépin C, Brettmann J, Kyrychenko A, Palchevskyy SS, Pucci B, Ladokhin AS (2008) Interactions of fluorinated surfactants with diphtheria toxin t-domain: testing new media for studies of membrane proteins. *Biophys J* 94:4348–4357
- Rosen MJ (2004) *Surfactants and interfacial phenomena*. Wiley, Brooklyn
- Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SGF, Choi H-J, Devree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK (2011) Structure and function of an irreversible agonist- $\beta(2)$ adrenoceptor complex. *Nature* 469:236–240
- Sadtler VM, Giulieri MP, M.P. K, Riess JG (1998) Micellization and adsorption of fluorinated amphiphiles: questioning the 1CF2 ~ 1.5CH2 rule. *Chem Eur J* 4:1952–1957
- Screpanti E, Padan E, Rimon A, Michel H, 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
- Selao TT, Branca R, Chae PS, Lehtiö J, Gellman SH, Rasmussen SGF, Nordlund S, Norén A (2011) Identification of chromatophore membrane protein complexes formed under different nitrogen availability conditions in *rhodospirillum rubrum*. *J Proteome Res* 10:2703–2714
- Singh R, Flowers RA (2010) Efficient protein renaturation using tunable hemifluorinated anionic surfactants as additives. *Chem Commun* 46:276–278
- Soderberg I, Drummond CJ, Furlong DN, Godkin S, Matthews B (1995) Non-ionic sugar-based surfactants: self-assembly and air/water interfacial activity. *Colloids Surf A* 102:91–97
- Sonoda Y, Newstead S, Hu N-J, Alguel Y, Nji E, Beis K, Yashiro S, Lee C, Leung J, Cameron AD, Byrne B, Iwata S, Drew D (2011) Benchmarking membrane protein detergent stability for improving throughput of high-resolution X-ray structures. *Structure* 19:17–25
- Srinivasan V, Blankschtein D (2005) Prediction of conformational characteristics and micellar solution properties of fluorocarbon surfactants. *Langmuir* 21:1647–1660
- Talbot J-C, Dautant A, Polidori A, Pucci B, Cohen-Bouhacina T, Maali A, Salin B, Brèthes D, Velours J, Giraud M-F (2009) Hydrogenated and fluorinated surfactants derived from tris(hydroxymethyl)-acrylamidomethane allow the purification of a highly active yeast F1-F0 ATP-synthase with an enhanced stability. *J Bioenerg Biomembr* 41:349–360
- Tanford C (1980) *The hydrophobic effect: formation of micelles and biological membranes*. Wiley, New York
- Tao H, Fu Y, Thompson A, Lee SC, Mahoney N, Stevens RC, Zhang Q (2012) Synthesis and properties of dodecyl trehaloside detergents for membrane protein studies. *Langmuir* 28:11173–11181
- Tate CG, Schertler GFX (2009) Engineering G protein-coupled receptors to facilitate their structure determination. *Curr Opin Struct Biol* 19:386–395
- Ujwal R, Bowie JU (2011) Crystallizing membrane proteins using lipidic bicelles. *Method San Diego Calif* 55:337–341
- Venkatesan P, Cheng Y, Kahne D (1994) Hydrogen bonding in micelle formation. *J Am Chem Soc* 116:6955–6956
- Wenk MR, Seelig J (1997) Interaction of octyl-beta-thioglucoopyranoside with lipid membranes. *Biophys J* 73:2565–2574
- Westfield GH, Rasmussen SGF, Su M, Dutta S, DeVree BT, Chung KY, Calinski D, Velez-Ruiz G, Oleskie AN, Pardon E, Chae PS, Liu T, Li S, Woods VL Jr, Steyaert J, Kobilka BK, Sunahara RK, Skiniotis G (2011) Structural flexibility of the G alpha s alpha-helical domain in the beta-2-adrenoceptor Gs complex. *Proc Natl Acad Sci U S A* 108:16086–16091

- Yildirim MA, Goh K-I, Cusick ME, Barabási A-L, Vidal M (2007) Drug-target network. *Nat Biotechnol* 25:1119–1126
- Zhang T, Marchant RE (1996) Novel polysaccharide surfactants: the effect of hydrophobic and hydrophilic chain length on surface active properties. *J Coll Interface Sci* 177:419–426
- Zhang Q, Ma X, Ward A, Hong W-X, Jaakola V-P, Stevens RC, Finn MG, Chang G (2007) Designing facial amphiphiles for the stabilization of integral membrane proteins. *Angew Chem Int Ed* 46:7023–7025
- Zhang Q, Tao H, Hong W-X (2011) New amphiphiles for membrane protein structural biology. *Methods* 55:318–323

Chapter 9

Building Model Membranes with Lipids and Proteins: Dangers and Challenges

James N. Sturgis

9.1 The Objectives: Why Make Artificial Membranes?

First, it is necessary to understand why biochemists wish to reconstitute membrane proteins into lipid vesicles and nonnative systems, and indeed there are several different reasons why this path could be taken. It is important to realize that there can be multiple objectives, as the shortcomings of a particular method for one objective might be an advantage for another. Historically, reconstitution has been aimed at both structural and functional studies. Structural studies can include not only the obvious determination of structure but also more subtle aspects of structure and the interplay of environment and structure, or the dynamics of the structure. Similarly, there are different aspects of function that can be targeted, many of which are accessible only in reconstituted systems.

Structural studies, such as high-resolution three-dimensional (3D) structural determination, can be conducted using 2D crystals formed in lipid membranes by using electron diffraction (e.g., Wang and Kühlbrandt 1992); however, these structures rarely compete with the structures obtained from X-ray diffraction (XRD) of 3D crystals. However, this approach can be particularly rewarding when the organization of lipids around a membrane protein is of interest. For example, initial studies of aquaporins by X-ray crystallography did not find any associated lipids (Murata et al. 2000); however, 2D crystals of aquaporin can be formed in several different lipids, and indeed in the structure obtained by electron diffraction (Gonen et al. 2005), many lipids were resolved in the crystal structure, including a few that were not in direct contact with the protein. Reconstitution for such structural studies, which can provide high quality 3D structures and information on the immediate environment, this is one possible reason to make artificial membranes. The objective is high protein density, and so reconstitutions are typically at lipid-to-protein ratio of less than 0.5 mg/mg.

For functional studies, reconstitution is often necessary because membrane proteins have transport activities. So ion channels are typically reconstituted into lipid bilayers for electrophysiological or transport measurements; for example, the KcsA

J. N. Sturgis (✉)

Engineering Laboratory of Macromolecular Systems, UMR-CNRS 7255 and
Aix-Marseille University, Marseille, France

e-mail: james.sturgis@imm.cnrs.fr

prokaryotic potassium channel was reconstituted at a lipid-to-protein ratio of about 1,000 mg/mg for Rb^{2+} transport measurements (Heginbotham et al. 1998). This approach allows detailed functional measurements. Similarly, transporters and pumps need to be incorporated into vesicles to measure transfer in and out of the vesicle, and if giant vesicles are used, even patch clamp measurements (Doeven et al. 2005). These functional studies on isolated systems in native-like environments can provide detailed information on mechanism and function, but rely on reconstitution into artificial membranes.

In general, the different reconstitution schemes aim to obtain reconstitution of a purified component into an environment as close to native as possible. However, there are constraints on this degree of nativeness; for example, for crystallization, it is necessary that the proteins are driven to organize themselves into 2D, and therefore, as for 3D crystallization, it is necessary to work at high concentrations and cause precipitation. Working at high concentrations with purified components means that the reconstitution yield is quite critical. At the other end of the scale, single-channel physiological methods require single channels and thus demand working at very low concentrations, and yield is thus much less critical. Equally, the lipid mixtures that are used are frequently much less complex and rich than the native membranes, varying from a single phospholipid to natural polar lipid extracts.

9.2 Current Reconstitution Methods

9.2.1 *The Standard Methods*

Reconstitution methods have been around for many years now and there is little mystery in the standard protocols. There are many published protocols that lead to different objectives, 2D crystallization, liposome formation, and single channel measurements. Figure 9.1 illustrates a typical reconstitution protocol based on detergent removal to form large unilamellar vesicles (LUVs). In the first few steps, multilamellar phospholipid vesicles are formed by hydration of a lipid film; during these steps, the organic solvent (typically chloroform–methanol) is removed. This step is important to avoid residual solvent that can lead to poorly sealed vesicles. After solvent removal and lipid dispersion in the aqueous buffer, the lipid mixture and the membrane protein of interest can be dissolved by detergent addition to obtain a homogeneous dispersion in the buffer. This aqueous dispersion is then driven to form proteoliposomes by detergent removal, typically by adsorption to polystyrene beads (Rigaud et al. 1997), or more classically, by dialysis. Finally, the proteoliposomes formed are purified by differential sedimentation, and/or flotation. This “standard” protocol leaves plenty of room for changing the lipids present, the detergents used, the protein added, and the buffers. However, as mentioned below, homogeneity can be a problem, and it is important that the lipids used be miscible and soluble in the detergent; furthermore, the protocol should be performed above the phase transition temperatures of the lipid mixture.

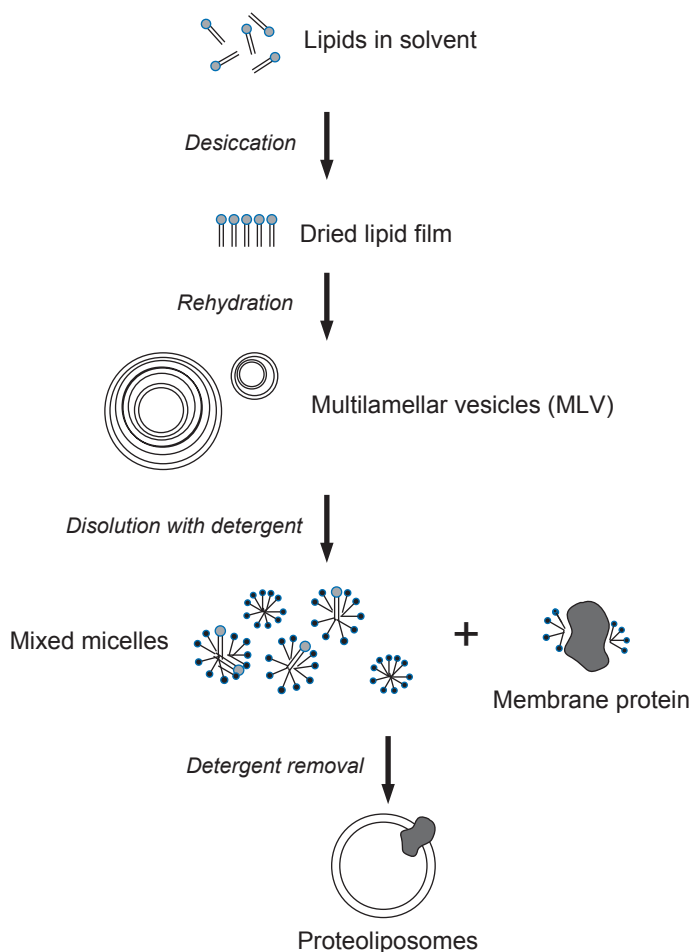


Fig. 9.1 Flow chart of a typical protocol for proteoliposome formation showing the main steps and cartoons of the objects present at different moments in the protocol

9.2.2 Variations on a Theme: Supported Lipid Bilayers and Giant Unilamellar Vesicles

Beyond the standard and frequently followed methods, there are extensions to the liposome reconstitution that have been used to generate either supported lipid bilayers (SLBs) or giant unilamellar vesicles (GUVs).

SLBs have become popular with the development of surface measurements, particularly atomic force microscopy (AFM; Levy and Milhiet 2013). With bilayers being formed on hard surfaces, for the membrane protein studies, the fusion of liposomes that rupture during adhesion to the surface is usually the pathway that is followed to produce these objects. This type of bilayer has the disadvantage that the interaction between one lipid monolayer and the support is very strong, possibly

perturbing the physicochemical properties. Thus, several groups have taken to producing tethered bilayers in which the lipids are attached to the support, leaving an aqueous space between the membrane and the support (Sumino et al. 2011). This is achieved by attaching the proteins and/or lipids to the surface with spacers.

An interesting development in this context is the advent of new supports containing holes and electrical contacts, and it is now becoming possible to use supported bilayers for functional measurements followed optically, electrically, and/or by AFM (Gonçalves et al. 2006). This approach allows transmembrane transport, for example, of protons, to be measured optically, while simultaneously observing the membrane structure using AFM.

Another direction that has been pursued over recent years is the formation of GUV-containing proteins. These vesicles provide the possibility of making measurements using conventional optical microscopes and even coupling this to micro-manipulation (Campillo et al. 2013). The majority of protocols reported for making giant vesicle-containing proteins proceed by drying proteoliposomes and then making giant vesicles by swelling or electro-formation on indium tin oxide (ITO) electrodes.

9.2.3 Problems with Standard Protocols

Two major problems have been identified with the currently popular methods that can render the resulting reconstituted systems difficult to use, or imperfectly mimetic of the natural system, which are: asymmetry and heterogeneity. We treat these problems individually subsequently.

9.2.3.1 Asymmetry

Native membranes are very asymmetric. Most proteins have a specific orientation in the membrane, and the lipids on the two leaflets of the membrane are usually different. In a few cases, proteins can be inserted into the membrane in both orientations, or be engineered to be inserted in both orientations (Rapp et al. 2007); however, this is exceptional. The membrane lipids are also asymmetric, and this asymmetry is maintained actively by a series of flippases and metabolic enzymes. However, the standard reconstitution protocol presented above typically produces symmetric membranes, i.e., with equal lipid composition in the two leaflets and proteins randomly oriented in the bilayer, with 50% of the proteins having each of the two possible orientations. This is clearly unnatural, sometimes useful, but often a problem.

9.2.3.2 Asymmetric Use

Even if the normal protocols are unable to yield membranes with asymmetrically inserted proteins, it is possible to use the proteins asymmetrically. This is typically

achieved by the provision of the substrate only to one side, typically the outside, of the reconstituted vesicles. For example, in studies of adenosine triphosphate (ATP)-binding cassette transporters, which use ATP as a substrate to pump various chemicals across the membrane, it is possible to add ATP to the outside of the reconstituted vesicles and measure net transport (Hofacker et al. 2007; Geertsma et al. 2008). This is possible since, though the membrane includes proteins with both orientations, only those proteins in which the ATP-binding site is exposed on the outside of the vesicle are active.

9.2.3.3 Heterogeneity

Ideally, a reconstituted membrane should be homogeneous, i.e., all parts of the membrane should be the same, and more importantly, different membrane fragments should be similar. While it is not possible to have all liposomes to contain the same number of molecules of each constituent, the aim is usually to have a relatively narrow distribution. Unfortunately, this is far from general; indeed, in the majority of cases that have been studied, the composition is often rather heterogeneous.

For example, in a work on light-harvesting complexes reconstituted into GUVs, we have observed lipid-to-protein ratios varying from 50 to 5,000 in different liposomes from the same preparation. In the same study, we also observed differences between two different proteins (core-complexes and light-harvesting complexes) present at a 1:1 ratio varying from 0.05 to 20 from liposome to liposome.

It should be emphasized that this heterogeneity can often apply to both the lipids and the proteins. So if complex lipid mixtures are used, it is far from certain that different lipids are present in the same proportion in various liposomes. This has been examined, for example, by the dispersion of phase transition temperatures from vesicle to vesicle in a preparation, where dispersions of 15° for this parameter can be observed (Baykal-Caglar et al. 2012). Thus, in principle, each liposome can have a completely different composition.

This heterogeneity is unfortunate for several reasons and illustrates that the current reconstitution methods are more art than science in that the system is incompletely mastered. First, it is often assumed that the distribution is narrow and the results are analyzed with this assumption—this, however, can considerably bias the interpretation. An example of this difficulty, albeit in a natural proteoliposome system rather than a reconstituted system, is illustrated by the discussions of the interpretation of kinetic data on chromatophores (photosynthetic vesicles) of the bacterium *Rhodobacter sphaeroides* (Crofts et al. 1998; Joliot et al. 1989; Lavergne et al. 1989), where two different interpretations of the data were proposed—one claiming specific protein associations and the other assigning the same results to a consequence of heterogeneity between vesicles. Second, this can often adversely affect the yield, that is, a certain fraction of the reconstituted system may be too lipid-rich or too protein-rich to meet the selection criteria for use or usefulness.

Several approaches have been suggested to alleviate problems of heterogeneity. Lipid heterogeneity appears to arise, at least in certain circumstances, from phase separations during solvent removal and/or detergent removal (not all lipids are equally soluble in all detergents). To reduce this, various suggestions have been made; an interesting recent proposal for the formation of GUVs is to form them from a damp lipid film (Baykal-Caglar 2012). In this protocol, the lipids are first made into small liposomes, using the rapid-solvent exchange protocol; these are then deposited on ITO slides and dried to a controlled humidity (55%) for over 22–25 h before the process of electro-formation as normal. The results presented show a considerable reduction in GUV-to-GUV heterogeneity.

Conceptually, a different approach to heterogeneity is to profit from it and measure liposomes individually, evaluating their composition and properties. This is feasible with GUVs, which are large enough to be visible, and individuals under microscopes are auscultated by different spectroscopic techniques, including marker fluorescence, intrinsic fluorescence, Raman spectroscopy, infrared absorption spectroscopy, coherent anti-Stokes Raman spectroscopy (CARS), or indeed any other spectroscopy adapted for imaging.

9.3 Recent Developments

9.3.1 *Creating Lipid Asymmetry*

To address the problems of lipid symmetry in vesicles, several methods have been developed to generate lipid asymmetry in initially symmetric vesicles, or more recently and excitingly, to form asymmetric vesicles.

Perhaps it should be noted here that lipid asymmetry is not infinitely stable, and indeed certain lipids are able to reorientate in the membrane relatively rapidly, for example, sterol flip-flop from one side of the membrane to the other takes place over a period of 1–2 h (Smith and Green 1974). Other lipids typically take many hours to exchange between the two leaflets, though this can be drastically increased by the presence of detergents (Pantaler et al. 2000) or certain proteins (scramblases; Bassé et al. 1996) in a lipid-dependent manner (Contreras et al. 2010). In natural membranes, asymmetry is therefore actively maintained, thanks to the asymmetric synthesis and degradation and a series of energy-dependent flippases dedicated to the reorientation of certain lipids (Poulsen et al. 2008). Clearly, the reconstitution of such a complex machinery is well beyond our current abilities.

9.3.1.1 Cyclodextrin Use

For several years, methyl- β -cyclodextrin (MBCD) has been used to modify the cholesterol content of membranes (Kilsdonk et al. 1995), initially native biological

membranes and later artificial membranes. This molecule is able to bind cholesterol inside its barrel-like structure, and so if added to the membrane, will selectively extract cholesterol from the membrane; conversely, if preloaded with cholesterol and then added to the membrane, it will be able to deliver the cholesterol to the membrane. Recently, it has been shown that the same molecule can be used to modify and tune the lipid content of the outer leaflet of a membrane beyond just altering the cholesterol content, considerably altering the content of different polar lipids in the outer leaflet of liposomes (Cheng et al. 2009).

In the novel protocol developed by Cheng et al. 2009, MBCD is charged with one type of lipid (sphingomyelin) and allowed to interact with preformed liposomes made from glycerophospholipids before being removed. This protocol was developed to prepare small unilamellar vesicles (SUVs) in which the outer leaflet was mainly sphingomyelin mixed with 1-palmitoyl-2-oleoyl-phosphatidyl choline (POPC) and the inner leaflet di-oleyl-phosphatidyl choline (DOPC), POPC, palmitoyl-oleoyl-phosphatidyl serine (POPS), and palmitoyl-oleoyl-phosphatidyl ethanolamine (POPE). That is a membrane with both asymmetric acyl-chains and asymmetric head groups. In a second step, cholesterol was introduced into the outer leaflet of these vesicles. Briefly, MBCD was charged with sphingomyelin by incubating at high concentration with multilamellar liposomes; this caused partial solubilization of the liposomes. To this suspension of lipids and MBCD was added a suspension of SUV-containing unsaturated glycerophospholipids. The MBCD catalyzes the exchange of sphingomyelin into the SUVs; the small vesicles with a modified outer leaflet were then recovered by centrifugation and size exclusion chromatography. In a second MBCD exchange step, cholesterol could be added complexed to MBCD and inserted into the vesicles. This protocol has now been extended and adapted to large (Cheng and London 2011) or giant liposomes (Chiantia et al. 2011).

The mechanism and energetics of cholesterol exchange have been examined by molecular dynamics simulations; these show how the dimers of β -cyclodextrins are able to adhere to the membrane surface with relatively high affinity. Bound dimers are then able to exchange lipids between the outer leaflet of the membrane and the interior of the cylindrical structure (López et al. 2013). This exchange appears to be controlled by both kinetic and energetic barriers.

Clearly, the use of this method to develop complex asymmetries is unrealistic; however, modification of the composition in one or two lipids is feasible, especially with the knowledge of the affinity of MBCD for different lipids. Unfortunately, there currently are little data on the selectivity of MBCD for different lipids.

9.3.1.2 Enzyme Use

An alternative approach has been to asymmetrically add purified enzymes to an initially symmetric membrane to selectively modify the lipids of the outer leaflet. This has been done with several different enzymes, for example, cholesterol oxidase (Ali et al. 2007) or phospholipase (Mouritsen 2011). The main disadvantage of

this approach is that the number of different reactions available and the asymmetries that can thus be developed are rather limited. However, the method has one important advantage over the use of MBCD in that the reactions can essentially go to completion, and thus it is possible to generate very highly asymmetric membranes.

9.3.1.3 Blowing Bubbles

An interesting new method for the generation of highly asymmetric lipid vesicles was recently proposed (Richmond et al. 2011). In this method, giant lipid vesicles are blown from a lipid film formed between two aqueous droplets. Initially, the aqueous droplets are in oil surrounded by a monolayer of lipids, either dissolved in the oil or delivered into the aqueous droplet as SUVs. Later, two monolayer-delimited aqueous droplets are brought into contact, along the lines developed by Bayley and coworkers for studying channels (Hwang et al. 2008). The two aqueous droplets are consequently separated by a bilayer, with one monolayer originating from each of the two original aqueous droplets, thus being potentially highly asymmetric. In the final step of formation, a micro-fluidic jet generated by the nozzle of an ink-jet printer deforms the bilayer to generate a giant vesicle.

This method, though technologically challenging, provides a clear route to highly asymmetric membranes, with asymmetric protein insertion and asymmetric content. It would seem unlikely, however, that the method is able to provide large samples easily, and there remains the difficulty of the inclusion of oil lenses in the resulting membranes. To date, it has been used for monotopic membrane proteins; however, the directional insertion of polytopic membrane proteins will remain challenging.

Perhaps one of the major interests of this approach is that it follows the path of sample engineering, in that each individual vesicle is manufactured in a more or less controlled manner. This can be opposed to the majority of the other vesicle-preparation methods which rely essentially on the self-assembly properties of the membrane components.

9.3.2 Protein Insertion into Preformed Vesicles

The two methods presented above address the problem of lipid asymmetry, concentrating on obtaining liposomes with controlled and asymmetric membranes. However, a major interest in forming artificial membranes is to study proteins in these membranes and often proteins inserted asymmetrically into the membranes. This remains challenging. Two main methods exist; first is the fusion of liposomes either with native membranes or with other liposomes containing asymmetrically inserted proteins, and second is the direct asymmetric insertion of proteins into liposomes.

9.3.2.1 Fusion

One possibility is to fuse liposomes with fragments of native membranes or asymmetric liposomes. This relies on obtaining oriented membrane vesicles containing the protein of interest. This has been achieved with certain bioenergetic systems, such as chromatophores from *R. sphaeroides*, or inverted inner membrane vesicles from *Escherichia coli* cells. Fusion between liposomes and these native membranes can then be driven by various methods. Freeze–thaw sonication (Pennoyer et al. 1985), calcium-induced fusion (Theiler and Niederman 1991), or fusion with detergent-destabilized vesicles (Rigaud et al. 1995) have all been reported. However, these methods are limited in that the proteins are not purified and only certain systems seem amenable.

9.3.2.2 Direct Asymmetric Insertion

The key to asymmetric insertion is protein insertion during detergent-mediated reconstitution. During early experiments, it was observed that this could give rise to highly asymmetric insertion (Helenius et al. 1981); this has since been much studied and the mechanisms are relatively well established. In this approach, liposomes destabilized by detergents are mixed with solubilized membrane proteins under conditions that avoid complete solubilization of the liposomes. The detergent is then removed and the proteins are inserted into the membrane from the outer leaflet in the kinetically most favorable orientation. That is, the orientation is determined by the energy necessary to move the hydrophilic portion of the protein through the membrane, and thus the largest, most hydrophilic part of the membrane protein will end up outside the vesicle. The protocol for this method is illustrated in Fig. 9.2, derived from the work of Rigaud and coworkers (Rigaud and Levy 2003, Geertsma et al. 2008).

The first part of the procedure is the preparation of liposomes. While different types and compositions of liposomes can be used, the procedure is typically performed using LUVs. It appears important that the diameter is greater than about 150 nm to avoid liposome fusion during the second step; it also appears to be important that the sample is as homogeneous as possible (see note on heterogeneity above). In a typical procedure, for the preparation of LUVs, purified lipids are dispersed in water as multilamellar vesicles (MLVs), as described in Fig. 9.1, and then a homogeneous LUV suspension is formed by extrusion through polycarbonate filters.

In the second step in the procedure, the liposomes (typically 1–10 mg/ml of lipid) are titrated with detergent in order to destabilize the lipid bilayer. The objective is to saturate the lipid bilayer with detergent. The progress can be conveniently followed by following absorption changes (in a spectrophotometer) or light-scattering changes (in a fluorimeter) as detergent is added, and tables (Rigaud and Lévy 2003) showing typical concentrations for different detergents and empirical approach should be used. It appears that the degree of asymmetry obtained during insertion depends on the amount of solubilization; the closer the system is to R_{sat} (see Fig. 9.2), the

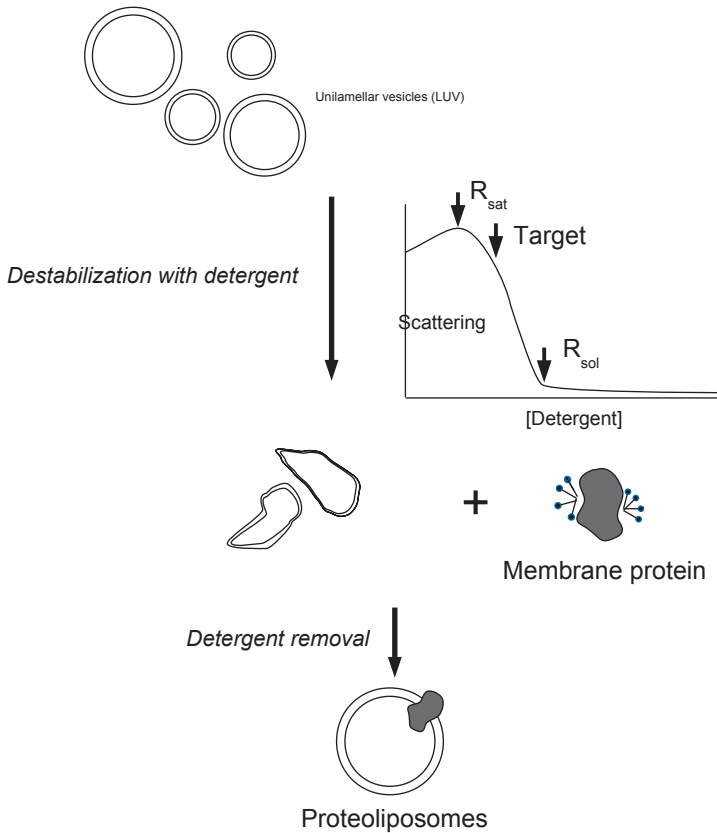


Fig. 9.2 Protocol for the asymmetric insertion of proteins into preformed liposomes, showing the two main steps. The graph shows the titration of liposomes with detergent and the positions of R_{sat} and R_{sol} , the detergent concentrations necessary to saturate the liposome with detergent, and to solubilize the liposome, respectively. For optimum yield and orientation, the detergent concentration for reconstitution needs to be above but close to R_{sat} ("Target" on figure). Sometimes, if care is not taken, the purified protein can contain a considerable quantity of extra detergent that can make optimization of the procedure difficult

higher the asymmetry of insertion, while if the system is totally solubilized (above R_{sol}), then insertion is symmetric. However, lower detergent concentrations can also adversely affect the yield of incorporation into the proteoliposomes. Thus, this titration step is critical for the final results obtained. The lipid–detergent system also takes some time to equilibrate; therefore, it should be allowed to equilibrate for at least an hour before moving on. The third step is to add the protein; here, there are several factors that should be borne in mind. First, for reproducibility, a homogeneous monodisperse protein solution should be used. Second, the amount of excess detergent brought to the mixture should be kept to a minimum, since this would perturb the careful titration performed in the second step; so practically, this implies using relatively concentrated protein solutions. Third, the reconstitution again takes

time and so the solution should be mixed and allowed to equilibrate typically for an hour. The final step is detergent removal; though several protocols can be used, removal by adsorption to polystyrene beads (Bio-Beads SM-2) is both efficient and rapid. In a typical protocol, methanol-washed, water-rinsed Bio-Beads are added directly to the detergent–protein–lipid solution at a ratio of 10 mg of beads per 1 mg of detergent for detergent removal. This is stirred for an hour before a second aliquot of beads is added, and then an hour later, a final aliquot of Bio-Beads is added, and the solution is allowed to equilibrate for an additional 2 h. Some particularly hydrophilic detergents need extra additions of Bio-Beads to be completely removed. The Bio-Beads are added in several stages in order to control the rate of detergent removal, which affects the final yield of proteoliposomes.

In a recent development of this method, Dezi and coworkers (Dezi et al. 2013) have adapted this method to allow the asymmetric incorporation of proteins into giant vesicles. This is important, since as mentioned above, heterogeneity is a point of concern and the use of visible objects allows the heterogeneity to be controlled and vesicles to be assessed individually. Furthermore, many experiments require being able to visualize the vesicles; this is the case with many of the example measurements made by Dezi et al. (2013).

To achieve oriented insertion of proteins into GUVs, the giant vesicles were made by electro-formation in the presence of sub-solubilizing concentrations of detergent. The concentrations used are above the critical micelle concentration for the detergent, thus assuring that there is sufficient detergent in the aqueous phase to maintain protein solubility and sufficient detergent in the membrane to aid oriented insertion. The giant vesicles that were thus formed were then incubated with a small volume of protein solution for several hours before detergent removal. Detergent was then removed by adding a few milligrams (a large excess) of either Bio-Beads or MBCD.

This method was demonstrated to allow successful oriented insertion of several membrane proteins, including the *E. coli* outer membrane proteins, FhuA and bacteriorhodopsin, into the GUVs. It also could be used to induce fusion of small native membrane vesicles with GUVs.

These direct insertion methods manage to obtain relatively high protein densities with good orientation. For example, with bacteriorhodopsin densities of about 2,000 μm^{-2} are reported, which correspond to a lipid-to-protein mole ratio of about 2,000, and even with this protein, the orientation appears almost completely asymmetric despite the small size of the hydrophilic domains. However, there is little quantitative information on the reconstitution yields of the proteins, which is expected to depend on the rate of detergent removal, with high yield being favored by slow removal; however, such information seems to be largely lacking from the literature.

Though these methods are relatively easy and effective, it is important to point out that it is the protein that determines the orientation and not the experimenter. Apparently, in each case, the larger, more hydrophilic part of the protein is left on the outside of the vesicle. Thus, for systems involving multiple proteins, it is hard to ensure that they all have the correct topology.

9.4 Conclusions

In this review, we have given an overview of several recent protocols for reconstitution of membrane proteins into a lipid environment. Many of these protocols require attention to detail for efficient implementation. Furthermore, reconstituted samples need careful analysis not only of average characteristics but also of heterogeneity and dispersion to ensure correct interpretation of results. For example, the phase diagrams of multiple lipid and detergent mixtures are particularly rich and varied, and this can easily result in unexpected demixing of components with a consequent heterogeneity in the final sample.

Reconstitution of membrane proteins into artificial membranes is now a relatively standard protocol for the assembly of unoriented systems into simple lipid mixtures. For several types of study, this is sufficient to obtain structural and functional information. However, for other purposes, it is useful to have more asymmetric and complex systems, possibly containing multiple lipids and multiple proteins. Much progress has been made recently on fabricating such asymmetric systems, allowing the formation of liposomes with asymmetrically distributed lipids and oriented membrane proteins. However, there is a long way to go before truly biomimetic systems can be constructed bottom up. The major current difficulty is studying multiple oriented proteins in such systems, as we do not yet have reliable and easily applicable protocols to impose the orientation during reconstitution.

References

- Ali MR, Cheng KH, Huang J (2007) Assess the nature of cholesterol-lipid interactions through the chemical potential of cholesterol in phosphatidylcholine bilayers. *Proc Natl Acad Sci U S A* 104:5372–5377
- Bassé F, Stout JG, Sims PJ, Wiedmer T (1996) Isolation of an erythrocyte membrane protein that mediates Ca^{2+} -dependent transbilayer movement of phospholipid. *J Biol Chem* 271:17205–17210
- Baykal-Caglar E, Hassan-Zadeh E, Saremi B, Huang J (2012) Preparation of giant unilamellar vesicles from damp lipid film for better lipid compositional uniformity. *Biochim Biophys Acta* 1818:2598–2604
- Campillo C, Sens P, Köster D, Pontani LL, Lévy D, Bassereau P, Nassoy P, Sykes C (2013) Unexpected membrane dynamics unveiled by membrane nanotube extrusion. *Biophys J* 104:1248–1256
- Cheng HT, London E (2011) Preparation and properties of asymmetric large unilamellar vesicles: interleaflet coupling in asymmetric vesicles is dependent on temperature but not curvature. *Biophys J* 100:2671–2678
- Cheng HT, Megha, London E (2009) Preparation and properties of asymmetric vesicles that mimic cell membranes: effect upon lipid raft formation and transmembrane helix orientation. *J Biol Chem* 284:6079–6092
- Chiantia S, Schwille P, Klymchenko AS, London E (2011) Asymmetric GUVs prepared by M β CD-mediated lipid exchange: an FCS study. *Biophys J* 100:L1–L3

- Contreras FX, Sánchez-Magrner L, Alonso A, Goñi FM (2010) Transbilayer (flip-flop) lipid motion and lipid scrambling in membranes. *FEBS Lett* 584:1779–1786
- Crofts AR, Guergova-Kuras M, Hong S (1998) Chromatophore heterogeneity explains phenomena seen in *Rhodobacter sphaeroides* previously attributed to supercomplexes. *Photosynth Res* 55:357–362
- Dezi M, Di Cicco A, Bassereau P, Lévy D (2013) Detergent-mediated incorporation of transmembrane proteins in giant unilamellar vesicles with controlled physiological contents. *Proc Natl Acad Sci U S A* 110:7276–7281
- Doeven MK, Folgering JHA, Krasnikov V, Geertsma E, van den Bogaart G, Poolman B (2005) Distribution, lateral mobility and function of membrane proteins incorporated into giant unilamellar vesicles. *Biophys J* 88:1134–1142
- Geertsma ER, Nik Mahmood NA, Schuurman-Wolters GK, Poolman B. (2008) Membrane reconstitution of ABC transporters and assays of translocator function. *Nat Protoc* 3:256–266
- Gonçalves RP, Agnus G, Sens P, Houssin C, Bartenlian B, Scheuring S (2006) Two-chamber AFM: probing membrane proteins separating two aqueous compartments. *Nat Methods* 3:1007–1012
- Gonen T, Cheng Y, Sliz P, Hiroaki Y, Fujiyoshi Y, Harrison S, Walz T (2005) Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* 438:633–638
- Heginbotham L, Kolmakova-Partensky L, Miller C (1998) Functional reconstitution of a prokaryotic K⁺ channel. *J Gen Physiol* 111:741–749
- Helenius A, Sarvas M, Simons K (1981) Asymmetric and symmetric membrane reconstitution by detergent elimination. Studies with Semliki-Forest-virus spike glycoprotein and penicillinase from the membrane of *Bacillus licheniformis*. *Eur J Biochem* 116:27–35
- Hofacker M, Gompf S, Zutz A, Presenti C, Haase W, van der Does C, Model K, Tampé R (2007) Structural and functional fingerprint of the mitochondrial ATP-binding cassette transporter Mdl1 from *Saccharomyces cerevisiae*. *J Biol Chem* 282:3951–3961
- Hwang WL, Chen M, Cronin B, Holden MA, Bayley H (2008) Asymmetric droplet interface bilayers. *J Am Chem Soc* 130:5878–5879
- Joliot P, Verméglio A, Joliot A (1989) Evidence for supercomplexes between reaction centers, cytochrome *c*₂ and cytochrome *bc*₁ complex in *Rhodobacter sphaeroides* whole cells. *Biochim Biophys Acta* 975:336–345
- Kilsdonk EP, Yancey PG, Stoudt GW, Bangerter FW, Johnson WJ, Phillips MC, Rothblat GH (1995) Cellular cholesterol efflux mediated by cyclodextrins. *J Biol Chem* 270:17250–17256
- Lavergne J, Joliot P, Verméglio A (1989) Partial equilibration of photosynthetic carriers under weak illumination: a theoretical and experimental study. *Biochim Biophys Acta* 975:347–355
- Levy D, Milhiet PE (2013) Imaging of transmembrane proteins directly incorporated within supported lipid bilayers using atomic force microscopy. *Methods Mol Biol* 950:343–357
- López CA, de Vries AH, Marrink SJ (2013) Computational microscopy of cyclodextrin mediated cholesterol extraction from lipid model membranes. *Sci Rep* 3:2071
- Mouritsen OG (2011) Lipids, curvature, and nano-medicine. *Eur J Lipid Sci Technol* 113:1174–1187
- Murata K, Mitsuoka K, Hirai T, Walz T, Agre P, Heymann JB, Engel A, Fujiyoshi Y (2000) Structural determinants of water permeation through aquaporin-1. *Nature* 407:599–605
- Pantaler E, Kamp D, Haest CW (2000) Acceleration of phospholipid flip-flop in the erythrocyte membrane by detergents differing in polar head group and alkyl chain length. *Biochim Biophys Acta* 1509:397–408
- Pennoyer JD, Kramer HJ, van Grondelle R, Westerhuis WH, Amesz J, Niederman RA (1985) Excitation energy transfer in *Rhodospseudomonas sphaeroides* chromatophore membranes fused with liposomes. *FEBS Lett* 182:145–150
- Poulsen LR, López-Marqués RL, Palmgren MG (2008) Flippases: still more questions than answers. *Cell Mol Life Sci* 65:3119–3125
- Rapp M, Seppälä S, Granseth E, von Heijne G (2007) Emulating membrane protein evolution by rational design. *Science* 315:1282–1284

- Richmond DL, Schmid EM, Martens S, Stachowiak JC, Liska N, Fletcher DA (2011) Forming giant vesicles with controlled membrane composition, asymmetry, and contents. *Proc Natl Acad Sci U S A* 108:9431–9436
- Rigaud JL, Lévy D (2003) Reconstitution of membrane proteins into liposomes. *Methods Enzymol* 372:65–86
- Rigaud JL, Pitard B, Levy D (1995) Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *Biochim Biophys Acta* 1231:223–246
- Rigaud J-L, Moser G, Lacapere JJ, Olofsen A, Levy D, Ranck JL (1997) Bio-Beads: an efficient strategy for two-dimensional crystallization of membrane proteins. *J Struct Biol* 118:226–235
- Smith RJ, Green C (1974) The rate of cholesterol ‘flip-flop’ in lipid bilayers and its relation to membrane sterol pools. *FEBS Lett* 42:108–111
- Sumino A, Dewa T, Takeuchi T, Sugiura R, Sasaki N, Misawa N, Tero R, Urisu T, Gardiner AT, Cogdell RJ, Hashimoto H, Nango M (2011) Construction and structural analysis of tethered lipid bilayer containing photosynthetic antenna proteins for functional analysis. *Biomacromolecules* 12:2850–2858
- Theiler R, Niederman RA (1991) Localization of chromatophore proteins of *Rhodobacter sphaeroides*. I. Rapid Ca(2+)-induced fusion of chromatophores with phosphatidylglycerol liposomes for proteinase delivery to the luminal membrane surface. *J Biol Chem* 266:23157–23162
- Wang D, Kühlbrandt W (1992) Three-dimensional electron diffraction of plant light-harvesting complex. *Biophys J* 61:287–297

Chapter 10

Analytical Ultracentrifugation and Size-Exclusion Chromatography Coupled with Light Scattering for the Characterization of Membrane Proteins in Solution

Aline Le Roy, Cécile Breyton and Christine Ebel

10.1 Introduction

Membrane proteins are physiologically embedded in a lipid bilayer. Detergents are used to extract them from the membrane to allow their purification and characterization. Protein association state and sample homogeneity, together with the size of the membrane protein–detergent complex, have to be determined in view of functional and structural studies. Detergents are characterized by their critical micelle concentration (*cmc*) above which detergent monomers auto-assemble to form small, compact aggregates of aggregation number N_{agg} . Above the *cmc*, detergent will bind to the hydrophobic transmembrane domain of the proteins, thereby solubilizing them. The simplest sample of a membrane protein therefore comprises the protein–detergent complex and free detergent micelle, in addition to monomers of detergent. Protein and detergent will be labeled with subscript *p* and *d* below. The “real” sample is often more complex (Fig. 10.1). The membrane protein is possibly glycosylated to an extent often not well defined. Lipids, which often are essential for stability and/or function, and/or hydrophobic cofactors may be solubilized in the free micelles or specifically bound to the membrane protein. In the following

C. Ebel (✉) · C. Breyton

Institute of Structural Biology, French National Centre for Scientific Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France
e-mail: christine.ebel@ibs.fr

A. Le Roy

Institute of Structural Biology and The European Molecular Biology Laboratory, Integrated Structural Biology Grenoble, French National Centre for Scientific Research, Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

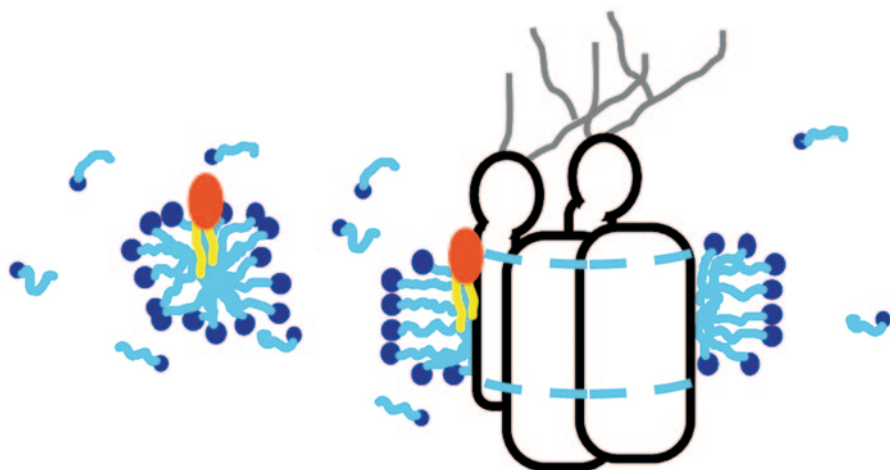


Fig. 10.1 Drawing of the protein sample solubilized in detergent. *Black lines* delimit the membrane protein, which may consist of several similar or no subunits. *Gray lines* represent glycosylation. Lipids are represented with a *red hydrophilic head* and *yellow hydrophobic tails*. Detergent molecules are represented with a *dark blue hydrophilic head* and a *light blue hydrophobic tail*

sections, we will consider a two-component protein–detergent complex, glycosylation and lipid binding being possibly integrated in the analysis (see, e.g., Dach et al. 2012).

Analytical ultracentrifugation (AUC) sedimentation velocity (SV) and size-exclusion chromatography coupled with static and dynamic light scattering, absorbance, and refractive index detection (SEC/MALS) are two techniques that combine separation and analysis, in an absolute manner, of the mass and size of macromolecules in solution. We present here how they can be applied to the study of solubilized membrane proteins to determine the presence of aggregates, the association state of the protein, an often controversial question, and the amount of bound surfactants. AUC also allows the characterization of the auto-associative properties of detergent or surfactant themselves, which can form aggregates of varying geometry, as rods (see Breyton et al. 2009).

We will pay attention here to the complementarity and differences between these two techniques, since they provide nearly the same information without too much details: Over the years, detailed protocols or specific methods have been published. For AUC, see, e.g., Lebowitz et al. (2002); le Maire et al. (2008); Ebel (2011); Le Roy et al. (2013). For SEC/MALS, see, e.g., Hayashi et al. (1989); Slotboom et al. (2008). Table 10.1 presents a comparison of some features related to the experimental requirements, and Table 10.2 presents a comparison of some features related to their analysis.

This chapter describes and compares the two techniques, for the case of a sample consisting of a membrane protein solubilized with detergent. The example of the

Table 10.1 AUC SV and SEC/MALS experimental requirements

	AUC SV	SEC/MALS
Molecular parameters for species separation	s , i.e., M_b/R_H	R_H
Experimental parameter for optimizing separation	Angular velocity; solvent density	Type of chromatographic column
Minimal equipment	Analytical ultracentrifuge (AUC) equipped with UV and interference detectors (A/J), with rotor and cells	HPLC; SEC column; LS, RI, and UV online detectors
Equipment, optional	Fluorescence detector for AUC; density-meter; viscosity-meter	DLS detector; thermostated sample changer; column oven; thermostated collector of eluted fractions
Software for data acquisition	From the equipment supplier	From the SEC/LS supplier
Duration of the experiment	Typically overnight	One hour per sample, after one night of column equilibration
Easiness of the setup of the experiment	Requires training	Easy
Caution regarding instrumentation	Instructions of the supplier to be followed	Take care of the column and avoid release of particles from the matrix in the detectors by centrifugation or filtration of the samples, using a constant flow rate
Caution regarding analysis	Temperature equilibration before the experiment; absorbance of interest below 1.2 in the cell	Choice of the appropriate column for a good separation; perfect equilibration with the eluant
Number of samples per experiment	With eight-hole/four-hole rotors: 7/3 samples with A/J detection and 14/6 samples for fluorescence detection	Different samples may be stored in a sample changer, and loaded automatically every 40–60 min
Sample volume	60 μL for $l=1.5$ mm; 120 μL for $l=3$ mm; 450 μL for $l=12$ mm	20–200 μL
Sample concentration	Absorbance (typically 280 nm) in the AUC: 0.1–1.2; Interference: conc. >0.1 mg/mL; fluorescence ^a : conc. of 1 pM–14 μM	Typically 2 mg/mL
Solvent volume	Some mLs with and without detergent	1 L without detergent; 200 mL with detergent
Solvent requirement	>100 mM salt recommended; limited absorbance	>100 mM salt recommended; pH compatibility with column; limited absorbance
Versatility of solvent composition	Large	Large but each solvent requires \approx 1 day equilibration
Sample dilution during the experiment	No	Yes

Table 10.1 (continued)

	AUC SV	SEC/MALS
Possibility of combination of detectors	UV: 3 λ s in the range 230–700 nm can be combined with interference. Fluorescence detection cannot be combined	3 λ s from the UV detector can be combined with RI and DLS
Temperature range	4–40 °C ^b	4–30 °C

^a Fluorescein-probe measured with Aviv fluorescence detector

^b For recent models equipped with a turbo pump; previously: 4–20 °C

Escherichia coli outer-membrane ferrichrome transporter (FhuA) is described in more details. Note that the characterization by AUC and other techniques of FhuA solubilized in n-dodecyl- β -D-maltopyranoside (DDM) was described in Boulanger et al. 1996.

10.2 AUC and SEC/MALS Are Separative Techniques

10.2.1 AUC

Macromolecules subjected to a large centrifugal field—up to 300,000 g in the commercially available analytical ultracentrifuge—will redistribute in solution. Particle distribution is measured at different times. The two very usual types of experiment are SV and sedimentation equilibrium (SE). SE uses a relatively low centrifugal field, and equilibrium condition is obtained after typically 24 h. The resulting concentration profiles, measured as a function of the radial position, in SE, do not depend on the shape of the species in solution—analysis is mathematically very simple and yields signal-average buoyant molar mass (M_b , defined below)—which are interpreted in terms of thermodynamic association models. However, the samples have to be extremely well defined and stable (duration of the experiments is typically some days). While extremely useful for specific studies or to ascertain specific results, it will not be described more here. We will focus on SV, which is much more versatile.

SV experiments are performed at high centrifugal fields. SV profiles, measuring a signal related to concentration as a function of the radial distance, are measured at different times, for example every 10 min. The experiment duration is typically a couple of hours. Different types of macromolecules sediment at different speeds. This allows the study of complexes and heterogeneous and/or interacting systems. The different detections (absorbance and interference, or fluorescence) provide a complementary means to investigate multicomponent systems (the sample of membrane protein), since the different components (detergent, protein) will be detected independently. Some experimental requirements are presented in Table 10.1.

Table 10.2 Comparison of AUC SV and SEC/MALS data analysis

	AUC SV	SEC/MALS
Software for analysis	Packages are freely available on the web	From the equipment supplier
Easiness of data analysis	Homogeneity: easy; composition: may be tricky	Easy
Limit	Analysis with floating detergent is difficult ^a ; analysis with interference with complex solvent (e.g., with glycerol) may be difficult ^b	Poor separation of the different species; possible interaction with the SEC matrix; solvent exchange to the elution solvent; possible delipidation; dilution leading to the dissociation of complexes
Parameters required for analysis	Solvent: ρ , η ; protein, detergent: $E_{0.1\%, 280^\circ}$, $\partial n/\partial c$	Solvent: ρ , η ; protein, detergent: $E_{0.1\%, 280^\circ}$, $\partial n/\partial c$
Other optional or required parameters	Detergent: M , N_{agg} , cmc . Protein: R_{H} , M , glycosylation, bound lipids	Detergent: M , N_{agg} , cmc . Protein: M , glycosylation, bound lipids
Raw data	A set of SV profiles obtained at different times of centrifugation for each of the detection	Superposition of the elution profiles obtained with the different detectors
Transformed data	For the set of SV profiles, for each optics: $c(s)$	
Evidence of aggregates	More or less easy depending of their size/amount	Very easy
Evidence of free detergent micelle	Easy if the micelle sediments, i.e., depends on \bar{v}_d	Easy if the micelle is separated from the protein–detergent complex
Results from signal intensities	For a selected peak in the $c(s)$ analysis; B_d ; c (mg/mL)	At each volume of elution: B_d ; M_p ; M_{pd} under the elution peak: amount (g) and mean B_d ; M_p , M_{pd}
Evidence of interacting systems	From the superposition of $c(s)$ at various conc.: s is constant for each noninteracting systems, mean s increases with conc. for interacting systems	From the plot of the M_p ; M_{pd} versus the elution time: values validate the analysis in terms of species Note the complex may dissociate during elution
Molar mass determination	For a homogeneous sample, analysis in noninteracting species provides M_b and s . M_b is related to the M_p and B_d ; for a heterogeneous sample of noninteracting species, the s -value is analyzed in combination, with other knowledge: B_d , R_{H} , M_b , from AUC or other techniques	See results from the signal intensities

The table refers to a protein–detergent complex. For the case of glycosylated protein and/or for bound lipid or other ligand, the additional components have to be considered in the protein or detergent part. Symbols are defined in the text

^a LDAO is an example of a detergent that floats

^b The sedimentation of glycerol is detected in interference, and we use boundary-forming centerpieces to allow a perfect subtraction of solvent contributions to sedimentation (Le Roy et al. 2013)

10.2.2 SEC/MALS

Size exclusion chromatography (SEC) is a common, simple, robust, and rapid laboratory technique allowing the separation of particles according to their hydrodynamic radii, R_H . The R_H of the protein–detergent complex can be easily estimated by SEC by comparison with globular water-soluble protein standards. The resolution is in principle lower as compared to SV, since the latter separates particles of same density and shape as $M^{2/3}$ (see the equations below) and SEC as $M^{1/3}$, with M the molar mass. Molar masses of membrane proteins can be estimated upon elution in a rigorous way, as well as the amount of bound detergent, when SEC is coupled “online” with MALS, refractive index (RI), and UV detections. Coupling SEC with dynamic light scattering (DLS) provides estimates of R_H without calibration. We here name SEC/MALS, the assembly of SEC and of these detectors.

10.3 Detection in AUC and SEC/MALS

10.3.1 Absorbance in AUC and SEC/MALS

Absorbance, A , is related to optical path length, l (cm), and macromolecule concentration, c (mg/mL), via the macromolecule extinction coefficient, $E_{0.1\%}$ ((mg/mL) $^{-1}$ cm $^{-1}$):

$$A = E_{0.1\%}lc. \quad (10.1)$$

10.3.2 Interference in AUC and Estimate of Bound Detergent

Interference detection measures the variation of the concentration of the particles, in fringe shifts, ΔJ . ΔJ is related to c (g/mL) via the refractive index increment, $\partial n/\partial c$ (mL/g), with λ being the wavelength (cm) of the laser light:

$$\Delta J = [(\partial n/\partial c)/\lambda]lc. \quad (10.2)$$

For a protein–detergent complex comprising the protein at concentration c_p and bound detergent in amount B_d (g/g), if protein and detergent are characterized by different $E_{0.1\%}$ over $\partial n/\partial c$ ratios, combining A and ΔJ allows determining c_p and B_d , from the knowledge of $E_{0.1\%}$ and $(\partial n/\partial c)$ for both the protein and the detergent.

10.3.3 Refractive Index in SEC/MALS and Estimate of Bound Detergent

The difference, Δn , in refractive index between the solution and the elution buffer is related to c (g/mL):

$$\Delta n = (\partial n/\partial c)c. \quad (10.3)$$

For a protein–detergent complex, similarly to above, combining A and Δn allows determining c_p and B_d , from the knowledge of $E_{0.1\%}$ and $\partial n/\partial c$ for both the protein and the detergent.

10.3.4 Fluorescence in AUC

Fluorescence detection (wavelength of excitation 488 nm; of emission 535 ± 30 nm) in AUC has also been available over a couple of years (Aviv Biomedical, Lakewood, NJ, USA). The signal is proportional (in arbitrary units) to the amount of fluorescent probes, but not quantitative, since fluorescence intensity is amplified differently for each cell in the ultracentrifuge.

10.3.5 Light Scattering in SEC/MALS

According to the Rayleigh–Debye–Gans model, the intensity, I , of the light scattered (LS) by the solution is directly proportional to the concentration c (g/ml), the molar mass M (g/mol), and the square of the RI increment $\partial n/\partial c$ of the macromolecule in solution:

$$I = kcM(\partial n / \partial c)^2, \quad (10.4)$$

with

$$k = 4\pi n_0^2 / \lambda^4 N_A, \quad (10.5)$$

where k depends on the refractive index n_0 of the solvent of the wavelength λ of the laser, N_A being Avogadro's number. For heterogeneous protein samples, with c the total concentration, the weight-average molar mass is obtained. For macromolecular sizes >20 nm, the scattered intensity decreases with the diffusion angle θ and M is obtained by extrapolation of I at angle 0. This is why light-scattering intensity is measured at different angles (MALS stands for multi-angle laser scattering).

The scattered intensity of a mixture of species is dominated by the signal of the larger species, because the intensity is proportional to the product of the concentrations and molar masses. Even a very small amount of large aggregates avoids the interpretation of static light scattering and DLS. The combination of SEC with LS allows the analysis of a homogeneous eluate without contaminating aggregates.

10.3.5.1 Static Light Scattering and Measure of Molar Masses in SEC/MALS

MALS analyzes the time-averaged (1 s) scattered light intensity, I , and allows determining molar mass M if the concentration c is known (Eq. 10.4) (and radius of gyration if $R_g >20$ nm, not developed here).

In SEC/MALS, the concentration of soluble proteins is determined online, from the RI detector (Eq. 10.3). It is thus possible to determine M from the intensity of the scattered light (Eq. 10.4). For a membrane protein, the concentrations of protein and detergent, within the membrane protein complex, i.e., c_p and B_d , are estimated from the UV and RI signals and combined to the MALS signal to estimate the protein molar mass, M_p , the mass of bound detergent, and the mass of the protein–detergent complex.

10.3.5.2 Dynamic Light Scattering

DLS analyzes the fluctuations of the scattered light intensity, I , as a function of time in terms of shape of macromolecule (see Sect. 10.4.7).

10.4 Theoretical Background for AUC and SEC/MALS

10.4.1 Sedimentation Velocity Theoretical Background

The sedimentation is related, for each solute, to the sedimentation coefficient, s , which translates the velocity of the particle, v , in response to the centrifugal field, $\omega^2 r$ ($s = v/\omega^2 r$; where ω is the angular velocity, r the radial position), and to the diffusion coefficient, D , which translates the capacity of the particle to diffuse in response to concentration gradient. For particles in interaction, it is necessary to introduce parameters translating the fact that some of the particles “disappear” while others “appear” (kinetic constant of association/dissociation, equilibrium constant). Lamm’s equation describes the evolution of the concentration, c , of one type of ideal particles (without interaction) as a function of time, t , and the radial position, r

$$(\partial c / \partial t) = -1/r \partial r (cs \omega^2 r - D \partial c / \partial r) / \partial r. \quad (10.6)$$

10.4.2 Analysis of SV in Terms of Noninteracting Species (NIS)

The position and shape of the boundary, for an ideal solution of noninteracting species, are related to their sedimentation, s , and diffusion, D , coefficients. Analysis of the set of sedimentation profiles can be realized globally, thanks to numerical solutions of the Lamm equation. When the solution contains a limited number (up to three) of different independent (noninteracting) species, analysis in term of s and D can be done by least-squares boundary modeling of sedimentation velocity data. However, it is important to assess the hypothesis of the number of species and their noninteractivity, using the $c(s)$ analysis presented below for different samples at different concentrations. Indeed, a boundary may represent the sedimentation of, e.g., two types of particles that sediment only slightly differently. If there is a rapid equilibrium of association between the species in the sample, the boundary is then a

reaction boundary, which moves at a velocity intermediate between s of the smaller species and s of the larger complex, and whose shape may be simple or complex (Schuck 2010a, b; Zhao et al. 2011). For concentrated solutions ($> \text{mg/mL}$), nonideality due to excluded volume effects changes the velocity and shape of the boundary (Solovyova et al. 2001; Salvay et al. 2010). In these cases, the estimate of s , and D from boundary spreading, will be obviously wrong.

10.4.3 Analysis of SV in $c(s)$

The $c(s)$ method proposed in 2000 (Schuck 2000) has constituted a major advance for the analysis of SV experiments. The sample is described as an assembly of non-interacting particles, e.g., $N=200$ particles, defined by their minimum and maximum s values. These particles are assumed to have the same density and shape (same partial specific volume \bar{v} and frictional ratio ff_{min}), which provides a way to link an approximate D -value to each s -value, and thus efficiently account for diffusion broadening of the SV profiles (ff_{min} can be fitted, which compensates for the possibly erroneous value of \bar{v}). It results in a high-resolution continuous distribution of sedimentation coefficients, $c(s)$. Detergent micelles and protein–detergent complex boundaries are most often separated.

10.4.4 The Sedimentation Coefficient and Buoyant Molar Mass

For one noninteracting species behaving ideally, the sedimentation coefficient, s , depends on the particle buoyant molecular weight, M_b , and the hydrodynamic radius R_H . s also depends on the solvent density and viscosity, ρ and η , respectively. The Svedberg's equation

$$s / D = M_b / RT, \quad (10.7)$$

with R being the gas constant and T the temperature, can be written as:

$$s = M_b / (N_A 6\pi\eta R_H), \quad (10.8)$$

with

$$M_b = M(1 - \bar{v}\rho). \quad (10.9)$$

When different species are sedimenting, including the case of equilibrium, mean s -values are measured that represent weight-average values. The species concentrations in weight/volume unit are weighted by their “extinction coefficient,” and thus the mean s may differ for data obtained using interference and absorbance, in the case of multicomponent systems, e.g., samples of membrane proteins.

For a protein–detergent complex, M_b can be written, with p and d the subscripts for protein and detergent, respectively, and B_d the amount of bound detergent (g/g):

$$M_b = M_p(1 - \bar{v}_p \rho) + B_d(1 - \bar{v}_p \rho).$$

Detergent can float or sediment. Numerical values of \bar{v}_d can be found in le Maire et al. (2000); Ebel et al. (2007); le Maire et al. (2008); Ebel (2011).

10.4.5 *The Diffusion Coefficient, Hydrodynamic Radius, and Frictional Ratio*

For one noninteracting particle, the Einstein–Stokes relation relates D and R_H :

$$D = RT / N_A 6 \pi \eta R_H. \quad (10.10)$$

The frictional ratio ff_{\min} is the ratio of R_H to the radius, R_{\min} , of the anhydrous volume. For globular, compact particles, a typical value for ff_{\min} is 1.25 (Manon and Ebel 2010; Salvay et al. 2012).

10.4.6 *Static Light-Scattering Intensity*

See above subsection 10.3.5.1.

10.4.7 *Dynamic Light Scattering*

From the time, t , and fluctuations of the scattered light intensity, $I(t)$, the auto-correlation function $g^{(2)}(\tau)$, with τ a time interval, is

$$g^{(2)}(\tau) = \langle I(t)I(t + \tau) \rangle / \langle I(t) \rangle^2. \quad (10.11)$$

For a homogeneous solution

$$g^{(2)}(\tau) = 1 + \beta \exp(-2D q^2 \tau), \quad (10.12)$$

where β is a geometrical factor, and, θ the scattering angle:

$$q = (4\pi n_0 / \lambda) \sin(\theta / 2). \quad (10.13)$$

DLS coupled with SEC allows the determination of the diffusion coefficient D , and thus hydrodynamic radius R_H (Eq. 10.10), according to the quite simple equation (10.12), because the eluate is monodispersed. This is in contrast to the analysis usually made in terms of size distribution when measurements are performed in a batch mode. In the case of a solubilized membrane protein, free detergent micelle and protein–detergent complex should, however, contribute. The analysis is thus relevant only if the free detergent signal measured in the buffer is negligible compared to the detergent–protein complex.

10.5 Steps for AUC and SEC/MALS analysis

Table 10.2 compares some features concerning AUC SV and SEC/MALS data analysis. Some complementary details are given below.

10.5.1 Steps for AUC Analysis

10.5.1.1 From the Superposition of the $c(s)$

The analysis of the $c(s)$ (see Sect. 10.4.3) and the superposition of the $c(s)$ obtained using different optics, or from different samples, with different experimental parameters, e.g., protein or detergent concentration and buffer composition, allow to:

1. Characterize sample homogeneity by the number of peaks in the $c(s)$
2. Evaluate if there are association equilibria, from the comparison of the $c(s)$ for samples at different protein concentrations, since a complex sediments faster than the dissociated component (see, e.g., Josse et al. 2002)
3. Determine s easily for a protein–detergent complex that is not involved in an association equilibrium
4. Decide for a homogeneous protein–detergent complex that is not involved in an association equilibrium, an analysis in terms of noninteracting species (see below)
5. Determine a mean s -value: This allows to estimate the dissociation equilibrium constant if proteins are involved in such an equilibrium (see, e.g., Echaliier et al. 2013, for a soluble protein)
6. Estimate the amount of bound detergent within the protein–detergent complex from the integration of the $c(s)$ peak from data acquired in absorbance and interference
7. Estimate the amount of free detergent micelles from the integration of the $c(s)$ signals in fringe (reference channel being filled with buffer without detergent; Salvay and Ebel 2006)

8. Interpret the value of s , using estimates of the bound detergent and the protein–detergent complex hydrodynamic radius, or assuming a reasonable shape (i.e., f/f_{\min} value) for the complex.

10.5.1.2 Noninteracting Species Analysis

An analysis in terms of noninteracting species, leading to independent determination of s and D , and thus M_b and R_H , can also be made for a system containing a limited number of rather different noninteracting species, e.g., micelle and protein–detergent complex (see Sect. 10.4.2).

10.5.1.3 Complementary Analysis

A large variety of complementary analysis, not detailed here, are possible, including:

- Varying the solvent density by exchanging the solvent with a solvent containing D_2O will change the buoyant properties of the protein–detergent complex and help to define the association state of membrane proteins within the complex (Nury et al. 2008; Le Roy et al. 2013).
- Datasets obtained with different optical systems (e.g., absorbance interference) may be analyzed globally to derive the relative concentration (in micromolar units) of detergent and protein in the $c(s)$ analysis (Balbo et al. 2005; Salvay et al. 2007).
- Complex SV simulation for interacting systems with defined kinetics and equilibrium association constants, taking into account different extinction coefficients can be made and used for complex data analysis of interacting systems, from datasets obtained at different concentrations, with different optical systems (Schuck 2010a, c; Zhao et al. 2011). This has not yet been applied to protein–detergent interacting complexes.
- The highly specific and sensitive fluorescent detection allows to specifically follow complex formation of green fluorescent protein (GFP)-fused or fluorescein isothiocyanate (FITC)-labeled proteins with unlabeled, possibly membrane protein, partners, even in very complex mixtures of unlabeled proteins (under investigation).

10.5.1.4 Systems that Are Difficult to Analyze

Systems difficult to analyze are, concerning the analysis in $c(s)$:

- When the detergent floats (e.g., protein solubilized in lauryldimethylamine N-oxide (LDAO), which has a partial specific volume above the solvent density), $c(s)$ cannot be obtained
- When the solvent is complex, i.e., comprises solvent components, their sedimentations are detected in interference

- When the solvent or the detergent, such as Triton X-100, absorbs strongly light at 280 nm (Nury et al. 2008)
- When the shape and partial specific volume of the different species in the sample differ strongly
- When concentrations are larger than 2 mg/mL (repulsive interactions affect the shape of the boundaries), and the sample heterogeneous.

The analysis in terms of M_p and B_d is difficult:

- When the protein has an unknown glycosylation level or lipid content
- For distinguishing close number of subunits: monomers and dimers are much more easily distinguished than tetramers and pentamers
- When extinction coefficients are ill defined, B_d from absorbance and interference optics are ill defined
- When association constant is weak, and concentrations cannot be explored above and below the dissociation constant, K_d
- Concentrations above 2 mg/mL are associated with weak repulsive interparticle interaction, inducing in particular a decrease of the sedimentation coefficient when compared with the diluted system. Homogeneous systems are easily analyzed (Solovyova et al. 2001; Salvay et al. 2010). For interacting systems, weak interparticle interaction may mask complex formation (Irimia et al. 2003).

10.5.2 Steps for SEC/MALS Analysis

10.5.2.1 From the Superposition of the Chromatograms

The analysis is easy.

10.5.2.2 Systems that Are Difficult or Impossible to Analyze

- Detergent micelle and protein–detergent complexes should be well separated. This is not the case, for example, of small proteins, because their size is similar to that of the micelles.
- Separation between the different protein–detergent complexes has to be efficient.
- Analysis in a two-component mode: Proteins with an unknown glycosylation level or lipid content are more difficult to analyze (as in SV).
- Distinguishing close number of subunits is hard: Monomers and dimers are much more easily distinguished than tetramers and pentamers (as in SV).
- When extinction coefficients are ill defined, B_d from absorbance and RI optics are ill defined (as in SV): An approximate molar mass for the complex is however obtained from RI and LS detectors (and not absorbance), considering a mean $\partial n/\partial c$ value.

10.6 Example of a Membrane Protein, FhuA

10.6.1 AUC: FhuA in F₆-DigluM

10.6.1.1 Aim of the Experiment

Our aim was to investigate the homogeneity of FhuA after a surfactant exchange step to a fluorinated surfactant F₆-DigluM (Abla et al. 2012).

10.6.1.2 Sample, AUC Experiment, and Parameters for the Analysis

One-milliliter FhuA (0.2 mg/mL) in 20 mM Tris–HCl pH 8 (buffer A) with 0.05 % LDAO (Flayhan et al. 2012) was loaded onto a 150- μ L nickel–nitrilotriacetic acid (NiNTA) home-packed column (Qiagen resin) equilibrated in the sample buffer. Surfactant exchange was performed by washing the column with 2 mM F₆-DigluM in buffer A, and the protein was eluted with 1 mL buffer A with 4 mM F₆-DigluM and 200 mM Imidazol, in two fractions of 0.5 mL. The absorbance at 280 nm of the fraction we used was 0.3, estimated with a reference buffer containing 100 mM Imidazol (200 mM provided negative values of absorbance, the solvent concentration is thus in between that of the washing and elution buffers). The reference solvent in AUC was thus chosen as buffer A with 100 mM Imidazol, without F₆-DigluM. Indeed, the reference channel is filled in general with solvent without detergent (see, however, e.g., Dach et al. 2012; Le Roy et al. 2013, for a study and a protocol using reference solvent containing detergent). The FhuA sample was used for SV experiments conducted in an XLI analytical ultracentrifuge (Beckman, Palo Alto, CA, USA), using an ANTi-60 rotor, at 42,000 rpm and 20 °C overnight, using double-channel center pieces (Nanolytics, Potsdam, Germany) of 12 mm optical path length (loaded volume: 420 μ L) equipped with sapphire windows. The parameters that were used for the analysis are tabulated in Table 10.3. The parameters characterizing usual solvents can be determined from tabulated data with the program SEDNTERP (available free at: <http://sednterp.unh.edu/>), those for the protein from amino acid composition with the same program and SEDFIT (available free at: <http://www.analyticalultracentrifugation.com>). The latter also allows the calculation of $\partial n/\partial c$; $\partial n/\partial c=0.187$ mL/g is typical of membrane proteins (Hayashi et al. 1989). For detergents, values have to be measured or obtained from literature or technical notices (see e.g. <http://www.anatrace.com>). Some values have been tabulated in le Maire et al. 2000.

10.6.1.3 $c(s)$ Analysis

Data analysis were done with the program SEDFIT (available free at: <http://www.analyticalultracentrifugation.com>). Figure 10.2 and Table 10.4 show the analysis in terms of $c(s)$ analysis. In principle, the comparison of the $c(s)$ obtained at different protein and/or detergent concentrations has to be made (see Salvay et al. 2007 for a protocol) to get confidence in the details of the $c(s)$ distribution, and to check the absence of auto-association equilibrium (see Josse et al. 2002 for an example).

Table 10.3 Parameters for SV analysis

<i>Buffer</i> ^a	ρ	g/mL	1.000
	η	mPa s	1.003
<i>FhuA</i> ^b	M^d	Da	79,971
	\bar{v}	mL/g	0.724
	$E_{0,\%1,280}$	$\text{cm}^{-1}(\text{mg/mL})^{-1}$	1.297
	$\partial n/\partial c$	mL/g	0.190
<i>F₆-DigluM</i> ^c	M^d	Da	863.6
	\bar{v}	mL/g	0.585
	$E_{0,\%1,280}$	$\text{cm}^{-1}(\text{mg/mL})^{-1}$	n.d.
	$\partial n/\partial c$	mL/g	0.091
	cmc^d	mM	0.38–0.48
	N_{agg}^d		51–54

^a Values are measured at 20 °C

^b From the amino acid sequence, using the program SEDFIT; we neglect the possible bound polyliposaccharide

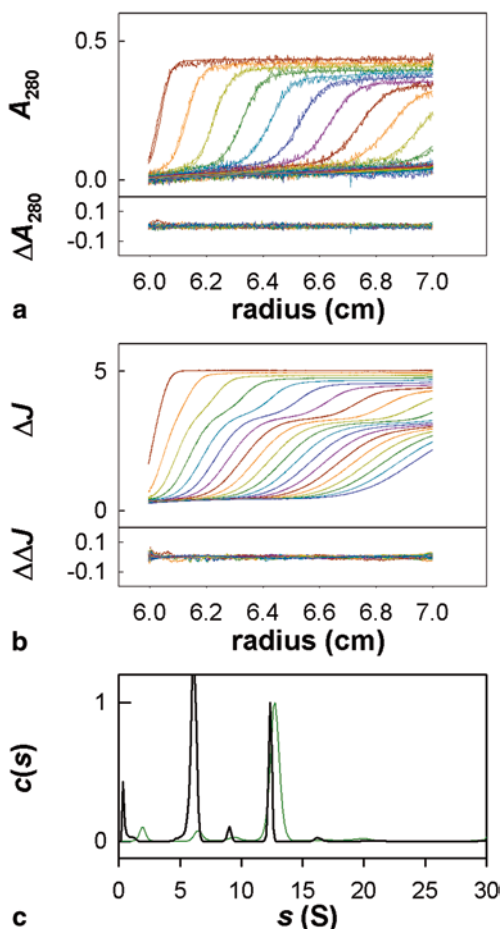
^c From Abia et al. (2012)

^d Optionally required

The data at 280 nm (Fig. 10.2a, c) show one main boundary at 12.5 ± 0.2 S corresponding to the sedimentation of the protein–detergent complex. Homogeneity is the first result of such analysis. The data from interference (Fig. 10.2b, c) show in addition the sedimentation of the surfactant micelles, at an s -value of 6.1 S from interference, close to that 6.4 ± 0.1 S reported in H₂O (Abia et al. 2012). From the fringe shift number of this boundary, we determine a micelle concentration of 2.15 mM, to which has to be added the cmc to obtain experimentally the value for the total free surfactant concentration of about 2.5 mM, which is that expected from our sample preparation protocol. For the protein–detergent complex boundary, the ratio of the signals (absorbance and interference) is used to determine the amount of bound surfactant. This precision of this determination depends on the accuracy of the extinction coefficients for the detergent, which are sometimes ill defined (adsorption of absorbing hydrophobic impurities or bound ligands). We have used here for the surfactant, the experimental extinction coefficient of $0.01 \text{ (mg/mL)}^{-1} \text{ cm}^{-1}$ determined from the integration of the $c(s)$ contribution of the free micelles to derive 143 ± 10 mol/mol bound surfactant. Assuming this value and a monomeric FhuA, we can determine a frictional ratio of ca. 1.4 for the protein–detergent complex, which is slightly larger than 1.25, expected for a globular compact assembly. With only these data, it is hard to certify the anisotropy of the shape of the complex. Furthermore, if we analyze the experimental value of s , considering a globular compact shape ($f/f_{\min} = 1.25$)—which is likely, since FhuA does not bear large extra-membrane domains—we have a second evaluation of the bound surfactant of 117 ± 3 mol/mol, which is not very far from that determined above from the integration of the related peak in the $c(s)$ obtained from absorbance and interference. In the absence of other data, we have to consider the two sets of values as possible.

Because the sedimentation of FhuA observed at 280 nm shows essentially a main species, we attempted an analysis considering the model of one interacting

Fig. 10.2 Sedimentation velocity analysis of FhuA in F_6 -DigluM. Superposition of 20 experimental and fitted sedimentation velocity profiles (*top sub-panels*) obtained during 3.5 h at 42,000 rpm and 20°C, in 12 mm optical path length centerpieces, at 280 nm (a) and with interference optics (b), and their differences (*bottom sub-panels*). Superposition of the sedimentation coefficient distributions $c(s)$ (c) from the absorbance (*green line*) and interference (*black line*) data. ΔJ is in fringe shift displacement unit



species to determine s and M_b (technically, this is done by setting in the analysis program, $\bar{v}=0.75$ mL/g and $\rho=1$ g/mL, which provides a buoyant factor $(1-\bar{v}\rho)=0.25$, thus $M_b=M_{app}/4$, with M_{app} , the output apparent molar mass from the fit). The fit (not shown) is as good as the $c(s)$ analysis (Fig. 10.2a), and leads to the same s -value (12.7 S), but the numerical analysis leads clearly to an improbable solution, i.e., the derived M_b is 49.5 kDa. The contribution from FhuA is 22.1 kDa, the remaining attributed to F_6 -DigluM would be the difference, and would lead to the calculation of 80 mol/mol bound F_6 -DigluM, which is not compatible with the s -value (i.e., it would correspond to a frictional ratio of 1.0, which is too low). This illustrates the fact that the analysis in term of noninteracting species requires strict sample homogeneity.

Table 10.4 Results from the $c(s)$ analysis

Species	Parameter	Value	Line number
F6-DigluM micelle	s (S) from A_{280}	6.5	(1)
	s (S) from ΔJ	6.1	(2)
	A_{280} , 1 cm	0.02	(3)
	ΔJ , 1 cm	2.58	(4)
	c (mg/ml) from line (4)	1.86	(5)
	c (mM) from line (4)	2.15	(6)
	$E_{0.1\%, 280 \text{ nm}}$ ((mg/mL) ⁻¹ cm ⁻¹) from lines (3), (5)	0.01	(7)
FhuA complex ^a	s (S) from A_{280}	12.7	(8)
	s (S) from J	12.4	(9)
	s (S) consensus	12.5±0.2	(10)
	A_{280} , 1 cm	0.30	(11)
	ΔJ , 1 cm	1.11	(12)
	$\Delta J/A_{280}$ from lines (11), (12)	3.7±0.1	(13)
	B_d (g/g) from line (13)	1.55±0.1	(14)
	B_d (mol/mol) from line (14)	143±10	(15)
	ff_{\min} from lines (10), (13)	1.38±0.07	(16)
	B_d (g/g) from line (10) assuming $ff_{\min} = 1.25$	1.26±0.04	(17)
	B_d (mol/mol) from line (17)	117±3	(18)

^a The analysis considers FhuA as a monomer without associated lipids

10.6.2 SEC/MALS: FhuA in LDAO

10.6.2.1 Aim of the Experiment

Our aim was to quantify the amount of bound detergents LDAO to FhuA and test the possibilities of the DLS detector for membrane proteins.

10.6.2.2 Sample, SEC/MALS Experiment, and Parameters for the Analysis

Thirty microliters of FhuA at 3 mg/mL was injected at 0.5 mL/min, with a high-performance liquid chromatography (HPLC) system LC20AD with a thermostated at 15 °C sample changer (Shimadzu France), onto a Superdex 200 10/300 GL column (GE Healthcare) at 20 °C (oven Wynsep, Ste Foy d'Aigrefeuille, France) with a solvent, 20 mM Tris-HCl pH 8, 0.15 M NaCl, containing 0.05 % LDAO. Elution profiles were followed online at 280 nm (SPD-M20A Shimadzu) and RI, static and dynamic light scattering (Optilab rEX, miniDAWN TREOS and Dynapro Nanostar, Wyatt Technology, Santa Barbara, CA, USA) were measured using a laser emitting at 658 nm. Data were analyzed using the protein conjugate template of the ASTRA software V 5.3.4.18. Input values were, for the sample, those of water; for FhuA, $E_{0.1\%, 280}$ and $\partial n/\partial c$ given in Table 10.3; and for LDAO, $E_{0.1\%, 280} = 0.02$ mL/(g/cm) and $\partial n/\partial c = 0.138$ – 0.148 mL/g (Strop and Brunger 2005). We considered as default

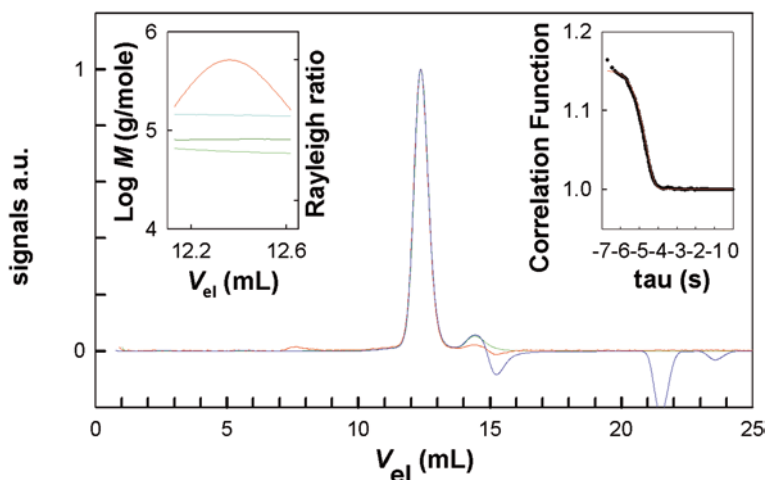


Fig. 10.3 Molar masses determination of FhuA solubilized in LDAO by SEC/MALS. *Main panel:* Superposition of A_{280} (green), differential refractive index (blue), light scattering (red) signals, in arbitrary units, as a function of the elution volume V_{el} . *Left insert:* Molecular masses calculation during elution, of the complex (light blue), with the protein (dark green) and detergent (light green) contributions. *Right insert:* Autocorrelation function of the scattered light at the maximum of the main peak (black symbols) and fit (red line)

value the second virial coefficient of 2×10^{-4} mol mL/g². Other optional information concerning LDAO, not relevant on the present analysis are: $M=229.41$ g/mol, $\bar{v}=1.128-1.134$ mL/g (le Maire et al. 2000), $cmc \sim 0.14$ mM in 0.1 M NaCl and 1–2.2 mM in H₂O (le Maire et al. 2000), and $N_{agg}=69-76$ (le Maire et al. 2000).

10.6.2.3 Result of the Analysis

Figure 10.3 and Table 10.5 show the analysis and results. The chromatograms from the three detectors show essentially the same features. We do not observe here a large peak in the void volume of the column (7.6 mL), which indicates an insignificant amount of aggregates. LS (red line) indeed emphasizes the presence of large aggregates, easily detected even if they are present in minor amounts (see, e.g., Ebel 2011; Dach et al. 2012). FhuA main complex is eluted at 12.4 mL as the dominant species. Detergent micelles at ≈ 15 mL are not detected very well, indicating that the detergent was at the same concentration in the sample and elution buffer. The total volume of the column at ≈ 22.5 mL is indicated by negative (as here) or positive contributions arising from minor differences in buffer composition.

From the analysis, considering the whole peaks for the aggregates and the main FhuA complex, we obtain an estimate of the mass recovered upon elution, and thus the percentage of each type of species. It confirms the very minor proportion of aggregate of this sample (0.003%). The recovered concentration of FhuA (95 μ g) is close to the nominally injected concentration (≈ 90 μ g), and that of associated

Table 10.5 Results for FhuA/LDAO from SEC/MALS

Species	Parameter	Value
Aggregates	V_{el} (mL)	7.6
	mass fraction (%) ^a	0.003
Complex	V_{el} (mL)	12.6
	mass fraction (%) ^a	99.997
	protein mass (μg) ^a	95
	detergent mass (μg) ^a	74
	M_p (kDa) ^b	80.8 (0.4%)
	M_d (kDa) ^b	65.7 (0.9%)
	B_d (g/g) ^b	0.77
	B_d (mol/mol)	268
R_H (nm) ^b	4.4 (4%)	

^a Considering the volume range for the whole peaks (7–9 mL and 11–13.6 mL)

^b Obtained considering a volume range of 0.5 mL centered on the maximum of the peak

detergent (74 μg) is in the same order of magnitude. If there were an excess of detergent in the sample, the analysis of the micelle peak would allow the determination of the eluted amount, which could be transcribed in excess concentration in the sample.

Further analysis is made considering a volume range of 0.5 mL centered on the FhuA complex. The range corresponds nearly to the half height, and determines the errors in the resulting values indicated in percentage in Table 10.5. The molecular masses during elution, shown in the left insert of Fig. 10.3, for the complex, the protein, and detergent contributions, are nearly constant. The mass of protein is close to the theoretical value of 80 kDa, and the amount of bound detergent, 0.77 g/g or 268 mol/mol, is thus determined with confidence. Changing the parameters of the analysis takes a couple of seconds, and we check easily that the incertitude on the value of $E_{0.0\%,1,280}$ or $\partial n/\partial c$ for the detergent does not significantly influence the outputs of the analysis.

The correlation function from DLS at the maximum of the elution peak is shown in the right insert of Fig. 10.3. The scattered light at the maximum of the main peak is dominated by the signal of the complex. Analysis of the eluant, at elution volumes where the chromatogram is flat, is extremely noisy, showing that detergent micelles in the solvent do not contribute significantly to the signal. The derived R_H for FhuA–LDAO complex is 4.4 nm in the range of the values given for FhuA–DDM (4.2–4.5 nm; Boulanger et al. 1996).

10.6.3 Conclusion

AUC is a very resolutive technique, and particularly appropriate for the study of complex systems. The separation of the particle depends on their mass, hydrodynamic radius, and buoyancy. It requires small volumes of sample and solvent.

Samples with different concentrations of protein and detergent, and different solvent densities can be easily investigated. The $c(s)$ analysis is easy, robust, and allows to characterize the number of boundaries, i.e., the minimum number of sedimenting species. Comparison of data obtained with different samples allows to determine if proteins auto-associate. In favorable cases, the related association constants, M_p , B_d , and possibly R_H , may be obtained.

SEC/MALS requires large volume of solvent (and detergent) for equilibrating the column, and protein–detergent complexes may evolve upon chromatography. It emphasizes the presence of large-size particles, and is therefore particularly appropriate for quantifying aggregates. Given protein–detergent complexes are well separated from detergent micelles and aggregates, SEC/MALS is an easy technique providing precise estimates of M_p , B_d (reasonably precise R_H), and free detergent concentration, and appears appropriate for systematic studies.

Acknowledgment This work used the platforms AUC and PAOL of the Grenoble Instruct centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

References

- Abla M, Durand G, Breyton C, Raynal S, Ebel C (2012) A diglycosylated fluorinated surfactant to handle integral membrane proteins in aqueous solution. *J Fluor Chem* 134:63–71
- Balbo A, Minor KH, Velikovskiy CA, Mariuzza RA, Peterson CB, Schuck P (2005) Studying multiprotein complexes by multisignal sedimentation velocity analytical ultracentrifugation. *Proc Natl Acad Sci U S A* 102(1):81–86
- Boulanger P, leMaire M, Bonhivers M, Dubois S, Desmadril M, Letellier L (1996) Purification and structural and functional characterization of FhuA, a transporter of the Escherichia coli outer membrane. *Biochemistry* 35(45):14216–14224
- Breyton C, Gabel F, Abla M, Pierre Y, Lebaupain F, Durand G, Popot J-L, Ebel C, Pucci B (2009) Micellar and biochemical properties of (hemi)fluorinated surfactants are controlled by the size of the polar head. *Biophys J* 97(4):1077–1086
- Dach I, Olesen C, Signor L, Nissen P, le Maire M, Møller JV, Ebel C (2012) Active detergent-solubilized H⁺,K⁺-ATPase Is a Monomer. *J Biol Chem* 287(50):41963–41978
- Ebel C (2011) Sedimentation velocity to characterize surfactants and solubilized membrane proteins. *Methods* 54:56–66
- Ebel C, Møller JV, le Maire M (2007) Analytical ultracentrifugation: membrane protein assemblies in the presence of detergent. In: Pebay-Peyroula E (ed) *Biophysical analysis of membrane proteins. Investigating structure and function*. Wiley, Weinheim, pp 91–120 (p. Chap. 4)
- Echalier A, Pan YB, Birol M, Tavernier N, Pintard L, Hoh F, Ebel C, Galophe N, Claret FX, Dumas C (2013) Insights into the regulation of the human COP9 signalosome catalytic subunit, CSN5/Jab1. *Proc Natl Acad Sci U S A* 110(4):1273–1278
- Flayhan A, Wien F, Paternostre M, Boulanger P, Breyton C (2012) New insights into pb5, the receptor binding protein of bacteriophage T5, and its interaction with its Escherichia coli receptor FhuA. *Biochimie* 94(9):1982–1989
- Hayashi Y, Matsui H, Takagi T (1989) Membrane protein molecular weight determined by low-angle laser light-scattering photometry coupled with high-performance gel chromatography. *Method Enzymol* 172:514–528
- Irimia A, Ebel C, Madern D, Richard SB, Cosenza LW, Zaccari G, Vellieux FM (2003) The Oligomeric states of Haloarcula marismortui malate dehydrogenase are modulated by solvent components as shown by crystallographic and biochemical studies. *J Mol Biol* 326(3):859–873

- Josse D, Ebel C, Stroebel D, Fontaine A, Borges F, Echalié A, Baud D, Renault F, Le Maire M, Chabrieres E, Masson P (2002) Oligomeric states of the detergent-solubilized human serum paraoxonase (PON1). *J Biol Chem* 277(36):33386–33397
- Lebowitz J, Lewis MS, Schuck P (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci* 11(9):2067–2079
- le Maire M, Champeil P, Møller JV (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta* 1508(1-2):86–111
- le Maire M., Arnou B, Olesen C, Georjgin D, Ebel C, Møller JV (2008) Gel chromatography and analytical ultracentrifugation to determine the extent of detergent binding and aggregation, and Stokes radius of membrane proteins using sarcoplasmic reticulum Ca²⁺-ATPase as an example. *Nat Protoc* 3(11):1782–1795
- Le Roy A., Nury H, Wiseman B, Sarwan J, Jault J-M, Ebel C (2013) Sedimentation velocity analytical ultracentrifugation in hydrogenated and deuterated solvents for the characterization of membrane proteins. In: Springer (ed) *Membrane biogenesis: methods and protocols*. Humana Press, New York pp 219-251
- Manon F, Ebel C (2010) Analytical ultracentrifugation, a useful tool to probe intrinsically disordered proteins. In: Uversky VN, Longhi S (eds) *Instrumental analysis of intrinsically disordered proteins: assessing structure and conformation*. Wiley, Hoboken, pp 433–449
- Nury H, Manon F, Arnou B, le Maire M, Pebay-Peyroula E, Ebel C (2008) Mitochondrial bovine ADP/ATP carrier in detergent is predominantly monomeric but also forms multimeric species. *Biochemistry* 47(47):12319–12331
- Salvay AG, Ebel C (2006) Analytical ultracentrifuge for the characterization of detergent in solution. *Progr Colloid Polym Sci* 131:74–82
- Salvay AG, Santamaria M, le Maire M, Ebel C (2007) Analytical ultracentrifugation sedimentation velocity for the characterization of detergent-solubilized membrane proteins Ca⁺⁺-ATPase and ExbB. *J Biol Physics* 33:399–419
- Salvay AG, Gabel F, Pucci B, Santos J, Howard EI, Ebel C (2010) Structure and interactions of fish type III antifreeze protein in solution. *Biophys J* 99(2):609–618
- Salvay AG, Communie G, Ebel C (2012) Sedimentation velocity analytical ultracentrifugation for intrinsically disordered proteins. In: Uversky VN, Dunker AK (eds) *Intrinsically disordered protein analysis*. Springer, New York, pp 91–104
- Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys J* 78(3):1606–1619
- Schuck P (2010a) Diffusion of the reaction boundary of rapidly interacting macromolecules in sedimentation velocity. *Biophys J* 98(11):2741–2751
- Schuck P (2010b) On computational approaches for size-and-shape distributions from sedimentation velocity analytical ultracentrifugation. *Eur Biophys J* 39(8):1261–1275
- Schuck P (2010c) Sedimentation patterns of rapidly reversible protein interactions. *Biophys J* 98(9):2005–2013
- Slotboom DJ, Duurkens RH, Olieman K, Erkens GB (2008) Static light scattering to characterize membrane proteins in detergent solution. *Methods* 46(2):73–82
- Solovyova A, Schuck P, Costenaro L, Ebel C (2001) Non-ideality by sedimentation velocity of halophilic malate dehydrogenase in complex solvents. *Biophys J* 81(4):1868–1880
- Strop P, Brunger AT (2005) Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci* 14(8):2207–2211
- Zhao H, Balbo A, Brown PH, Schuck P (2011) The boundary structure in the analysis of reversibly interacting systems by sedimentation velocity. *Methods* 54(1):16–30

Chapter 11

Lipidic Cubic Phase Technologies for Structural Studies of Membrane Proteins

Andrii Ishchenko, Enrique Abola and Vadim Cherezov

11.1 Introduction

Membrane proteins (MPs) are essential components of cellular membranes. They constitute about 30% of the proteome in most organisms (Fagerberg et al. 2010) and play crucial roles in many cellular and physiological processes. Due to their involvement in a large number of human pathological conditions, MPs serve as important therapeutic targets; more than 60% of drugs currently in the market bind to and modulate MP function (Pieper et al. 2013). MPs provide links between the extracellular and the intracellular environments and include receptors that are central to signal transduction, transporters, ion channels, enzymes, and others, such as adhesion proteins involved in cell–cell recognition. Attempts at developing a full understanding of their mechanisms of action require not only biophysical and biochemical characterization, but more importantly the knowledge of high-resolution structures of MPs and their complexes. X-ray crystallography has become the most successful method for generating atomic resolution structures, but requires the availability of sufficiently large and well-ordered crystals that diffract to high resolution. Crystallization of MPs, however, is challenging, because of their low stability outside their natural membrane environment as well as the dual nature of their surfaces, parts of which are hydrophobic and therefore embedded in the membrane and the rest are solvent exposed and hydrophilic. Thus, despite the high biomedical value of these molecules, only 1% of all Protein Data Bank (PDB) entries currently belong to MPs.

The complexity of MP structural biology studies is exemplified by difficulties that are encountered in attempting to solve the high-resolution structures of human G protein-coupled receptors (GPCRs; Stevens et al. 2013). This protein superfamily is responsible for cellular signaling, and represents the most important class of

V. Cherezov (✉) · A. Ishchenko · E. Abola
Department of Integrated Structural and Computational Biology,
The Scripps Research Institute, La Jolla, California, USA
e-mail: vcherezov@scripps.edu

drug targets (Rask-Andersen et al. 2011; Congreve et al. 2011). The study of GPCR structure and function is of great fundamental and practical importance as it can be applied to design more efficient and safer therapeutics. These receptors represent highly challenging heterologous expression targets and are also very difficult to handle because of their intrinsic conformational flexibility, which is essential for their role in signaling but detrimental for crystallization. In addition, elucidation of their mechanism of action requires not only the availability of receptor structures in different conformational states bound to different ligands, but also structures of their complexes with signaling partners, such as G proteins or β -arrestins. After about 20 years of extensive research, multiple breakthroughs were achieved with the establishment of a robust protein production platform in insect cells (Hanson et al. 2007), receptor stabilization by fusion partners or antibodies (Rosenbaum et al. 2007; Rasmussen et al. 2007; Chun et al. 2012), development of crystallization methods in lipidic environment (Cherezov et al. 2007; Cherezov 2011), and advancements in micro-crystallography (Cherezov et al. 2009; Smith et al. 2012). By 2013, structures of about 20 unique receptors as well as a complex between a GPCR and G protein, most of which were crystallized in the lipidic cubic phase (LCP), became available in the PDB (Katritch et al. 2013).

Although MP crystallization in LCP was introduced more than 17 years ago (Landau and Rosenbusch 1996), the initial progress of structure determination by this method was slow (Fig. 11.1), and its application was limited to studies of colored proteins from the microbial rhodopsin family. The primary reason for a slow adoption rate for this method was due to difficulties in handling the sticky, gel-like LCP material and challenges in the detection of small colorless protein crystals. This method was therefore not fully embraced by the structural biology community until recently when automated instruments and dedicated LCP tools became commercially available. These include tools for mesophase preparation and manipulation, pre-crystallization assays to monitor MP properties and automated systems for setting up crystallization trials and detecting crystals. Additionally, improvements in hardware and crystallographic data collection methodology at microfocus beamlines allowed structure determination from relatively small crystals grown in LCP. These recent advances have now made the use of LCP for crystallization generally accessible to laboratories of all sizes that wish to pursue the high-resolution structures of MPs. In this chapter, we discuss the main principles of this method and provide highlights of recent advancements in LCP-related technologies.

11.2 Crystallization of MPs

In general, MP crystallization methods can be divided into two categories depending on the protein-hosting environment: *in surf*o and *in meso* (Fig. 11.2). The *in surf*o method has been widely used since the first MP structure of the photosynthetic

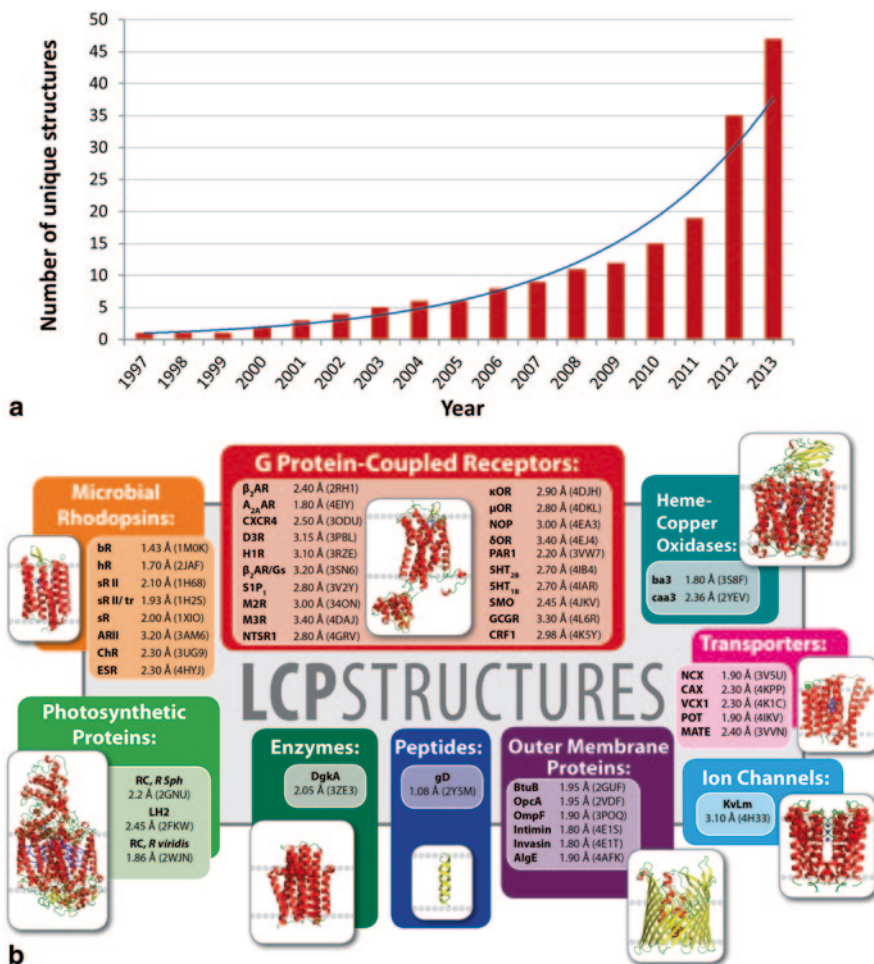


Fig. 11.1 Progress in membrane protein (MP) structure determination by the lipidic cubic phase (LCP) method. **a** Cumulative increase in the number of unique MP structures solved by using LCP crystallization and deposited to the Protein Data Bank (PDB). The data do not follow a simple exponential growth shown as the *blue curve*, but rather indicate an initial lag phase associated with developments of LCP technology leading to a sharp increase in the structure determination rate started a few years ago when many of these developments became commercially available. **b** Collection of all unique MPs and peptides crystallized by the LCP method, showing the achieved resolution and the corresponding PDB ID

reaction center was solved in 1985 (Deisenhofer et al. 1985). This technique relies on solubilization of MPs in detergent micelles. The use of surfactants helps to extract proteins from the membrane and to maintain the insoluble MP in solution. The protein detergent complex (PDC) can then be treated as a soluble protein and crystallized using conventional crystallization methods including vapor diffu-

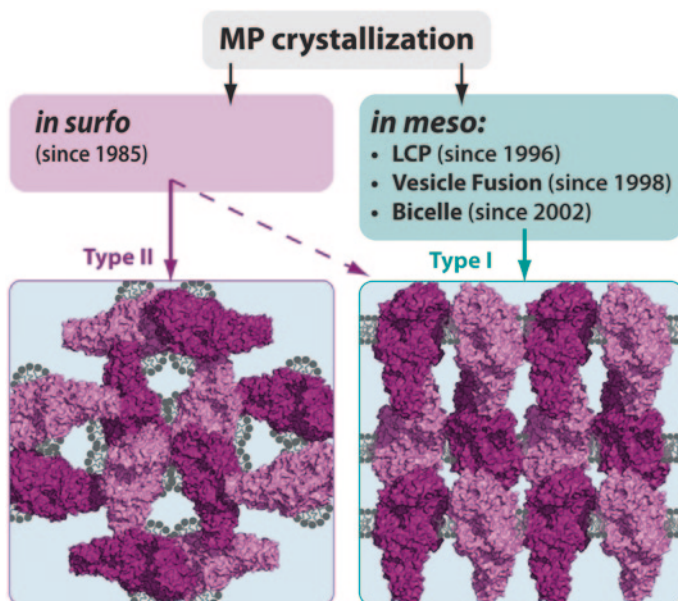


Fig. 11.2 Two approaches to MP crystallization. *In surfo* methods typically produce type II crystals, and occasionally lead to type I crystals. *In meso* crystallization always results in type I crystal packing. Crystal packing diagrams are shown for the photosynthetic reaction center from *B. viridis* crystallized in detergent solutions (type II packing, PDB ID 1PRC) and in LCP (type I packing, PDB ID 2WJN)

sion, batch, free-interface diffusion, etc. (McPherson 2004). Although the *in surfo* method remains the most frequently used approach for the majority of published MP structures, the environment of detergent micelles is not ideal for crystallization. Many detergents are too harsh and can destabilize MPs, while milder detergents tend to form larger micelles that shield a substantial part of an MP and thus limit possibilities for protein–protein contacts that contribute to the generation of higher-quality crystals. The choice of an optimal detergent for solubilization and subsequent crystallization is often not a trivial task and requires elaborate and expensive screening tests (Wiener 2004). Finally, *in surfo* crystallization typically results in type II crystals (Michel 1991; see Fig. 11.2), where crystal contacts are formed only by polar parts of the protein leading to a large solvent content due to the porous nature of the packing and resulting in crystals with low order and poor diffraction quality.

An alternative method for MP crystallization (referred to here as the *in meso* approach) is based on employment of a quasi-native lipid environment, where MPs are reconstituted in a lipid bilayer and crystallization is initiated from a lipidic mesophase. Three technologies have been developed that fall into this category: LCP (Caffrey and Cherezov 2009), vesicle fusion (Takeda et al. 1998), and bicelles (Faham and Bowie 2002). LCP crystallization was first introduced in 1996 (Laudau and Rosenbusch 1996), resulting in the first successful determination of the

high-resolution structure of bacteriorhodopsin (Pebay-Peyroula et al. 1997), a protein that was exhaustively studied, but which resisted the formation of well-ordered crystals using conventional crystallization methods in detergent micelles. LCP crystallization, for the first time, provided a universal approach to produce MP crystals with type I packing. The vesicle fusion method was published in 1998 (Takeda et al. 1998). Following the reconstitution of an MP into spherical lipid bilayers (vesicles), precipitants are added which trigger fusion of vesicles resulting in a subsequent nucleation process. Despite its earlier success in bacteriorhodopsin crystallization, no other MPs have been crystallized with this procedure. Finally, the last *in meso* method, based on MP reconstitution into discoidal lipid/detergent micelles called bicelles, appeared in 2002 (Faham and Bowie 2002). The flat surface of the bicelle is formed by a lipid bilayer and the hydrophobic rim is stabilized by detergents. The size and thickness of bicelles depend on the type of the lipid, detergent, and the molar ratio between them (Whiles et al. 2002). Bicelle solutions are typically liquid and easy to manipulate using a pipette at low temperatures and become more viscous and gel-like as temperature increases due to bicelle fusion and formation of a perforated lamellar phase (Wang et al. 2003), which is conducive to MP crystallization. Several proteins have been crystallized in bicelles, including one GPCR (Rasmussen et al. 2007).

As mentioned earlier, crystallization in LCP is the most successful of the *in meso* methods to date. By August 2013, representatives of at least nine different families of MPs have been crystallized in LCP, resulting in 151 entries in the PDB, among them 47 unique structures (<http://cherezov.scripps.edu/structures.htm>; Fig. 11.1). Examples include alpha-helical and beta-barrel proteins, peptides, photosynthetic proteins, receptors, enzymes, transporters, and ion channels. The highest resolution of an MP crystallized from LCP was achieved for bacteriorhodopsin (1.43 Å, PDB IDs 1M0K, 1M0M). A hydrophobic peptide, gramicidin D, was also crystallized in LCP and its structure was solved at 1.08 Å (PDB ID 2Y5M).

There are several factors that contribute to the success of using LCP for MP crystallization. Just as bicelles and lipid vesicles, LCP provides a native-like lipid environment that stabilizes sensitive MPs. Crystals grown in LCP belong to type I, where crystal contacts are made not only by the polar parts, but also by the non-polar parts of MPs (Fig. 11.2). Such crystals have tight packing and lower water content often resulting in small but well-ordered crystals. LCP crystallization depends strongly on protein diffusion within the LCP matrix (Cherezov et al. 2008). Large-sized impurities and protein aggregates cannot diffuse through LCP and are therefore excluded from the crystal formation process (Kors et al. 2009). This advantage together with a large curvature of LCP membranes can turn into a disadvantage when a large MP or a complex of MPs has to be crystallized. The protein-size limitation, however, can be overcome by swelling LCP and transforming it into a liquid-like sponge phase (Cherezov et al. 2006a; Wadsten et al. 2006).

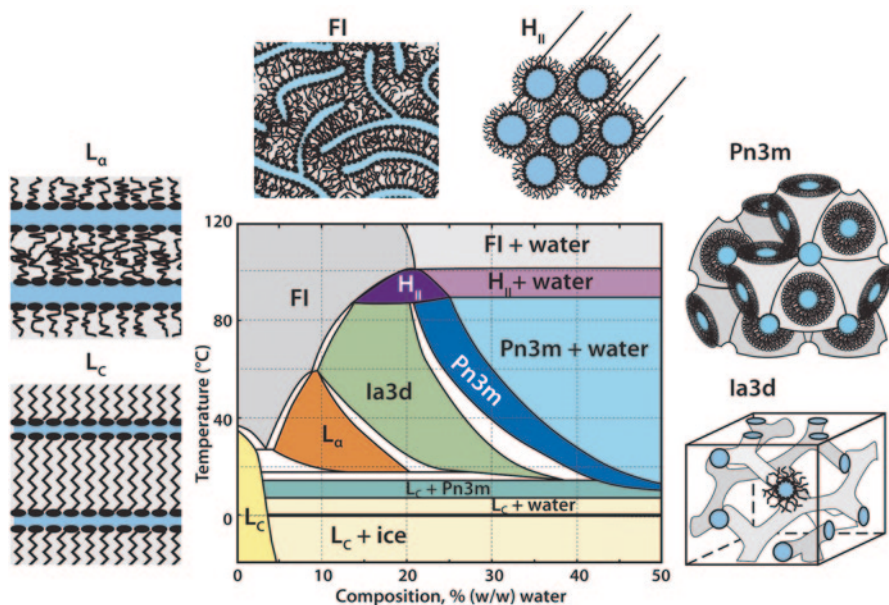


Fig. 11.3 Equilibrium temperature–composition phase diagram for monoolein (redrawn from Qiu and Caffrey 2000, with permission). Cartoon representations of different phases are shown around the phase diagram

11.3 LCP Structure and Properties

Lipids are amphiphilic molecules, which, upon mixing with water, form a variety of liquid crystalline phases depending on their chemical structure, temperature, and hydration level (Qiu and Caffrey 2000; Briggs et al. 1996). The most common class of lipids used for LCP crystallization is monoacylglycerols (MAGs). They contain a hydrophilic glycerol headgroup attached to a hydrophobic monounsaturated fatty acid chain through an ester bond. Here we refer to monounsaturated 1-MAG lipids using the *N.T* MAG abbreviation scheme, where “*N*” represents the number of hydrocarbons between the ester bond and the double bond and “*T*” represents the number of hydrocarbons between the double bond and the end of the acyl chain. The temperature–composition phase diagram of the lipid most successfully used for LCP crystallization, monoolein or 9.9 MAG, is shown in Fig. 11.3. At low hydration and temperatures below 37°C, 9.9 MAG exists in a solid lamellar crystalline, L_C phase. Above 37°C it melts, forming a fluid isotropic, FI phase. The addition of water leads to the formation of several lyotropic mesophases, the identity of which depends on the temperature and water content. The middle part of the phase diagram at moderate hydration (5–20% water) and temperatures (18–60°C) is occupied by a lamellar liquid crystalline, L_α phase. High hydration levels are dominated by cubic phases in a wide range of temperatures from 18 to 90°C. Above 90°C, lipidic phase converts into an inverted hexagonal, H_{II} phase.

The part of the phase diagram occupied by LCP is the most interesting for crystallization. LCP is composed of a single lipid bilayer that separates the space into two interpenetrating nonintersecting networks of water channels. Both lipid bilayer and water channels are continuous in all three dimensions and, therefore, this type of LCP is often referred to as a bicontinuous phase. This continuity allows MPs to diffuse within the single lipid bilayer and precipitants to penetrate inside the LCP and induce crystallization. The lipid bilayer of LCP follows an infinite periodic minimal surface (IPMS) with a cubic symmetry (Hyde et al. 1984). Each point on an IPMS corresponds to a saddle point with the two principal orthogonal curvatures of the same absolute values but opposite directions resulting in the mean curvature of zero (Rummel et al. 1998). LCP lattices with three different space groups have been observed: Ia3d (gyroid), Pn3m (diamond), and Im3m (primitive). The typical lattice parameter of a Pn3m-LCP is 110 Å with a water channel diameter of about 50 Å.

Macroscopically, LCP appears as a sticky, viscous, transparent, and optically isotropic gel. This material does not flow on its own or under gravity and cannot be pipetted, but it can be extruded through a needle using a positive displacement syringe. Due to its unique properties, LCP has been used in many applications, including drug delivery, removal of contaminations, biosensors, and protein crystallization (Shah et al. 2001; Clogston 2005; Nazaruk et al. 2008; Cherezov 2011).

11.4 LCP Crystallization

The underlying procedure for LCP crystallization is relatively simple and straightforward:

1. Solubilized protein is combined with a host lipid at a certain proportion dictated by the phase diagram and mixed together until transparent LCP is formed.
2. A precipitant is added triggering crystal nucleation and growth.

Despite the apparent simplicity, there are several important factors to consider, some of which have been addressed by a number of studies on the mechanism of MP crystallization in LCP (Nollert et al. 2001; Caffrey 2008). Upon mixing of lipid with detergent-solubilized protein solution, LCP is formed spontaneously and protein is reconstituted in the lipid bilayer of LCP, as demonstrated by fluorescence quenching (Liu and Caffrey 2005; Cherezov et al. 2006a). Protein insertion in LCP is usually performed at room temperature, and along this process the protein experiences a variety of changing environments; therefore, fast incorporation of protein into LCP, as that achieved by mechanical mixing (Caffrey and Cherezov 2009), is preferred. MPs are generally more stable in LCP than in detergent solutions (Liu et al. 2010); however, it is recommended to set up crystallization trials immediately upon protein reconstitution. Crystals that grow in LCP are formed by layers of two-dimensional MP crystals stacked in the third dimension through hydrophilic interactions of solvent exposed MP parts; this arrangement is known as type I packing (Fig. 11.2). Growing crystals are attached to the bulk LCP through a multi-

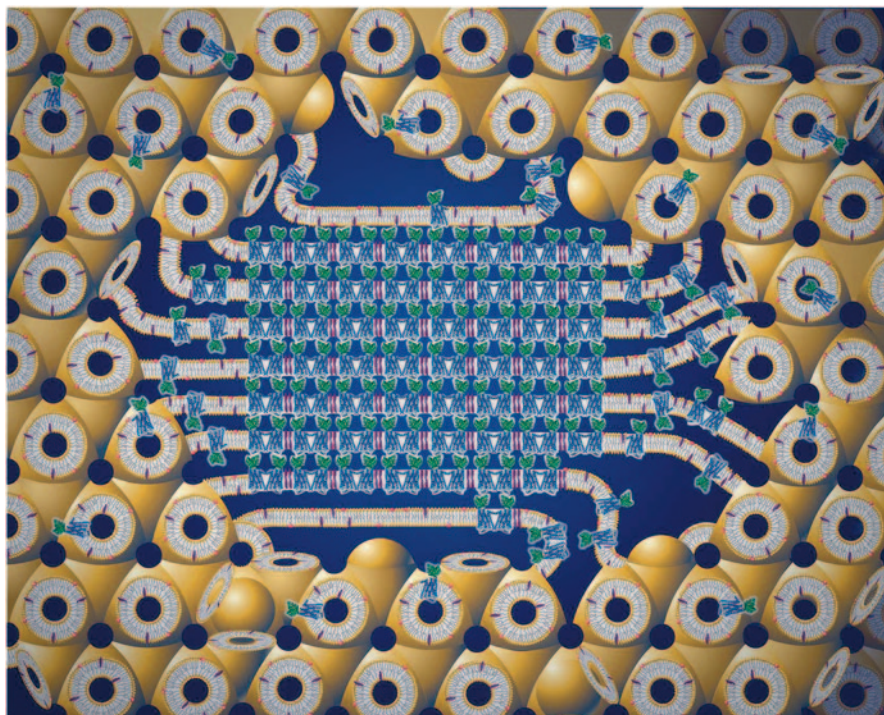


Fig. 11.4 Cartoon representation of a hypothetical mechanism of MP crystal growth in LCP. MPs (blue/green) are reconstituted in the lipid bilayer of LCP (yellow), which provides a stabilizing environment and allows MPs to diffuse in the three-dimensional space. Addition of a precipitant induces favorable protein interactions leading to formation of a crystal nucleus. The growing crystal consists of layers of two-dimensional crystals stacked in the third dimension (type I packing). The crystal is attached to the bulk LCP through a multilamellar phase portal, allowing MPs to approach and join the growing crystal. (Reproduced from Caffrey and Cherezov 2009)

lamellar lipid phase (Cherezov and Caffrey 2007), which serves as a conduit for MPs, allowing them to reach and incorporate into the growing crystal (Fig. 11.4). Sudden temperature fluctuations should be avoided as they induce changes in the LCP lattice parameter and can lead to a detachment of growing crystals from the bulk LCP phase and termination of crystal growth. Therefore, crystallization trials should be incubated and imaged at a constant temperature as much as possible. Crystals typically nucleate within hours to a few days and grow for several days to several weeks. Diffusion of precipitants in LCP is relatively slow, depending strongly on the size of the molecule with respect to the diameter of water channels, for example, polymers diffuse much slower than ions. Overlaying an LCP bolus with a precipitant solution leads to the formation of various concentration and hydration gradients across LCP changing in time and lasting from minutes to days. Therefore, the size of an LCP bolus and its shape can play a profound role in crystal nucleation and growth, and these effects can be explored during crystal optimization. Several different crystal space groups have been observed, but no

apparent correlation between the LCP lattice parameter and crystal lattice has been found. The mismatch is apparently accommodated by adjustments in the stacking distance within the lamellar phase.

Addition of precipitants can affect the lattice parameter of LCP and the phase identity (Cherezov et al. 2001). In general, kosmotropic agents, such as most salts, decrease the hydration capacity of lipid headgroups increasing the lipid's intrinsic curvature, which results in shrinkage of LCP and eventually its transition into a hexagonal H_{II} phase. Large molecular weight polymers, such as PEG 10,000, also shrink LCP but via a different mechanism of applying an osmotic pressure and withdrawing water resulting in transformation to a lamellar L_{α} phase. On the other hand, most chaotropic agents, as well as low molecular weight polymers (PEG 400), alcohols (MPD, propanediol, butanediol), and polymers (jeffamine, pentaerythritol propoxylate), swell LCP and at higher concentrations transform it into a sponge phase (Cherezov et al. 2006a; Wadsten et al. 2006). Combinations of two or more precipitants with opposite effects are not in general additive, but often show a more complex behavior. Small-angle X-ray scattering (SAXS) is the most straightforward method to obtain lipidic phase identity and structural parameters. Several high-throughput (HT) approaches have recently been introduced to perform SAXS analysis of lipidic samples under a variety of conditions in a 96-well plate format (Joseph et al. 2011; Conn et al. 2012). These HT LCP-SAXS approaches can help to better understand the mechanism of LCP crystallization and design special crystallization screens.

Typically, MPs are solubilized and purified in detergent solutions prior to their reconstitution in LCP. Detergents become incorporated in LCP along with the protein. Most detergents are well tolerated by LCP up to concentrations of several times their CMC (critical micelle concentration; Ai and Caffrey 2000; Sennoga et al. 2003; Misquitta and Caffrey 2003). In most cases the detergent identity is not critical for crystallization, therefore it is recommended to use the most stabilizing detergent for MP purification. At high concentrations, detergents can transform LCP into a lamellar L_{α} phase. The effect of detergents on LCP, however, can be counteracted to a degree by increasing the salt concentration (Misquitta and Caffrey 2003), if such an increase in salt concentration does not destabilize the protein.

11.4.1 LCP Host Lipids

The pool of lipids that form stable LCP at or below room temperature and, therefore, are suitable for serving as a host lipid for protein crystallization is relatively limited. Monoolein (9.9 MAG) is the most common and the most successful LCP host lipid to date (Kulkarni et al. 2011). Several other MAGs can also form LCPs, structural properties and phase behavior of which strongly depend on the acyl chain length and the position of *cis*-olefinic bond (Caffrey 2009). Protein stability, diffusion, and interactions are affected by the thickness, curvature, and lateral pressure profile across the lipid bilayer of LCP. Therefore, it is important to search for a host

lipid that would be optimal for crystallization of a particular protein. For example, specific MAGs have been designed for low-temperature crystallization, such as 7.9 MAG (Misquitta et al. 2004a), and for crystallization of large proteins and complexes, such as 7.7 MAG (Misquitta et al. 2004b; Rasmussen et al. 2011). Lipid host screening was successfully applied to improve crystal growth and diffraction for an outer MP OprB (Li et al. 2011) and a membrane enzyme diacylglycerol kinase (DgkA; Li et al. 2013b). Several MAGs are commercially available from Nu Chek Prep, Sigma, and Avanti Polar Lipids; others can be synthesized following published protocols (Coleman et al. 2004; Fu et al. 2010). While largely successful, one of the drawbacks of MAGs is their relatively low stability and susceptibility to hydrolysis at basic or acidic conditions (Murgia et al. 2002).

Another class of LCP-forming lipids based on isoprenoid chains has recently been explored for their use in crystallization. It includes a common cosmetic and food additive lipid phytantriol (Barauskas and Landh 2003) and a library of synthetic lipids with variable hydrocarbon chain length and different polar headgroups (Yamashita et al. 2008; Hato et al. 2009). The distinguishing feature of these lipids is their regularly branched hydrocarbon chain structure, as opposed to a linear hydrocarbon chain with a different degree of unsaturation in MAGs. Isoprenoid-chained lipids have a number of unique and attractive properties, such as low transition temperatures between solid and liquid crystalline phases, lower than in MAGs solute permeability, and higher salt tolerance. Many of these isoprenoid-chained lipids lack an ester linkage between the polar headgroup and the hydrophobic tail making them highly stable with respect to hydrolysis. Phytantriol and β -XyIOC(16+4) lipid were used to successfully crystallize bacteriorhodopsin (Borshchevskiy et al. 2010).

LCP host lipids used for MP crystallization are not common native lipids of biological membranes. It is, however, possible to dope the host lipid, i.e., monoolein, with different natural membrane lipids (Cherezov et al. 2002), which may preferentially interact with MPs incorporated in LCP. For example, addition of cholesterol to monoolein-based LCP improves stability of GPCRs (Liu et al. 2010) and is instrumental for their crystallization (Cherezov et al. 2007; Hanson et al. 2008).

11.5 LCP Assays

The need to screen for LCP host lipids and lipid additives substantially increases the number of crystallization variables. Several LCP pre-crystallization assays have been developed in order to evaluate MP behavior in LCP and approach crystallization in a rational manner. These assays profile activity, stability, and diffusion of MPs in LCP at a variety of conditions, and help to identify the best protein constructs, ligands, LCP host lipids, and substantially reduce the range of precipitant parameters for screening in subsequent crystallization trials.

11.5.1 MP Activity in LCP

Optical transparency of LCP along with a high density of lipid bilayers makes it a convenient system to assay MP activity by a variety of spectroscopic approaches. Such assays also take advantage of the bicontinuous nature of LCP, where both cytoplasmic and extracellular sides of reconstituted MPs can be easily accessed by water-soluble ligands through the network of water channels that span through the mesophase.

One example of a successful application of functional assays in LCP involves the enzyme DgkA, which catalyzes the ATP-dependent phosphorylation of diacylglycerols, converting them into phosphatic acid (Li and Caffrey 2011). Since DgkA can also phosphorylate different MAGs including monoolein, the lipid in this assay performs both roles as the host LCP lipid as well as the substrate. The change in adenosine diphosphate (ADP) concentration was monitored in an indirect way, in which a coupled enzyme system that included pyruvate kinase, lactate dehydrogenase, phosphoenolpyruvate, and nicotinamide adenine dinucleotide (NADH) was used. In this system, the production of ADP leads to a decrease in NADH concentration that can be measured by absorption at 340 nm. The assays have shown that the substrates follow classic Michaelis–Menten saturation behavior and, thus, convincingly demonstrated that DgkA is enzymatically active in LCP. DgkA was successfully crystallized in LCP and its structure was solved at 2.05 Å (Li et al. 2013a).

Another example of activity assays in LCP includes measuring ligand-binding affinity of the cobalamin transporter BtuB (Cherezov et al. 2006b). The binding of the ligand cyanocobalamin (CNCbl) to BtuB reconstituted in LCP was quantified in two ways. First, it was monitored as a decrease in absorption at 361 nm, the characteristic absorption maximum of CNCbl. Second, it was quantified by monitoring ligand-induced intrinsic protein fluorescence quenching. Scatchard analysis was performed to obtain the binding constant, the value of which was similar to that obtained by traditional methods for BtuB in membranes. Ligand bindings have also been measured in LCP using radioligands (Darmanin et al. 2012).

Finally, light-dependent activation of transducin by a prototypical GPCR, visual rhodopsin, incorporated in LCP was studied by UV-visible absorption and Fourier transform infrared spectroscopy (FTIR; Navarro et al. 2002). Transducin was shown to diffuse freely through the LCP matrix within aqueous channels and its ability to form functional complexes with rhodopsin was confirmed. Although the photocycle kinetics was altered, the LCP matrix appeared as a suitable environment for maintaining the GPCR in a functional form. Recently, a structure of β_2 -adrenergic receptor-G protein complex was determined using crystallization in LCP (Rasmussen et al. 2011).

11.5.2 *MP Stability in LCP (LCP-Tm)*

In the case of unstable proteins, such as many MPs, the protein unfolding temperature (T_m) often correlates with crystallization propensity. To evaluate MP stability directly in LCP, an accurate and robust thermal stability assay, LCP-Tm, has been developed (Liu et al. 2010). The assay uses spectroscopic measurements facilitated by the transparency of LCP when formed under proper conditions. Protein unfolding is monitored by either intrinsic protein fluorescence or by using a thiol-reactive fluorescent probe, 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM). The assay protocol addresses issues related to the clouding of LCP with increasing temperature, in which shrinkage of the cubic phase and the shedding of water is observed. Samples are brought back into a transparent state by cooling to 20 °C and centrifugation at 5,600 g for 10–15 min, after which spectroscopic measurements are performed. Therefore, a protein unfolding curve is recorded by employing cycles of 5-min incubation at desired temperature followed by cooling to 20 °C, centrifugation, and spectroscopic readings, which are repeated with increments of 5 °C until the maximum temperature is reached, such as 80 °C in the case of monoolein-based LCP.

LCP-Tm assay has been used to compare effects of ligands, LCP host lipids, and lipid additives on the stability of β_2 -adrenergic receptor in LCP (Liu et al. 2010). The most stabilizing lipid, monoolein, and lipid additive, cholesterol, as identified by LCP-Tm, formed the most successful lipid host mixture for crystallization of GPCRs in LCP. The stabilizing rank of ligands correlated well with the resolution of obtained receptor–ligand co-crystal structures (Cherezov et al. 2007; Hanson et al. 2008; Wacker et al. 2010).

11.5.3 *MP Diffusion in LCP (LCP-FRAP)*

One of the prerequisites for successful crystallization in LCP is the ability of MPs to diffuse within the folded lipid bilayer. Spatial constraints in LCP, such as high local membrane curvature and narrow parts of the water channels, limit diffusion of large MPs or oligomeric aggregates of smaller MPs. The mobility of MPs in LCP is affected by a number of factors: identity of the host lipid and LCP composition, size and aggregation state of the protein, precipitant composition, and temperature. Fluorescence recovery after photobleaching (FRAP) is a convenient technique for measuring protein diffusion. It has been adapted to assay and study MP diffusion in LCP under different conditions (Cherezov et al. 2008). The LCP-FRAP assay allows fast and reliable quantification of protein diffusion characteristics in LCP with a minimal amount of protein.

In this assay, proteins are labeled with hydrophilic 5,5'-disulfato-1'-ethyl-3,3,3',3'-tetramethylindocarbocyanine (Cy3) fluorescent dye. There are two options for Cy3 conjugation carriers: Cy3 mono-maleimide reacting with free

sulfhydryl groups of cysteine residues and N-hydroxysuccinimidyl (NHS) ester reacting with free amino groups. The first labeling method requires free cysteine residues exposed at the protein surface. The amino-reactive conjugation suffers from inadvertent labeling of free-amine-containing lipids like phosphatidylethanolamine (PE) lipids, which are produced by expression hosts and co-purified with the protein. Labeled protein is incorporated in LCP, dispensed in a 96-well glass sandwich plate, overlaid with different screening solutions and incubated at 20°C for 12 h. Measurements are taken by bleaching a spot with a laser and following the recovery of the intensity inside the bleached spot in time. Due to a relatively slow diffusion of MP in LCP, recording a full FRAP curve takes 20–30 min. The curve is then fitted by a diffusion equation giving values for the protein diffusion coefficient and the protein mobile fraction (Cherezov et al. 2008). In order to increase throughput and enable screening for protein diffusion at many different conditions, the assay was modified and automated (Xu et al. 2011). The optimized HT LCP-FRAP protocol includes recording pre-bleached images, sequential bleaching of 96 samples, 30-min incubation, and imaging of the recovered states, allowing measurement of the mobile protein fractions in 96 conditions within 2 h.

The LCP-FRAP method was validated with β_2 -adrenergic and adenosine A_{2A} receptors, for which a good correlation between a high mobile protein fraction and crystallization conditions have been observed (Cherezov et al. 2008; Xu et al. 2011). The HT LCP-FRAP assay was integrated in the GPCR structure determination pipeline by the GPCR Network (Stevens et al. 2013), and has proven to be essential for screening multiple receptor constructs and ligand combinations and identifying initial precipitant conditions for crystallization trials. An automated commercial LCP-FRAP imager capable of performing HT LCP-FRAP assays and taking full LCP-FRAP recovery curves is available from Formulatrix.

11.6 LCP Tools

Original LCP crystallization protocol (Landau and Rosenbusch 1996) was a time- and protein-consuming process, allowing only a few crystallization trials to be set up at a time. LCP was prepared by extensive centrifugation inside small glass tubes, the curvature and thickness of which prevented reliable detection of small colorless protein crystals. These and other issues have now been addressed by a variety of instruments and tools developed for manual and automatic crystallization setups, crystal detection and harvesting, and crystallographic data collection (Fig. 11.5). Handling LCP becomes routine when appropriate instruments and techniques are used. Comprehensive written and video protocols demonstrating optimized procedures employing some of these tools have been published (Caffrey and Cherezov 2009; Cherezov et al. 2010; Liu and Cherezov 2011; Caffrey and Porter 2010; Li et al. 2012).

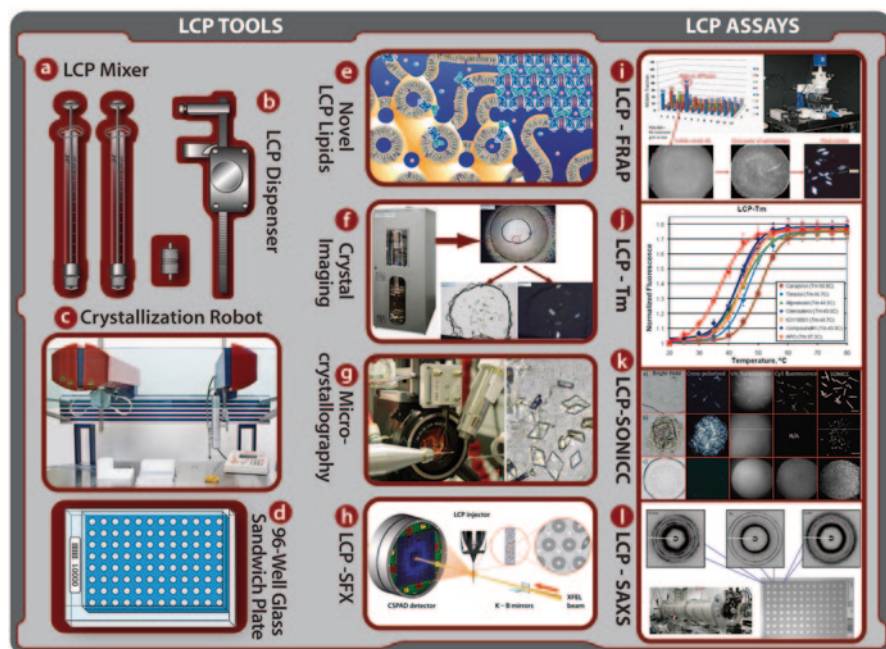


Fig. 11.5 LCP toolchest: collection of tools, instruments, and assays for LCP preparation and handling and MP characterization and structural studies. **a** LCP syringe mixer (Cheng et al. 1998). **b** Manual LCP dispenser (Cherezov and Caffrey 2005). **c** LCP crystallization robot (Cherezov et al. 2004). **d** Glass sandwich plates (Cherezov and Caffrey 2003; Cherezov et al. 2004). **e** Novel LCP lipids (Misquitta et al. 2004a, b; Yamashita et al. 2008; Hato et al. 2009). **f** Crystal imaging in LCP (Cherezov et al. 2010). **g** Micro-crystallography of LCP grown crystals (Cherezov et al. 2009). **h** Serial femtosecond crystallography in LCP (LCP-SFX; Liu et al. 2013; Weierstall et al. 2014). **i** MP diffusion in LCP, lipidic cubic phase fluorescence recovery after photobleaching (LCP-FRAP; Cherezov et al. 2008; Xu et al. 2011). **j** MP stability in LCP (LCP-Tm; Liu et al. 2010). **k** Advanced imaging of crystals in LCP, second order nonlinear optical imaging of chiral crystals (SONICC; Kissick et al. 2010). **l** High-throughput characterization of lipidic mesophases, lipidic cubic phase small-angle X-ray scattering (LCP-SAXS; Joseph et al. 2011)

11.6.1 Syringe Mixer

The most efficient and fast mixing of a lipid with protein solution can be achieved using a simple mechanical mixer, which was initially designed for studying lipid phase diagrams (Cheng et al. 1998). It consists of two coupled Hamilton gas-tight syringes. Detergent-solubilized and concentrated MP solution is loaded into one of the syringes and a molten lipid into another, and they are mechanically mixed by repeated transfer from one syringe to another through a narrow needle within the coupler. LCP is typically formed in about 100 passages taking less than 5 min, and protein volumes as low as 5–10 μL and up to 30–40 μL can be comfortably used with 100- μL syringes. Once a transparent LCP is formed, it is transferred into one

of the syringes, from which it can be delivered into crystallization wells either by using a manual syringe dispenser or an automatic LCP crystallization robot.

11.6.2 Syringe Dispenser

Purified protein solution is one of the most expensive ingredients of crystallization trials and therefore its consumption should be minimized as much as possible. Reproducible manual delivery of protein-laden LCP can be achieved by using a Hamilton repetitive syringe dispenser. When coupled to a 10- μ L Hamilton syringe, it can deliver 200 nL of LCP in a crystallization well upon pressing the button. The Hamilton dispenser was further modified to decrease the dispensing volume down to about 70 nL (Cherezov and Caffrey 2005), saving precious protein and lipids during manual LCP crystallization setup.

11.6.3 Glass Sandwich Plate

LCP crystallization trials can be set up in a number of different ways, including batch or vapor diffusion techniques, using many commercially available crystallization plates. Typically, a small volume of protein-laden LCP is dispensed in each well, overlaid by a larger volume of precipitant solution to reduce dehydration problems (vapor diffusion plates can have the same or different precipitant solution delivered in the reservoir well). The plate is then sealed and incubated at a constant temperature. The interface between LCP and precipitant solution, however, becomes ragged with time making it difficult to detect small colorless protein crystals growing in LCP in these plates. To overcome this problem, a 96-well glass sandwich plate was developed (Cherezov and Caffrey 2003; Cherezov et al. 2004). The plate is made from two flat hydrophobically coated glass slides. The 96 wells of 5-mm diameter are defined by a perforated double sticky tape with the thickness of about 150 μ m attached to the base glass slide. After samples are delivered, the wells are sealed by a thinner glass cover slide. LCP therefore becomes sandwiched between two flat glass slides providing excellent optical properties for crystal detection by different methods, including UV fluorescence. The plate has an SBS-compatible footprint and is suitable for robotic crystallization and automated imaging. Glass sandwich plates are commercially available from Marienfeld-Superior, Hampton Research, and Molecular Dimensions.

11.6.4 LCP Crystallization Robot

Most of the tools discussed so far were related to manual LCP crystallization, which could be implemented on a relatively low budget in a single principal investigator

(PI) laboratory. There are, however, many advantages in the automation of crystallization setups, including a substantial increase in throughput, allowing for better exploration of a vast crystallization space, and the ability to work in harsh conditions, such as low temperatures and in the darkness. Robots can also improve reproducibility and reduce sample consumption.

The first robotic system for LCP crystallization was built using a general-purpose liquid-dispensing station Xantus from Sias (Cherezov et al. 2004). The robot contained two arms, one of which was used for handling liquid precipitant solutions. Arm 2 was adapted to carry an automatic syringe microdispenser for LCP delivery. The robot was capable of delivering as low as 20 nL of LCP boluses into a 96-well glass sandwich crystallization plate and overlaying them with 0.8 μ L of precipitant solutions taken from a 96-well block. The entire plate could be set up within 5–10 min. Droplet evaporation during the crystallization setup was reduced by a homemade attachment delivering the mist produced by a humidifier into a general vicinity of the plate.

A similar principle of LCP delivery was later implemented in several commercial crystallization robots: Mosquito LCP (TTP LabTech), Griffon (Art Robbins), and NT8-LCP (Formulatrix). Modern instruments have high precision, accuracy, and reliability. They are versatile and capable of performing different crystallization setups, including vapor diffusion and bicelle crystallization. Some of them contain an isolated chamber with controlled humidity to minimize evaporation.

11.6.5 *Microfluidics*

Recent improvements in microfabrication processes have triggered rapid progress in microfluidics applications, including those related to macromolecular crystallography. Two novel microfluidics approaches for setting up LCP crystallization trials have recently been described. In one of them, an integrated microfluidic chip capable of mixing lipids with a protein solution inside individual microchambers was designed (Perry et al. 2009). The chip is made of polydimethylsiloxane (PDMS) and contains several microchannels with pneumatic valves, a special mixing chamber, a precipitant chamber, and a crystallization chamber. Protein solution is mixed with lipid by employing a chaotic, tendril-whorl type mixing producing nonbirefringent LCP within 1–2 min, which is then transferred into the crystallization chamber where a precipitant is added. The method was validated by successful crystallization of bacteriorhodopsin. The chip takes less than 20 nL of material per crystallization trial and has the potential for scaling down to sub-nanoliter volumes. Since LCP mixing is achieved in individual chambers, this approach could be useful for implementing host lipid screening, where MP solution is mixed with different lipids. However, the complexity of the valves and pneumatic lines, as well as potential evaporation and loss of chemicals through PDMS (Toepke and Beebe 2006), may prevent a wide acceptance of this method.

Another approach for setting up LCP crystallization trials is based on a plug-based microfluidic system (Li et al. 2010). The system contains two microfluidic devices: One is a flow-focusing device to generate nanoliter volume LCP plugs, and the other is a merging device that merges LCP plugs with precipitant plugs. Combined plugs of LCP and precipitant are separated by fluorinated oil and transferred into a long Teflon tubing for incubation. Crystallization trials can be set up in two ways, either using a premixed protein-laden LCP or by a post-LCP formation incorporation of protein (PLI). The PLI method was successfully used to crystallize bacteriorhodopsin and photosynthetic reaction center (Li et al. 2010). The plug-based system is relatively simple; however, reliable plug formation and merging, as well as plug storage and crystal imaging, present technical challenges.

11.6.6 *Crystal Imaging in LCP*

Imaging LCP crystallization drops requires, at minimum, a good microscope with a Koehler-type illumination and a strong but cold light source. The microscope should preferably be equipped with 10x and 40x objectives, and with rotating linear polarizers. Initial crystal hits in LCP are typically very small, micron-sized crystals, often in the form of thin needles or plates. Such crystals may have a very poor optical contrast inside of the gel-like LCP material, and their detection can be obscured by various LCP defects such as formation of cracks, droplets, or phase transformations. Using a linear polarizer and analyzer crossed at 90° can help to increase contrast and detect bright birefringent crystalline objects on the dark background of optically isotropic LCP. Most crystals, except for those with cubic symmetry, can rotate polarization of light; however, birefringence from protein crystals could be weak and often depends on crystal orientation with respect to the direction of polarization of the incident light. MP crystals can be distinguished from crystals of salts, ligands, lipids, etc., by imaging them with a UV-fluorescent microscope. This method of crystal detection is extremely useful for MPs enriched in tryptophans. Alternatively, MPs can be labeled with trace amounts of a fluorescent dye, which can facilitate detection of tiny crystals (Cherezov et al. 2010), but could also interfere with crystal nucleation and growth.

A new label-free technique for protein crystal detection, abbreviated as SONICC (second order nonlinear optical imaging of chiral crystals), has recently been introduced (Wampler et al. 2008). It is based on the effect of second harmonic generation (SHG) or frequency doubling, the propensity of which increases dramatically when chiral molecules are assembled in a crystal lattice. SONICC signal is not strongly affected by scattering and thus is very useful for crystal detection inside opaque or turbid media, such as different lipidic mesophases (Kissick et al. 2010).

All of these imaging modes have been automated and are available in modern crystal imagers. Most commercial imagers include a plate hotel, where crystallization plates can be incubated at a constant temperature and imaged following a user-defined schedule.

11.6.7 *Crystal Harvesting*

For harvesting purposes, crystals can be released from the gel-like LCP by several methods, such as enzymatic digestion by a lipase (Nollert and Landau 1998), incubation with low concentration of detergents (Luecke et al. 1999), dissolving lipids in oil (Cherezov et al. 2010), or transforming LCP into a sponge phase (Cherezov et al. 2006a). All of these interventions, however, can disturb crystals and affect their diffraction. The most gentle technique is to harvest crystals directly from LCP and immediately flash-freeze them in liquid nitrogen. MiTeGen micromounts are very convenient for this purpose as they are thin and rigid, allowing one to easily penetrate inside LCP and manipulate the crystals. The size of the micromount should be chosen so that it is sufficiently large to support the crystal, but not too large to avoid dragging excessive amounts of lipid along with the crystal. Crystal harvesting from LCP is usually performed under a stereomicroscope with a large working distance continuous zoom objective equipped with rotating linear polarizers. Polarizers are important because without them crystals become lost as soon as the surface of the LCP bolus is disturbed. With a little practice, harvesting of small crystals from LCP becomes easier than harvesting similar-sized crystals from aqueous solution, because crystals in LCP do not move away from the loop by the flow. When glass sandwich plates are used for crystallization, opening an individual well presents an additional challenge. Detailed written and visual protocols for crystal harvesting from glass sandwich plates have been published (Cherezov et al. 2010; Liu and Cherezov 2011; Li et al. 2012).

11.7 LCP Micro-Crystallography at Synchrotron Sources

MP crystals grown in LCP are often relatively small but better ordered compared to crystals obtained in detergent micelles. Developments of micro-crystallography during the past 15 years enabled high-resolution data collection from small MP crystals with sizes as small as 10 μm . Microfocused or collimated minibeam down to 5 μm in diameter are currently available at most modern synchrotron sources. Centering of small crystals harvested from LCP and surrounded by opaque frozen lipids has been facilitated by implementation of automated rastering procedures (Cherezov et al. 2009). SONICC has been integrated into beamlines to further shorten crystal location and centering time (Madden et al. 2013). Improvements in the X-ray beam stability, accuracy of the goniometers, detector sensitivity, and dynamic range allowed recording of crystallographic data from smaller and smaller crystals.

High-resolution data collection from small well-ordered crystals, however, is ultimately limited by the radiation damage. An estimated upper dose limit for frozen crystals is about 30 MGy (Owen et al. 2006). In the case of small crystals exposed to microfocus beams at the third-generation synchrotron sources, such a dose may correspond to only a few degrees of data. A common practice therefore consists

of merging 10–15° wedges of data obtained from dozens of microcrystals. Scaling data collected on multiple microcrystals is facilitated by high isomorphism, which is typically observed for LCP grown crystals in contrast to MP crystals grown in detergent micelles, which often are nonisomorphous (Pedersen et al. 2009).

Most structures of MPs crystallized in LCP were solved by molecular replacement. Experimental phasing of LCP-grown crystals is difficult due to the small crystal size. The first structures solved by experimental phasing from LCP-grown crystals started to appear only in 2012. The structure of channelrhodopsin was determined by multiple-wavelength anomalous diffraction (MAD) using mercury-derivatized crystals (Kato et al. 2012). The structure of a sodium/calcium exchanger was solved by single-wavelength anomalous diffraction (SAD) using crystals soaked with samarium compounds (Liao et al. 2012), and MAD phasing of SeMet crystals was used to determine the structure of an outer MP intimin (Fairman et al. 2012).

Recently, tantalum bromide cluster was used to obtain experimental phase information for two GPCRs: glucagon receptor from class B (Siu et al. 2013) and smoothed receptor from class Frizzled (Wang et al. 2013). Tantalum bromide cluster was incorporated by an overnight soaking, and a SAD dataset at the tantalum L3 edge (9.881 keV) was collected from a single crystal in each case. Although crystals were relatively small and weakly diffracting, the high phasing power of the tantalum bromide cluster enabled determination of phase information to approximately 6-Å resolution, followed by phase extension to the native high-resolution dataset.

11.8 LCP Serial Femtosecond Crystallography

While micro-crystallography at the third-generation synchrotron sources has come of age, new generation sources based on X-ray-free electron lasers (XFELs) are starting to emerge and demonstrate a promise for the next big leap toward nano-crystallography, room-temperature structures, and molecular movies (Fromme and Spence 2011). XFELs deliver ultimately short pulses of X-ray radiation of few femtoseconds in duration and extreme intensity that is a billion times higher than that of the most intense third-generation synchrotrons. Due to such an extremely short pulse duration, X-ray photons are able to diffract from macromolecules arranged in a crystal lattice before any conceivable damage sets in and the molecules are disintegrated by the extremely high radiation dose. The principle of “diffraction before destruction” was initially outlined in theoretical studies by Neutze et al. (2000).

The first experimental proof of this principle applied to macromolecular crystallography was achieved by Chapman et al. (2011) by determining an 8.4-Å resolution structure of a large MP complex, Photosystem I, using the first hard XFEL, at the Linac Coherent Light Source (LCLS) in Stanford, CA, USA. This work has introduced a new approach of serial femtosecond crystallography (SFX) that is based on a continuous supply of crystals to the interaction point with an XFEL beam. Hundreds of thousands of diffraction patterns are collected at room temperature from hydrated microcrystals at single orientations. These data are then processed by

Monte Carlo integration methods (Kirian et al. 2011) to yield the structure factors that are used for structure determination. Another low-resolution structure was also obtained for an MP, photosynthetic reaction center, grown in a lipidic sponge phase by the SFX approach (Johansson et al. 2012). The first high-resolution structure determined by SFX was a 1.9-Å structure of a well-behaving soluble protein lysozyme (Boutet et al. 2012). In all of these experiments, microcrystals were delivered within a liquid suspension streamed from a gas dynamic virtual nozzle (GDVN) injector (DePonte et al. 2008). The GDVN injector produces a fast-running liquid microjet focused by a gas flow to a diameter of 1–5 μm. The typical flow rate of a stable microjet is 10 μL/min, which is too fast for the 120-Hz pulse rate at LCLS, resulting in only one crystal out of about 10,000 being shot by the XFEL beam. Acquisition of a complete dataset therefore requires tens to hundreds of milligrams of crystallized protein, which is, in practice, unachievable for most MPs.

To overcome the problem of the fast-running microjet and to enable SFX data collection on MP microcrystals grown in LCP, a special LCP injector was designed (Weierstall et al. 2014). The LCP injector consists of a hydraulic stage, a sample reservoir, and a nozzle. The LCP reservoir is connected to a borosilicate capillary with 10–50-μm inner diameter, and the LCP is extruded out of this capillary into vacuum. This injector takes advantage of the viscous gel-like properties of LCP, allowing for the control of the flow rate within the optimal range matching the 120-Hz XFEL pulse rate at LCLS, so that the crystals are not wasted. The LCP injector was recently used to obtain the first room-temperature structure of a GPCR at 2.8-Å resolution using sub-10-μm crystals and less than 300 μg of crystallized receptor (Liu et al. 2013).

11.9 Conclusions

Significant progress has been achieved in our understanding of MPs in the past few years aided by the tremendous developments in structural biology tools and methods applicable to these molecules. The rate of structure determination for new MPs, however, is still lagging behind that for soluble proteins. Eukaryotic MPs are especially problematic because of their low expression yields and stability. LCP provides a convenient lipid matrix mimicking a native membrane, which is suitable for stabilization and crystallization of MPs. Recent developments of LCP-related techniques have facilitated the research in this field, leading to breakthroughs in structural studies of several crucial targets. These developments also simplified handling of LCP materials making it possible for laboratories of all sizes including small single PI groups to work with this powerful approach which was at one time thought to be too complex and difficult to pursue.

Several pre-crystallization LCP assays have been developed to rationally guide LCP crystallization trials. These procedures help save valuable MP samples and accelerate the process of obtaining initial crystal hits. MP function can also be characterized in a number of ways, taking advantage of the porous and bicontinuous

nature of LCP. During the last decade, a variety of LCP tools have appeared to help with handling this highly viscous gel-like material. Most of these instruments and tools for LCP preparation, dispensing, setting up crystallization trials, and harvesting crystals are now commercially available and accessible to researchers. Apart from manual tools, automated robotic systems were developed to perform crystallization trials in a HT manner.

Microfluidics show potential to become an alternative to HT robotic systems. The validity of this approach was demonstrated by crystallizing several model MPs in LCP. Further progress in this area can lead to a next step in minimizing material consumption and integration of all stages in structural studies of MPs, including cell-free expression, purification, crystallization, and crystallographic data collection within a single lab-on-a-chip.

New XFEL technologies continue to emerge as a major trend in modern structural biology. Using LCP for delivering microcrystals, the LCP-SFX approach to MP structure determination should accelerate the pace of structural work on challenging human MPs and their complexes with soluble partners, and enable time-resolved studies of intermediate states, thereby advancing our knowledge of the functional mechanisms of these biomedically important proteins.

Acknowledgments This work was supported in part by the NIH Common Fund in Structural Biology grant P50 GM073197. We thank K. Kadyshchinskaya for making the figures, A. Walker and C. Klasen for assistance with the manuscript preparation.

References

- Ai X, Caffrey M (2000) Membrane protein crystallization in lipidic mesophases: detergent effects. *Biophys J* 79:394–405
- Barauskas J, Landh T (2003) Phase behavior of the phytantriol/water system. *Langmuir* 19:9562–9565
- Borshchevskiy V, Moiseeva E, Kuklin A, Buldt G, Hato M, Gordeliy V (2010) Isoprenoid-chained lipid β -XylOC16+4-A novel molecule for in meso membrane protein crystallization. *J Cryst Growth* 312:3326–3330
- Boutet S, Lomb L, Williams GJ et al (2012) High-resolution protein structure determination by serial femtosecond crystallography. *Science* 337:362–364
- Briggs J, Chung H, Caffrey M (1996) The temperature-composition phase diagram and mesophase structure characterization of the monoolein/water system. *J Physique II* 6:723–751
- Caffrey M (2008) On the mechanism of membrane protein crystallization in lipidic mesophases. *Cryst Growth Des* 8:4244–4254
- Caffrey M (2009) Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu Rev Biophys* 38:29–51
- Caffrey M, Cherezov V (2009) Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc* 4:706–731
- Caffrey M, Porter C (2010) Crystallizing membrane proteins for structure determination using lipidic mesophases. *J Vis Exp* 45:e1712. doi:10.3791/1712
- Chapman HN, Fromme P, Barty A et al (2011) Femtosecond X-ray protein nanocrystallography. *Nature* 470:73–77

- Cheng A, Hummel B, Qiu H, Caffrey M (1998) A simple mechanical mixer for small viscous lipid-containing samples. *Chem Phys Lipids* 95:11–21
- Cherezov V (2011) Lipidic cubic phase technologies for membrane protein structural studies. *Curr Opin Struct Biol* 21:559–566
- Cherezov V, Caffrey M (2003) Nano-volume plates with excellent optical properties for fast, inexpensive crystallization screening of membrane proteins. *J Appl Crystallogr* 36:1372–1377
- Cherezov V, Caffrey M (2005) A simple and inexpensive nanoliter-volume dispenser for highly viscous materials used in membrane protein crystallization. *J Appl Crystallogr* 38:398–400
- Cherezov V, Caffrey M (2007) Membrane protein crystallization in lipidic mesophases. A mechanism study using X-ray microdiffraction. *Faraday Discuss* 136:195–212
- Cherezov V, Fersi H, Caffrey M (2001) Crystallization screens: compatibility with the lipidic cubic phase for in meso crystallization of membrane proteins. *Biophys J* 81:225–242
- Cherezov V, Clogston J, Misquitta Y, Abdel-Gawad W, Caffrey M (2002) Membrane protein crystallization in meso: lipid type-tailoring of the cubic phase. *Biophys J* 83:3393–3407
- Cherezov V, Peddi A, Muthusubramaniam L, Zheng YF, Caffrey M (2004) A robotic system for crystallizing membrane and soluble proteins in lipidic mesophases. *Acta Crystallogr D Biol Crystallogr* 60:1795–1807
- Cherezov V, Clogston J, Papiz MZ, Caffrey M (2006a) Room to move: crystallizing membrane proteins in swollen lipidic mesophases. *J Mol Biol* 357:1605–1618
- Cherezov V, Yamashita E, Liu W, Zhahnina M, Cramer WA, Caffrey M (2006b) In meso structure of the cobalamin transporter, BtuB, at 1.95 Å resolution. *J Mol Biol* 364:716–734
- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* 318:1258–1265
- Cherezov V, Liu J, Hanson MA, Griffith MT, Stevens R (2008) LCP-FRAP Assay for pre-screening membrane proteins for in meso crystallization. *Cryst Growth Des* 8:4307–4315
- Cherezov V, Hanson MA, Griffith MT, Hilgart MC, Sanishvili R, Nagarajan V, Stepanov S, Fischetti RF, Kuhn P, Stevens RC (2009) Rastering strategy for screening and centring of microcrystal samples of human membrane proteins with a sub-10 µm size X-ray synchrotron beam. *J R Soc Interface* 6:S587–S597
- Cherezov V, Abola E, Stevens RC (2010) Recent progress in the structure determination of GPCRs, a membrane protein family with high potential as pharmaceutical targets. *Method Mol Biol* 654:141–168
- Clogston J (2005) Applications of the lipidic cubic phase: from controlled release and uptake to in meso crystallization of membrane proteins. Dissertation, The Ohio State University
- Chun E, Thompson AA, Liu W, Roth CR, Griffith MT, Katritch V, Kunken J, Xu F, Cherezov V, Hanson MA, Stevens RC (2012) Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* 20:967–976
- Coleman BE, Cwynar V, Hart DJ, Havas F, Mohan JM, Patterson S, Ridenour S, Schmidt M, Smith E, Wells AJ (2004) Modular approach to the synthesis of unsaturated 1-monoacyl glycerols. *Synlett* 8:1339–1342
- Congreve M, Langmead C, Marshall FH (2011) The use of GPCR structures in drug design. *Adv Pharmacol* 62:1–36
- Conn CE, Darmanin C, Mulet X, Kirby N, Drummond CJ (2012) High-throughput in situ analysis of the structural evolution of the monoolein cubic phase under crystallogensis conditions. *Soft Matter* 8:2310–2321
- Darmanin C, Conn CE, Newman J, Mulet J, Seabrook SA, Liang Y-L, Hawley A, Kirby N, Varghese JN, Drummond CJ (2012) High-throughput production and structural characterization of libraries of self-assembly lipidic cubic phase materials. *ACS Comb Sci* 14:247–252
- Deisenhofer J, Epp O, Miki K, Huber R, Michel H (1985) Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3 Å resolution. *Nature* 318:618–624

- DePonte DP, Weierstall U, Schmidt K, Warner J, Starodub D, Spence JCH, Doak RB (2008) Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J Phys D: Appl Phys* 41:195505
- Fagerberg L, Jonasson K, von Heijne G, Uhlen M, Berglund L (2010) Prediction of the human membrane proteome. *Proteomics* 10:1141–1149
- Faham S, Bowie JU (2002) Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J Mol Biol* 316:1–6
- Fairman JW, Dautin N, Wojtowicz D, Liu W, Noinaj N, Barnard TJ, Udho E, Przytycka TM, Cherezov V, Buchanan SK (2012) Crystal structures of the outer membrane domain of intimin and invasins from enterohemorrhagic *E. coli* and enteropathogenic *Y. pseudotuberculosis*. *Structure* 20:1233–1243
- Fromme P, Spence JCH (2011) Femtosecond nanocrystallography using X-ray lasers for membrane protein structure determination. *Curr Opin Struct Biol* 21:509–516
- Fu Y, Weng Y, Hong W-X, Zhang Q (2010) Efficient synthesis of unsaturated 1-Monoacyl glycerols for in meso crystallization of membrane proteins. *Synlett* 2011:809–812
- Hanson MA, Brooun A, Baker KA, Jaakola VP, Roth C, Chien EY, Alexandrov A, Velasquez J, Davis L, Griffith M, Moy K, Ganser-Pornillos BK, Hua Y, Kuhn P, Ellis S, Yeager M, Stevens RC (2007) Profiling of membrane protein variants in a baculovirus system by coupling cell-surface detection with small-scale parallel expression. *Protein Expr Purif* 56:85–92
- Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola VP, Chien EY, Velasquez J, Kuhn P, Stevens RC (2008) A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor. *Structure* 16:897–905
- Hato M, Yamashita J, Shiono M (2009) Aqueous phase behavior of lipids with isoprenoid type hydrophobic chains. *J Phys Chem B* 113:10196–10209
- Hyde ST, Andersson S, Ericsson B, Larsson K (1984) A cubic structure consisting of a lipid bilayer forming an infinite periodic minimum surface of the gyroid type in the glycerol/monooleate-water system. *Z Kristallogr* 168:213–219
- Johansson LC, Arnlund D, White TA et al (2012) Lipidic phase membrane protein serial femtosecond crystallography. *Nat Method* 9:263–265
- Joseph JS, Liu W, Kunken J, Weiss TM, Tsuruta H, Cherezov V (2011) Characterization of lipid matrices for membrane protein crystallization by high-throughput small angle X-ray scattering. *Methods* 55:342–349
- Kato HE, Zhang F, Yizhar O, Ramakrishnan C, Nishizawa T, Hirata K, Ito J, Aita Y, Tsukazaki T, Hayashi S, Hegemann P, Maturana AD, Ishitani R, Deisseroth K, Nureki O (2012) Crystal structure of the channelrhodopsin light-gated cation channel. *Nature* 482:369–374
- Katritch V, Cherezov V, Stevens RC (2013) Structure-function of the G protein-coupled receptor superfamily. *Annu Rev Pharmacol Toxicol* 53:531–556
- Kirian RA, White TA, Holton JM, Chapman HN, Fromme P, Barty A, Lomb L, Aquila A, Maia FR, Martin AV, Fromme R, Wang X, Hunter MS, Schmidt KE, Spence JC (2011) Structure-factor analysis of femtosecond microdiffraction patterns from protein nanocrystals. *Acta Crystallogr A* 67:131–140
- Kissick DJ, Gualtieri EJ, Simpson GJ, Cherezov V (2010) Nonlinear optical imaging of integral membrane protein crystals in lipidic mesophases. *Analyt Chem* 82:491–497
- Kors CA, Wallace E, Davies DR, Li L, Laible PD, Nollert P (2009) Effects of impurities on membrane-protein crystallization in different systems. *Acta Crystallogr D: Biol Crystallogr* 65:1062–1073
- Kulkarni CV, Wachter W, Iglesias-Salto G, Engelskirchen S, Ahualli S (2011) Monoolein: a magic lipid? *Phys Chem Chem Phys* 13:3004–3021
- Landau EM, Rosenbusch JP (1996) Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci U S A* 93:14532–14535
- Li D, Caffrey M (2011) Lipid cubic phase as a membrane mimetic for integral membrane protein enzymes. *Proc Natl Acad Sci U S A* 108:8639–8644

- Li D, Boland C, Aragao D, Walsh K, Caffrey M (2012) Harvesting and cryo-cooling crystals of membrane proteins grown in lipidic mesophases for structure determination by macromolecular crystallography. *J Vis Exp* 67:e4001. doi:10.3791/4001
- Li D, Lee J, Caffrey M (2011) Crystallizing membrane proteins in lipidic mesophases. A host lipid screen. *Cryst Growth Des* 11:530–537
- Li D, Lyons JA, Pye VE, Vogeley L, Aragao D, Kenyon CP, Shah ST, Doherty C, Aherne M, Caffrey M (2013a) Crystal structure of the integral membrane diacylglycerol kinase. *Nature* 497:521–524
- Li D, Shah ST, Caffrey M (2013b) Host lipid and temperature as important screening variables for crystallizing integral membrane proteins in lipidic mesophases. Trials with diacylglycerol kinase. *Cryst Growth Des* 13:2846–2857
- Li L, Fu Q, Kors CA, Stewart L, Nollert P, Laible PD, Ismagilov RF (2010) A Plug-based microfluidic system for dispensing Lipidic Cubic Phase (LCP) material validated by crystallizing membrane proteins in lipidic mesophases. *Microfluid Nanofluidics* 8:789–798
- Liao J, Li H, Zeng W, Sauer DB, Belmares R, Jiang Y (2012) Structural insight into the ion-exchange mechanism of the sodium/calcium exchanger. *Science* 335:686–690
- Liu W, Caffrey MC (2005) Gramicidin structure and disposition in highly curved membranes. *J Struct Biol* 150:23–40
- Liu W, Cherezov V (2011) Crystallization of membrane proteins in lipidic mesophases. *J Vis Exp* 49:e2501. doi:10.3791/2501
- Liu W, Hanson MA, Stevens RC, Cherezov V (2010) LCP-Tm: an assay to measure and understand stability of membrane proteins in a membrane environment. *Biophys J* 98:1539–1548
- Liu W, Wacker D, Gati C et al (2013) Serial femtosecond crystallography of G protein-coupled receptors in lipidic cubic phase. *Science* 342:1521–1524
- Luecke H, Schobert B, Richter H-T, Cartailler J-P, Lanyi JK (1999) Structure of bacteriorhodopsin at 1.55 Å resolution. *J Mol Biol* 291:899–911
- Madden JT, Toth SJ, Dettmar CM, Newman JA, Oglesbee RA, Hedderich HG, Everly RM, Becker M, Ronau JA, Buchanan SK, Cherezov V, Morrow ME, Xu S, Ferguson D, Makarov O, Das C, Fischetti R, Simpson GJ (2013) Integrated nonlinear optical imaging microscope for on-axis crystal detection and centering at a synchrotron beamline. *J Synchrotron Radiat* 20:531–540
- McPherson A (2004) Introduction to protein crystallization. *Methods* 34:254–265
- Michel H (1991) General and practical aspects of membrane protein crystallization. *Crystallization of membrane proteins*. CRC Press Inc., Boca Raton, pp 73–88
- Misquitta Y, Caffrey M (2003) Detergents destabilize the cubic phase of monoolein: implications for membrane protein crystallization. *Biophys J* 85:3084–3096
- Misquitta Y, Cherezov V, Havas F, Patterson S, Mohan JM, Wells AJ, Hart DJ, Caffrey M (2004a) Rational design of lipid for membrane protein crystallization. *J Struct Biol* 148:169–175
- Misquitta LV, Misquitta Y, Cherezov V, Slattery O, Mohan JM, Hart D, Zhahnina M, Cramer WA, Caffrey M (2004b) Membrane protein crystallization in lipidic mesophases with tailored bilayers. *Structure* 12:2113–2124
- Murgia S, Caboi F, Monduzzi M, Ljusberg-Wahren H, Nylander T (2002) Acyl migration and hydrolysis in monoolein-based systems. *Progr Colloid Polym Sci* 120:41–46
- Navarro J, Landau EM, Fahmy K (2002) Receptor-dependent G-protein activation in lipidic cubic phase. *Biopolymers* 67:167–177
- Nazaruk E, Bilewicz R, Lindblom G, Lindholm-Sethson B (2008) Cubic phases in biosensing systems. *Analyt Bioanal Chem* 391:1569–1578
- Neutze R, Wouts R, Van der Spoel D, Weckert E, Hajdu J (2000) Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* 406:752–757
- Nollert P, Landau EM (1998) Enzymic release of crystals from lipidic cubic phases. *Biochem Soc Trans* 26:709–713
- Nollert P, Qiu H, Caffrey M, Rosenbusch JP, Landau EM (2001) Molecular mechanism for the crystallization of bacteriorhodopsin in lipidic cubic phases. *FEBS Lett* 504:179–186
- Owen RL, Rudiño-Piñera E, Garman EF (2006) Experimental determination of the radiation dose limit for cryocooled protein crystals. *Proc Natl Acad Sci U S A* 103:4912–4917

- Pebay-Peyroula E, Rummel G, Rosenbusch JP, Landau EM (1997) X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science* 277:1676–1681
- Pedersen BP, Morth JP, Nissen P (2009) Structure determination using poorly diffracting membrane-protein crystals: the H⁺-ATPase and Na⁺, K⁺-ATPase case history. *Acta Crystallogr D Biol Crystallogr* 66:309–313
- Perry SLS, Roberts GWG, Tice JDJ, Gennis RB, Kenis PJ (2009) Microfluidic generation of lipidic mesophases for membrane protein crystallization. *Cryst Growth Des* 9:2566–2569
- Pieper U, Schlessinger A, Kloppmann E, Chang GA, Chou JJ, Dumont ME, Fox BG, Fromme P, Hendrickson WA, Malkowski MG, Rees DC, Stokes DL, Stowell MH, Wiener MC, Rost B, Stroud RM, Stevens RC, Sali A (2013) Coordinating the impact of structural genomics on the human α -helical transmembrane proteome. *Nat Struct Mol Biol* 20:135–138
- Qiu H, Caffrey M (2000) The phase diagram of the monoolein/water system: metastability and equilibrium aspects. *Biomaterials* 21:223–234
- Rask-Andersen M, Almen MS, Schioth HB (2011) Trends in the exploitation of novel drug targets. *Nat Rev Drug Discov* 10:579–590
- Rasmussen SGF, Choi H-J, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* 450:383–387
- Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK (2011) Crystal structure of the β 2 adrenergic receptor-Gs protein complex. *Nature* 477:549–555
- Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC, Kobilka BK (2007) GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* 318:1266–1273
- Rummel G, Hardmeyer A, Widmer C, Chiu ML, Nollert P, Locher KP, Pedruzzi II, Landau EM, Rosenbusch JP (1998) Lipidic cubic phases: new matrices for the three-dimensional crystallization of membrane proteins. *J Struct Biol* 121:82–91
- Sennoga C, Heron A, Seddon JM, Templar RH, Hankamer B (2003) Membrane-protein crystallization in cubo: temperature-dependent phase behaviour of monoolein-detergent mixtures. *Acta Crystallogr D Biol Crystallogr* 59:239–246
- Shah JC, Sadhale Y, Chilukuri DM (2001) Cubic phase gels as drug delivery systems. *Adv Drug Deliver Rev* 47:229–250
- Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS, Liu W, Lau J, Cherezov V, Katritch V, Wang MW, Stevens RC (2013) Structure of the human glucagon class B G-protein-coupled receptor. *Nature* 499:444–449
- Smith JL, Fischetti RF, Yamamoto M (2012) Micro-crystallography comes of age. *Curr Opin Struct Biol* 22:602–612
- Stevens RC, Cherezov V, Katritch V, Abagyan R, Kuhn P, Rosen H, Wuthrich K (2013) The GPCR network: a large-scale collaboration to determine human GPCR structure and function. *Nat Rev Drug Discov* 12:25–34
- Takeda K, Sato H, Hino T, Kono M, Fukuda K, Sakurai I, Okada T, Kouyama T (1998) A novel three-dimensional crystal of bacteriorhodopsin obtained by successive fusion of the vesicular assemblies. *J Mol Biol* 283:463–474
- Toepke MW, Beebe DJ (2006) PDMS absorption of small molecules and consequences in microfluidic applications. *Lab Chip* 6:1484–1486
- Wacker D, Fenalti G, Brown MA, Katritch V, Abagyan R, Cherezov V, Stevens RC (2010) Conserved binding mode of human beta2-adrenergic receptor inverse agonists and antagonists revealed by X-ray crystallography. *J Am Chem Soc* 132:11443–11445
- Wadsten P, Wöhri AB, Snijder A, Katona G, Gardiner AT, Cogdell RJ, Neutze R, Engström S (2006) Lipidic sponge phase crystallization of membrane proteins. *J Mol Biol* 364:44–53

- Wampler RD, Kissick DJ, Dehen CJ, Gualtieri EJ, Grey JL, Wang HF, Thompson DH, Cheng JX, Simpson GJ (2008) Selective detection of protein crystals by second harmonic microscopy. *J Am Chem Soc* 130:14076–14077
- Wang H, Nieh M, Hobbie E, Glinka CJ, Katsaras J (2003) Kinetic pathway of the bilayered-micelle to perforated-lamellae transition. *Phys Rev E* 67:060902
- Wang C, Wu H, Katritch V, Han GW, Huang X-P, Liu W, Siu FY, Roth BL, Cherezov V, Stevens RC (2013) Structure of the human smoothed receptor bound to an antitumour agent. *Nature* 497:338–343
- Whiles JA, Deems R, Vold RR, Dennis EA (2002) Bicelles in structure-function studies of membrane-associated proteins. *Bioorg Chem* 30:431–442
- Wiener MC (2004) A pedestrian guide to membrane protein crystallization. *Methods* 34:364–372
- Weierstall U, James D, Wang D et al (2014) Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat Commun* 5:3309
- Xu F, Liu W, Hanson MA, Stevens RC, Cherezov V (2011) Development of an automated high throughput LCP-FRAP assay to guide membrane protein crystallization in lipid mesophases. *Cryst Growth Des* 11:1193–1201
- Yamashita J, Shiono M, Hato M (2008) New lipid family that forms inverted cubic phases in equilibrium with excess water: molecular structure–aqueous phase structure relationship for lipids with 5, 9, 13, 17-tetramethyloctadecyl and 5, 9, 13, 17-tetramethyloctadecanoyl chains. *J Phys Chem B* 112:12286–12296

Chapter 12

Micelles, Bicelles, Amphipols, Nanodiscs, Liposomes, or Intact Cells: The Hitchhiker's Guide to the Study of Membrane Proteins by NMR

Laurent J. Catoire, Xavier L. Warnet and Dror E. Warschawski

12.1 Introduction

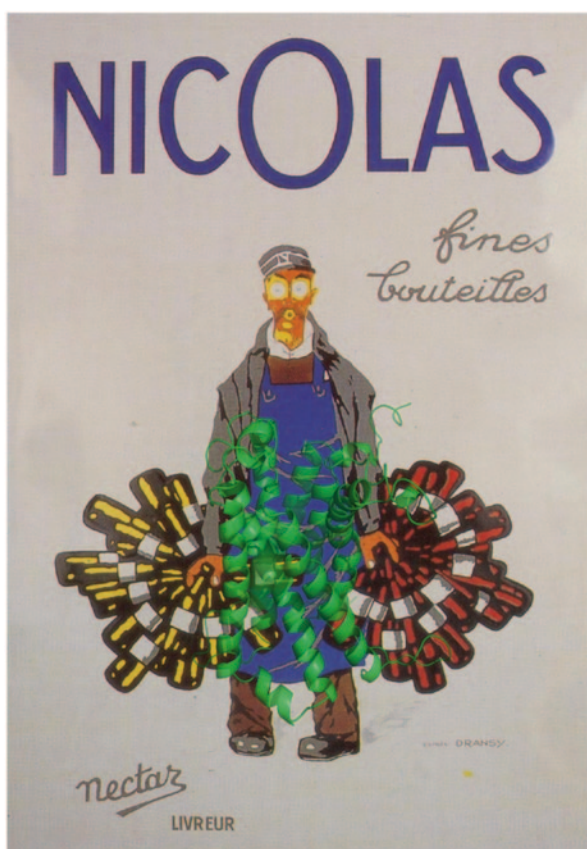
The vast majority of biophysical studies of membrane proteins (MPs) at the atomic scale are performed *in vitro* with preparations as homogeneous as possible, where the protein is isolated in a nonnative environment. MP samples for nuclear magnetic resonance (NMR) spectroscopy are no exception to the rule, in particular because purification helps to clearly detect and unambiguously identify signals from the protein of interest. By *native environment*, we consider the original membrane(s) where the protein exerts its biological role. It is difficult to place artificial systems on a scale defining how well they mimic native membranes, especially when a functional test is difficult or impossible to set up. Indeed, liposomes or nanometric lipid bilayers still represent artificial environments, and, on the contrary, exotic surfactants like amphipols, which could be thought to be inappropriate given their chemical structures, have proven to keep numerous MPs stable and active in solution. Over the past decades, various membrane mimetics have been developed, chosen on the basis of the compatibility with the technique of investigation used, sometimes at the expense of the functionality of the protein. Paradoxically, after so many efforts to improve membrane substitutes, *in-cell* NMR has known significant advances during the past few years (Selenko and Wagner 2007; Ito and Selenko 2010), which represents a very attractive potential for future NMR studies of MPs *in situ* (Renault et al. 2012a, 2012b).

L. J. Catoire (✉) · X. L. Warnet · D. E. Warschawski (✉)
Laboratory of Physico-Chemical Biology of Membrane Proteins, UMR-CNRS 7099,
Institute of Physico-Chemical Biology, and Université Paris Diderot, Paris, France
e-mail: laurent.catoire@ibpc.fr

D. E. Warschawski
e-mail: dror@ibpc.fr

In this chapter, we describe the different environments available and their applications to MP studies by NMR spectroscopy. We treated solution- and solid-state NMR separately because sample preparations and methodologies are different, even though some environments are common to these two subtypes of NMR. In theory, protein size for solution-state NMR is limited, not in solid-state (vide infra § 12.3.1). Additional equipment is also required for solid-state NMR, such as high-power amplifiers, air compressor and dryer, etc. In order to be concise, each MP environment, with its own advantages and drawbacks, is briefly described; readers interested in more complete descriptions can find an exhaustive bibliography in Warschawski et al. (2011).

12.2 Solution-State NMR



Adapted from Dransy, 1923, courtesy of Nicolas

12.2.1 Detergents

12.2.1.1 Generalities

For almost 40 years, detergents were used to characterize MPs in aqueous solutions (Helenius and Simons 1975; Tanford and Reynolds 1976), and most of NMR structural studies of MPs performed to date have been carried out in detergent solutions (Kang and Li 2011; Warschawski 2013). These molecules are amphiphilic, i.e., they possess both hydrophilic and hydrophobic parts, usually dubbed *head* and *tail*, respectively. Above a certain concentration and temperature, i.e., the critical micellar concentration (cmc) and *Kraft* temperature, detergent monomers form aggregates named *micelles*, and any addition of molecules of detergent create new micelles. In an aqueous solution, above the cmc, the hydrophilic heads are in contact with water molecules and the tails are in contact between each other. There is equilibrium between molecules of detergent associated in micelles with those existing as monomers (Fig. 12.1a). The form and size of micelles depend on the detergent chemical structure and experimental conditions, such as temperature, pH, and ionic strength. For instance, the detergent dodecyl- β -maltoside (β -DDM or C12-M), one of the most used detergents in structural biology, forms large oblate micelles in typical experimental conditions (Oliver et al. 2013), while dodecylphosphocholine (DPC or C12-PC, Fos-Choline-12 or MAPCHO-12) adopts preferentially a spherical shape (Tieleman et al. 2000). Detergents associated to MPs form complexes with a relatively small size compared to other solubilizing agents, and this is the main reason why they are the most frequently used molecules for solution-state NMR investigations of MPs.

Other detergent-based systems such as lipopeptide detergents (McGregor et al. 2003; Privé 2011) or peptide surfactants (Zhao et al. 2006) have been shown to substantially improve the stability of MPs in aqueous solutions compared to traditional detergents (McGregor et al. 2003; Yeh et al. 2005). The acyl-chain packing of these surfactants is more uniform compared to micelles, with a lateral pressure more comparable to the interior of a bilayer. They also display a low cmc, usually below the micromolar range, rendering them less dissociating than detergents (Privé 2011). These surfactants are interesting for solution-state NMR studies of MPs as they form complex sizes, once associated to MPs, similar to those measured with detergents. Indeed, the β -barrel protein PagP associated to lipopeptides gave rise to high-resolution NMR spectra (McGregor et al. 2003). Perhaps, one of the main drawbacks today remains the cost to produce these alternative molecules.

12.2.1.2 Illustrations

The glycoporphin A was the first MP structure to be solved by NMR (MacKenzie et al. 1997, Fig. 12.1b). The experiments were conducted in DPC micelles. This study represented a major achievement, in both NMR and biochemistry: (1) by demonstrating the capacity of solution NMR to determine MP structures and

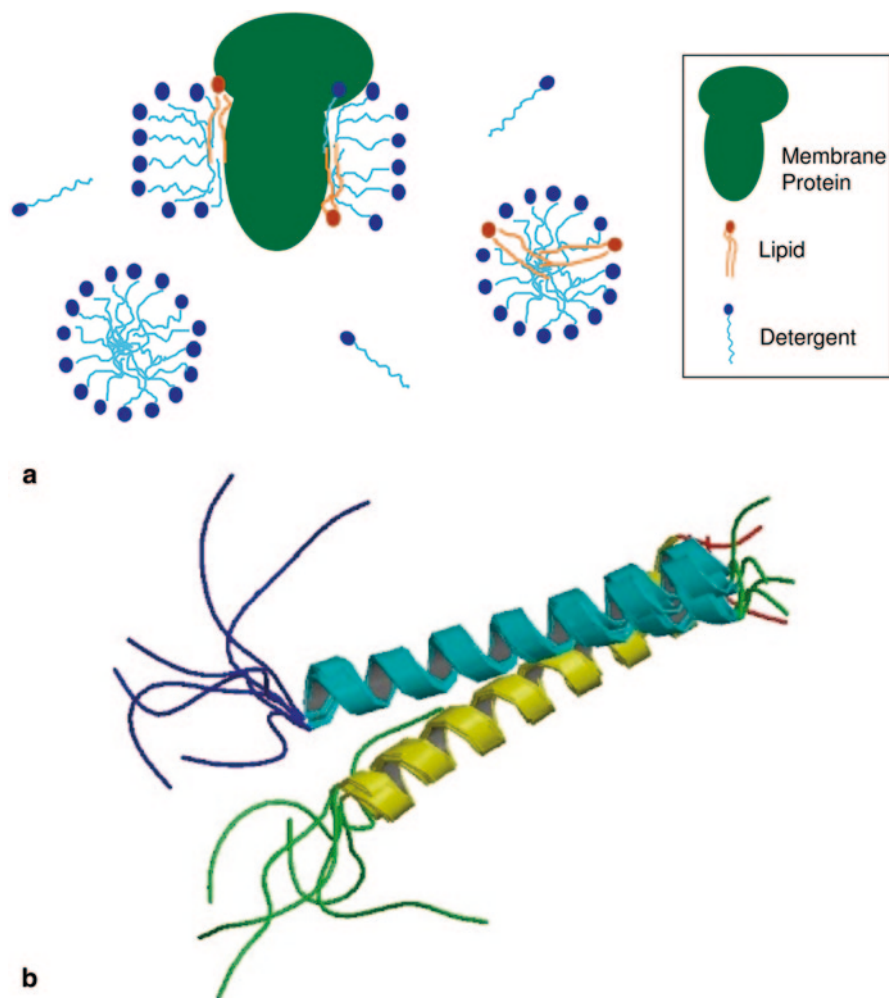


Fig. 12.1 NMR studies of MPs in detergent solutions. **a** Schematic view of the coexisting entities in a detergent solution, the MP (in *green*), lipid cofactors (in *orange*), and detergent molecules (in *blue*). **b** The first three-dimensional structure of an MP determined by NMR: the dimeric TM domain of glycophorin A (GpA), a 40-residue peptide, in a detergent DPC solution. (Image from the RCSB PDB (www.pdb.org) of PDB ID 1AFO; MacKenzie et al. 1997)

(2) by maintaining the oligomeric state of the protein in a detergent solution, which is essentially based on van der Waals interactions in this case. Since then, more than 100 MP structures have been solved de novo by NMR in detergent solutions (Warschawski 2013), from small size bitopic MPs (e.g., Teriete et al. 2007; Lau et al. 2009; Yang et al. 2009; Wittlich et al. 2010) to larger systems (>30 kDa, which is equivalent to >70 kDa with the belt of surfactant, e.g., Schnell and Chou 2008; Hiller et al. 2008; Gautier et al. 2010). In addition, solution-state NMR studies of

MPs in detergent solutions can be used to get important structural and dynamics information, without performing a full structure determination. For instance, with the help of G protein-coupled receptor (GPCR) crystal structures, NMR spectroscopy brought fundamental observations regarding the activation of these receptors (Bokoch et al. 2010; Kofuku et al. 2012; Liu et al. 2012; Nygaard et al. 2013).

12.2.1.3 Advantages

One of the best advantages using detergent for solution-state NMR is the resulting size of protein–detergents complexes that are usually smaller than complexes obtained with other classes of surfactant, despite their high propensity to inactivate MPs (Bowie 2001; Popot 2010). Novel promising detergents are regularly proposed, such as maltose-neopentyl glycol diacyl molecules (MNGs). Indeed, the thermal stability of several MPs could be substantially improved thanks to these new amphiphiles, such as the human β_2 adrenergic receptor-T4 lysozyme fusion protein or the muscarinic M3 acetylcholine receptor (Chae et al. 2010).

12.2.1.4 Drawbacks

Detergents tend to destabilize MPs, essentially by disrupting intraprotein, protein–protein, or protein–lipid interactions (Fig. 12.1a). These dissociating properties explain why they were originally used to extract MPs from their native or host membranes. For a given detergent, it is usually recommended to work close to the cmc in order to reduce the presence of protein-free micelles that could absorb lipid cofactors that are essential for the stability and/or activity of the protein. Regarding the concentrations of protein and detergent required to perform an NMR study (e.g., McDonnell and Opella 1993; Arora et al. 2001), i.e., well above the cmc, the probability of working with an inactive protein is high. This is why, following the structure of the glycophorin A, pioneering studies in detergent solutions by NMR were carried out on rugged β -barrel MPs from *Escherichia coli* (*E. coli*; Arora et al. 2001; Fernández et al. 2001; Hwang et al. 2002).

Sometimes, especially with detergents that form spherical micelles, the organization of hydrophobic chains could not always accommodate MPs very well, which can be revealed by variations in NMR protein chemical shifts compared to other media. For instance, OmpX exhibits various backbone $^{15}\text{N}/^1\text{H}^N$ chemical shifts depending on the surfactant used (Fernández et al. 2001; Lee et al. 2008; Hagn et al. 2013). This mainly reflects modifications in the structure of the protein instead of transmembrane (TM) electronic environment variations, as whatever the surfactant used, the amino acids pointing towards the membrane mostly face CH_n moieties of the surfactants.

Detergents can display a marked influence on the equilibrium kinetics between distinct MPs substates, depending of the cmc that is directly related to the detergent off-rate. This may be a drawback in studies that aim at looking at intra-MP

conformational exchanges. In a comparative study between the β 2-adrenergic receptor solubilized in either β -DDM or MNG3 detergent solutions, different conformational exchanges of the GPCR have been observed: With β -DDM, that has a cmc four orders of magnitude higher than MNG3, faster exchanges between distinct functional states could be observed compared to the receptor associated to MNG3 (Chung et al. 2012).

Besides difficulties to maintain active or native-folded MPs with detergents, another drawback concerns the choice of the detergent to be used to perform NMR studies. Usually, any NMR investigation relies on an empirical screening of detergents and concentrations, which is quite demanding in time and costly. This is also one of the reasons that led to the development of new alternatives to conventional detergents, like bicelles (Sanders and Landis 1995), amphipols (Tribet et al. 1996), nanodiscs (Bayburt et al. 2002), lipopeptide detergents (McGregor et al. 2003), peptide surfactants (Zhao et al. 2006), or new milder detergents (Chae et al. 2010). Some of those alternative media are quite universal and can be used by following general rules. Despite the fact that most alternative media give rise to larger ensembles compared to detergents, they represent a powerful substitute, thanks to improvements in NMR methodology and instrumentation, and also in the development of new isotope-labeling schemes dedicated to the study of large proteins or protein complexes (e.g., Plevin and Boisbouvier 2012). These new environments allow the detection of well-resolved MP NMR signals (*vide infra*).

12.2.2 *Mixed Detergent Solutions, Bicelles, or Detergents/Lipids Potpourris*

12.2.2.1 Generalities

In order to improve the stability of MPs in detergent solutions, the simplest solution consists of adding lipids to detergent micelles. These binary or more complex assemblies are usually named *mixed micelles*. Indeed, lipid cofactors are known to be crucial for the activity or stability of many MPs (Lee 2004). For instance, the *sn*-1,2-diacylglycerol kinase of *E. coli* requires lipid cofactors to be active (Walsh and Bell 1986). Assays performed in mixed octylglucoside/dimirystoylphosphatidylcholine (OG/DMPC) micellar systems, showed a protein 50-fold more active compared to pure OG micelles (Badola and Sanders 1997). More recently, high-resolution atomic structures of GPCRs have revealed the presence of a conserved sterol-binding site along some TM helices (Cherezov et al. 2007; Hanson et al. 2008; Wacker et al. 2010; Warne et al. 2011; Rosenbaum et al. 2011). Resulting tests of stability performed in β -DDM/cholesterol hemisuccinate mixed micelles demonstrated an increase in stability by $\sim 12^\circ\text{C}$ compared to the same measurements performed in a pure β -DDM solution (Thompson et al. 2011). Mixed micelles have also been used for solution-state NMR studies. For instance, the low-resolution structure of

the mitochondrial uncoupling protein 2 could be solved in mixed micelles of DPC and DMPC (Berardi et al. 2011), and low amount of mixed micelles were found to preserve the cytoplasmic domain of YgaP protein, in contrary to observations made with a protein solubilized in a pure detergent solution (Tzitzilioni et al. 2013).

From earlier studies of mixture of lipids and detergents in aqueous solutions (e.g., Gabriel and Roberts 1984), binary assemblies of detergents and lipids named *bicelles*, which contain usually a higher proportion of lipids than in mixed micelles, have been well characterized (Sanders and Prestegard 1990; Sanders and Schwonek 1992; Vold et al. 1997). Under appropriate conditions of temperature and concentration, bicelles are classically described as a planar bilayer of phospholipid stabilized by a swimming belt of detergents or short-chain lipids. Depending on the molar ratio q of lipids versus detergents, two kinds of bicelles can be distinguished: large anisotropic ($q > 0.5$, *vide infra* Fig. 12.6) and small isotropic ($q \leq 0.5$) bicelles (Fig. 12.2a). Larger bicelles can be used for solid-state NMR studies of MPs (e.g., Triba et al. 2006a, *cf.* § 12.3.2. herein), while isotropic bicelles are used for solution-state NMR investigations (e.g., Czernski and Sanders 2000). MPs associated to isotropic bicelles give rise to longer overall correlation times, but still to detectable NMR signals (e.g., Lee et al. 2008, Fig. 12.2b).

12.2.2.2 Illustrations

Complete structural studies of MPs in isotropic bicelles can be performed in solution by NMR (e.g., Bocharov et al. 2007, 2008; Lau et al. 2008). In an elegant comparison study, it has been shown that isotropic bicelles ($q = 0.33$) stabilize the functional form of a small multidrug-resistance transporter (Smr) compared to mono detergent solutions (Poget et al. 2007). Importantly, the authors succeeded to set up an *in vitro* ligand binding assay for this transporter, demonstrating that beautiful high-resolution two-dimensional (2D) ^1H , ^{15}N correlation experiments obtained in various pure detergent solutions do not necessarily mean the protein is active (Poget and Girvin 2007). Despite broader NMR signals, the authors succeeded to assign signals of ^1H , ^{13}C , and ^{15}N nuclei of the protein backbone in bicelles (Poget et al. 2007, 2010, Fig. 12.2b).

12.2.2.3 Advantages

Among some advantages, compared to pure detergent solutions, the immediate environment experienced by an MP is closer to a lipid bilayer. However, just as some mixed micelles may display some degree of organization, conversely, in the case of small isotropic bicelles, the architecture may be similar to mixed micelles rather than the idealized view of a well-segregated assembly between long-chain lipids and detergents (Triba et al. 2005, 2006b; Beaugrand 2014).

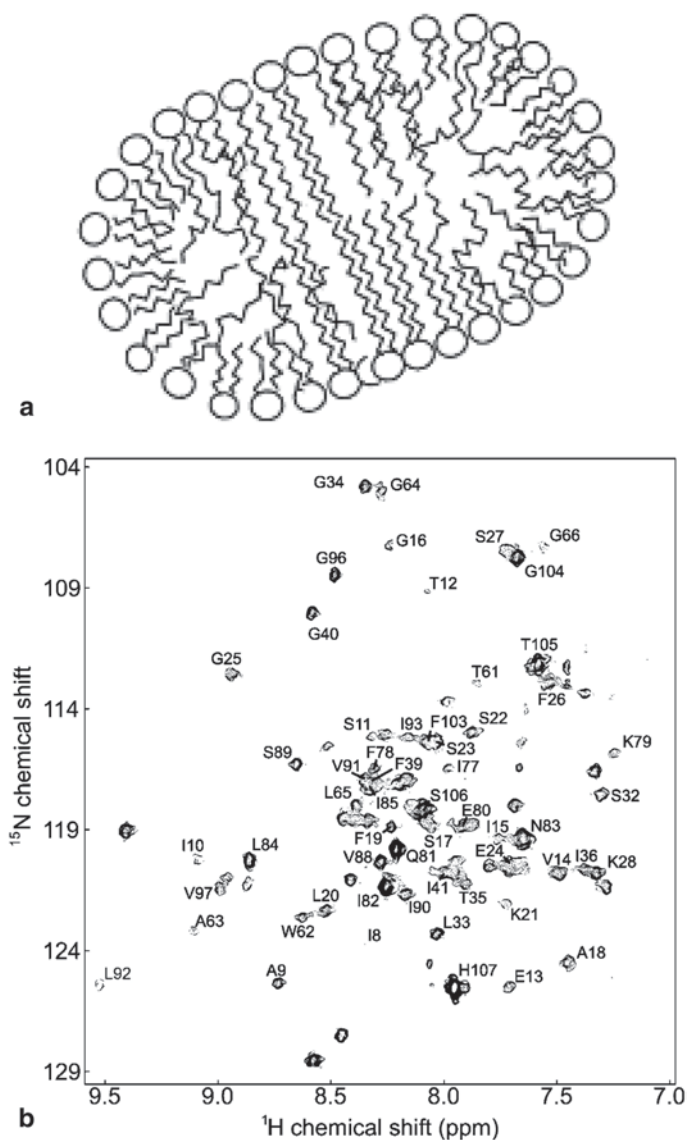


Fig. 12.2 NMR studies of MPs in mixed micelles or bicellar solutions. **a** The cross-section of an isotropic bicelle model, in which the disk-shaped bicelle consists of a small planar bilayer domain, predominately composed of long-chain phospholipids, coated by a rim of short-chain phospholipids or detergents (reprinted from Whiles et al. 2002 with permission from Elsevier). **b** Example of high-resolution NMR data of an MP associated to isotropic bicelles ($q=0.33$): ^1H , ^{15}N TROSY spectrum recorded at 900 MHz of 0.8 mM uniformly ^2H , ^{13}C , ^{15}N -labeled protein Smr (pH 6.5 and 47 °C). (Reprinted from Poget et al. 2007 with permission from the American Chemical Society)

12.2.2.4 Drawbacks

The detergent diffusion into the lipid disc (Triba et al. 2005, 2006b) may be a cause of protein instability. In addition, with isotropic bicelles, the current lipid compositions in use are limited, the best-characterized systems being composed of mixtures of DMPC and either dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC) or cholamidopropyl-dimethylammonio-hydroxy-propanesulfonate (CHAPSO) as detergents. Regarding sample preparations, the molar ratio q has to be kept constant in order to avoid any phase transition. This is not so trivial when buffer exchanges or temperature changes are required before reaching the NMR spectrometer or collecting NMR data.

12.2.3 Amphipols

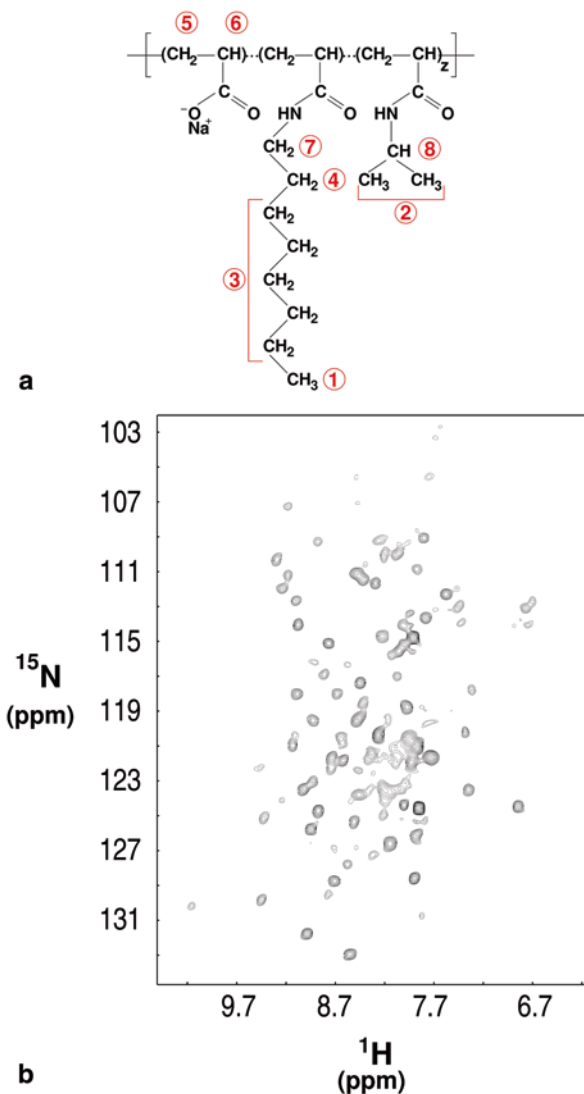
12.2.3.1 Generalities

The term *amphipol* (APol) refers to short amphiphilic polymers highly chemically stable, that can substitute for detergents to keep integral MPs water soluble (Tribet et al. 1996, Fig. 12.3a). These polymers were developed to find a relevant substitute to detergents by multiplying attachment points along the TM domain of MPs. They provide: (1) a low off-rate dissociation constant for these polymers from the TM domain, rendering their association to the MP quasi-irreversible and (2) a small equilibrium dissociation constant, which means low equilibrium concentration of free surfactant. For more than 15 years, these compounds have been proved to be a valuable alternative to detergents (Popot 2010; Popot et al. 2011; Elter et al. 2014 in preparation). In addition to their stabilizing properties, APols are compatible with many biophysical techniques, including NMR spectroscopy.

12.2.3.2 Illustrations

NMR studies of MP/APol complexes were carried out on several β -barrel MPs from the outer membrane of *E. coli* (Zoonens et al. 2005; Catoire et al. 2009, 2010a) or *Klebsiella pneumoniae* (Renault 2008; Planchard et al. 2014), and more recently with two α -helical MPs, the GPCR BLT2 (Catoire et al. 2010b, 2011) and the bacteriorhodopsin (Raschle et al. 2010; Etkorn et al. 2013). Even if no structure of MP associated to APols has been solved by NMR yet, all these studies clearly demonstrated that MP/APol complexes can give rise to highly resolved NMR signals in a short time: 2D heteronuclear $^1\text{H},^{15}\text{N}$ correlation experiments to attest β -barrels associated to APols are correctly folded (Zoonens et al. 2005; Catoire et al. 2010a; Raschle et al. 2010; Etkorn et al. 2013, Fig. 12.3b) or to look at slow dynamic chemical exchanges (Catoire et al. 2010a), 2D $^1\text{H},^{13}\text{C}$ heteronuclear nuclear

Fig. 12.3 Solution NMR studies of MPs trapped with APols. **a** Primary chemical structure of the polyacrylate APol A8-35 (Tribet et al. 1996; Popot et al. 2011). APol A8-35 can be easily partially deuterated (named DAPol, circled numbers 1, 2, 3, 4, 7, 8 indicating protons that are replaced by deuterons) or perdeuterated (named perDAPol). **b** Example of high-resolution NMR data of an MP associated to APols: $^1\text{H}, ^{15}\text{N}$ TROSY spectrum recorded at 800 MHz of uniformly $^2\text{H}, ^{15}\text{N}$ -labeled TM domain of OmpA ($[\text{OmpA}] = 1 \text{ mM}$, pH 7.9 and 30°C). (Reprinted from Zoonens et al. 2005 with permission from the National Academy of Sciences, USA)



Overhauser spectroscopy experiments (Catoire et al. 2009), or three-dimensional (3D) ^{15}N edited- $(^1\text{H}, ^1\text{H})$ HSQC-NOESY-TROSY experiments to look at the organization of APols around MPs (Renault 2008; Planchard et al. 2014). Structures of protonated organic ligands in their protein-bound states could also be determined with MPs associated to partially deuterated APols (Catoire et al. 2010b, 2011). All these studies are reviewed in Planchard et al. (2014).

12.2.3.3 Advantages

One of their major advantages over detergents is their stabilizing properties (e.g., Dahmane et al. 2009). This is particularly relevant in the context of NMR, which requires sometimes to work at high temperatures (typically 40–50 °C) during hours or days. Usually, MPs in APols come from a former stage where the protein is maintained soluble in a detergent solution. The oligomeric state of MPs in APols will depend whether or not the oligomerization has been conserved in detergents (see Planchard et al. 2014). Among some other advantages, these polymers have a very low critical aggregation concentration (equivalent to the cmc of detergents), which renders them irreversibly attached to MPs in the absence of competitive surfactants. This turns sample preparations and the handling of MPs associated to APols very easy. For instance, there is no need to add APols to the sample during buffer exchanges by dialysis. This has the practical advantage to limit the amount of APol consumed and to simplify sample preparations. Moreover, on an economical concern, APols are cost-effective compounds, which is quite interesting for NMR studies that usually require large amounts of material. These polymers can also be partially (Gohon et al. 2008) or totally deuterated (Giusti et al. 2014), which can greatly help to clearly identify protein signals in various homonuclear or heteronuclear NMR experiments.

12.2.3.4 Drawbacks

MP/APol complex sizes are usually larger than the same protein trapped with detergents (Popot 2010; Popot et al. 2011). But this increase in the overall correlation time does not preclude the observation of well-resolved NMR signals, thanks to the high magnetic fields available associated to relevant methodology and isotope-labeling schemes dedicated to the studies of large proteins or protein complexes. These broader NMR signals are largely compensated by an MP stable and active, which renders highly safe any MP/APol NMR studies. Perhaps, one of the major drawbacks concerning APols is psychological, as they do not resemble at all to a lipid bilayer. But, the only relevant answer to “how far can we move away from the native lipid membrane experienced by one MP or how artificial can be the swimming belt?” relies in the fact that the protein is active, i.e., correctly folded and stable. To be noticed, APols are an artificial medium that favor the retention of lipids, in contrary to detergents, thus providing an environment that is finally closer to a membrane than what could be told by their primary chemical structures. For NMR, the polyacrylate-based APol, historically named A8–35, is highly soluble at pH > 7. In acidic solutions, carboxylate groups start to protonate, leading to the progressive aggregation of the polymer. Working in basic solutions can be detrimental to observing exchangeable protons (Wüthrich 1986). Consequently, a bunch of different APols soluble in the 0–14 pH range are now available and have been validated for NMR (see Dahmane et al. 2011; Bazzacco et al. 2012).

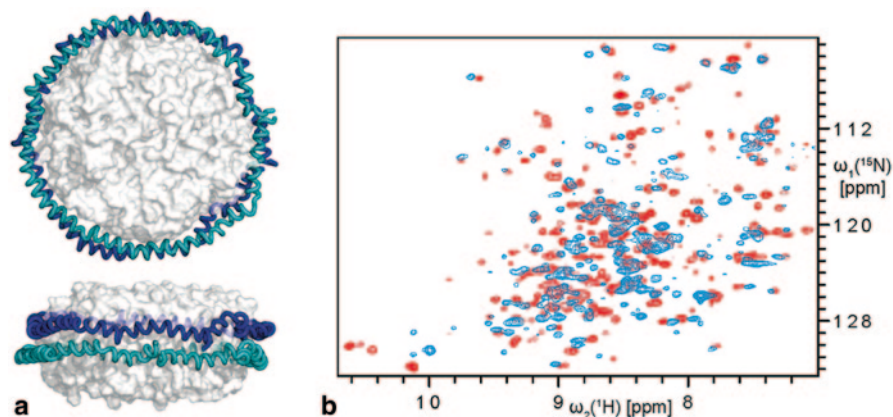


Fig. 12.4 Solution NMR studies of MPs embedded in nanodiscs. **a** Schematic view of nanodiscs, modeled with POPC as lipid. Lipid bilayer fragment (*white space filling*) is encircled by two amphipathic helices of membrane scaffold proteins (*blue ribbons*) (reprinted from Nath et al. 2007 with permission from the American Chemical Society). **b** 2D [^1H , ^{15}N]-TROSY spectrum of uniformly ^2H , ^{15}N -VDAC-1 in DMPC nanodiscs (*blue*) and in LDAO micelles (*red*). (Reprinted from Raschle et al. 2009 with permission from the American Chemical Society)

12.2.4 Nanometric Lipid Bilayers

12.2.4.1 Generalities

Nanometric lipid bilayers, often referred to as *nanodiscs*, have been designed to conduct *in vitro* biophysical studies of MPs (Bayburt et al. 2002; Denisov et al. 2004; Ritchie et al. 2009). A nanodisc is a non-covalent assembly of a phospholipid bilayer surrounded by a dimer of a genetically engineered lipoprotein named *membrane scaffold protein* (MSP; Fig. 12.4a). Various MSP have been engineered from the original sequence of human serum apolipoprotein apoA-I, which physiological role, through discoidal high-density lipoprotein particles, consists in reversing the transport of cholesterol (Ohashi et al. 2005). The size of the MSP defines the size of the nanodiscs and various lipids or mix of lipids that can be used to form the discoidal bilayer architecture (Ritchie et al. 2009).

12.2.4.2 Illustrations

A tremendous number of studies of MPs embedded in nanodiscs have been described in the literature (Nath et al. 2007; Ritchie et al. 2009; Bayburt and Sligar 2010). MPs associated to these discs are amenable to either solid-state (*vide infra* § 12.3.4) or solution-state NMR studies. Indeed, high-resolution NMR spectra can be obtained in solution (e.g., Tzitzilonis et al. 2013; Glück et al. 2009; Shenkarev

et al. 2010; Yokogawa et al. 2012; Shenkarev et al. 2013). Among these examples, the human mitochondrial voltage-dependent anion channel (VDAC-1), a 19-stranded β -barrel MP, in DMPC nanodiscs gave rise to a high-resolution 2D ^1H , ^{15}N -TROSY spectrum (Raschle et al. 2009). These NMR data were sufficiently different from those obtained with the protein in dodecyl-dimethylamine-oxide (LDAO or C12-DAO) micelles (Fig. 12.4b), requiring a novel sequence-specific resonance assignment of VDAC-1 in nanodiscs. More recently, the first structure of an MP embedded in a nanodisc, OmpX, has been determined by NMR in solution with the help of truncated MSP variants (Hagn et al. 2013). A comparison of the different OmpX structures obtained in different mimetics indicated differences in both the extracellular loops and the length and relative orientation of TM β -strands. Dynamics measurements also showed substantial differences, all the data underlying the impact of the artificial medium chosen on the structural and dynamical properties of MPs outside their native membranes.

12.2.4.3 Advantages

Nanodiscs offer a very convenient adjustable—in size and lipid composition—bilayer environment for *in vitro* studies of MPs. Like with APols, MPs associated to nanodiscs are highly stable, enabling the collection of NMR data at 50 °C during hours or days. Another advantage concerns studies of homo-oligomeric and hetero-oligomeric proteins. In most of the pre-discussed nonnative environments, it may sometimes be difficult to maintain the oligomeric state of proteins, especially in the framework of solution NMR studies that are conducted with high concentrations of surfactants. This is especially the case for systems involving the use of detergents. With nanodiscs, playing with the ratio of nanodiscs versus MP and the size of the nanodiscs at the reconstitution step, it is possible to adjust the oligomerization state of the protein (Ritchie et al. 2009).

12.2.4.4 Drawbacks

Even if desorption of lipids has been evaluated to be 20 times faster than in liposomes (Nakano et al. 2009), MPs embedded in nanodiscs at concentrations compatible with NMR studies in solution have been proven to be highly stable (Nath et al. 2007; Ritchie et al. 2009; Bayburt and Sligar 2010). One of the major drawback concerns sample preparations, even though once the MSP has been produced, the reconstitution procedure is quite straightforward and universal. MSP production and purification, along with production of the tobacco etch virus (TEV) protease, is time consuming, but the MSP can be kept stable frozen at -80°C for months. It may also be time consuming to find the appropriate lipid composition to get a fully active protein. Depending on the lipid composition used, sometimes it may be advantageous to work with deuterated lipids (e.g., Tzitzilonis et al. 2013). Unfortunately, many lipids are not available deuterated, even partially. Like with bicelles

and APols, the size of the MP/nanodisc complexes are much bigger than the corresponding size of the MPs in a detergent solution. This has led, for instance, to the construction of a new MSP variant leading to smaller nanodiscs (Hagn et al. 2013). Like with APols, broader NMR signals observed with nanodiscs can be compensated by working at higher temperatures.

12.3 Solid-State NMR



Courtesy of Edith Godard

12.3.1 A Little Bit of Theory

Solid-state NMR is the application of NMR to molecules that do not *tumble fast and isotropically*. This is due to the fact that, in such a case, the effects of *orientation-dependent couplings* often manifest themselves as severe line broadening on the NMR spectra. Depending on the molecule, the temperature, the nuclei that will be observed, the magnetic field, etc., several couplings will be dominating. Those couplings involve, as often in NMR, energy transfer that can be expressed in units of *frequency rate*. The difference between *fast* and *slow* tumbling is thereby defined by comparison to the strongest effective coupling present.

Without going into too many details, it is important to state that most orientation-dependent couplings bear a $(3\cos^2\theta - 1)$ angular dependence, where θ is the angle between the magnetic field and the line that connects the two coupled nuclei. Without such knowledge, one would be tempted to try to tumble the sample fast and

isotropically, which is not very practical. Knowing the simple $(3\cos^2\theta - 1)$ angular dependence allows to suggest ways to get rid of the broadening.

Since we cannot force θ for all pairs of atoms in the sample, and at all times, to be such that $(3\cos^2\theta - 1) = 0$, or $\theta \sim 55^\circ$ (also known as *the magic angle*), two options are offered:

- Aligning all molecules to a single orientation in the sample, so that instead of averaging out the couplings, they will have a single value throughout the sample, and the superimposition of resulting NMR spectra will be a narrow spectrum.
- Setting conditions to get the *average* value $(3\cos^2\theta - 1) = 0$, by placing the sample in a rotor macroscopically aligned at the magic angle, and spinning it fast, faster than the strongest effective coupling present.

In solution-state NMR, resonance lines broaden with protein size because they depend on the tumbling rate that averages $(3\cos^2\theta - 1)$ to zero when the molecular correlation time is around the nanosecond or faster. By contrast, in solid-state NMR, other approaches are at stake and linewidths are independent of the protein tumbling rate: therefore, theoretically, there is *no* protein size limit in solid-state NMR. Nevertheless, perfect alignment and infinite fast spinning rate are impossible to reach: Residual linewidths in solid-state NMR are broader than in solution-state NMR and ^1H – ^1H dipolar couplings that can reach 120 kHz are almost impossible to get rid of. In other words, biological solid-state NMR today is mostly applied to ^{13}C and ^{15}N NMR, in rotors of 2–7 mm diameter, 2–300- μl volume, and 5–50-kHz maximum spinning rate.

Combined with those mechanical approaches, another way to remove couplings between two nuclei is also to suppress one of the nuclei by replacing it with a *cold* isotope. Biochemistry has provided the NMR spectroscopist with a variety of isotopic labeling schemes, especially for the suppression of most large couplings in neighboring nuclei, leaving untouched either the smaller couplings among nuclei that are far away, or among selected isolated pairs of nuclei (Abdine et al. 2011). Typical examples involve growing proteins on selectively labeled glycerol (Castellani et al. 2002, Fig. 12.5) or glucose (Loquet et al. 2011).

12.3.2 Aligned Solid-State NMR: Glass Plates and Large Bicelles

12.3.2.1 Generalities

Aligning all biomolecules to a single orientation at the blow of a whistle may seem like a foolish dream, but in the case of lipids, we are approaching this possibility (Dürr et al. 2007a; Warschawski et al. 2011). First, they spontaneously align on a glass plate, with the bilayer normal perpendicular to the plate plane. This approach was used in the 1980s and helped determine the first MP structures, mostly by the

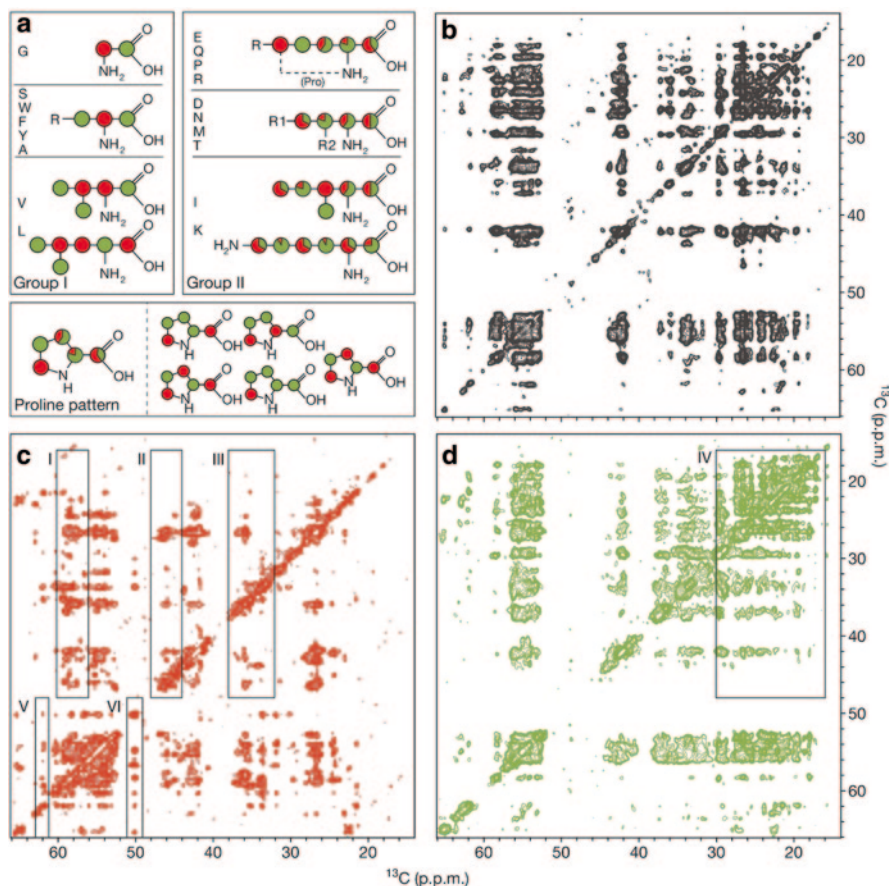


Fig. 12.5 Labeling patterns and NMR spectra for different protein preparations. Schematic representation of the effective ^{13}C enrichment for indicated amino acids by growth on $[1,3\text{-}^{13}\text{C}]$ glycerol (green) or $[2\text{-}^{13}\text{C}]$ glycerol (red) (a). 2D $^{13}\text{C}\text{-}^{13}\text{C}$ solid-state NMR spectra under magic-angle spinning at 8 kHz on α -spectrin SH3 domain grown on uniformly labeled glucose (b), $[2\text{-}^{13}\text{C}]$ glycerol (c), or $[1,3\text{-}^{13}\text{C}]$ glycerol (d). (Reprinted from Castellani et al. 2002 with permission from Macmillan)

groups of Tim Cross and Stan Opella. Second, combined with short-chain lipids, long-chain lipids may form bicelles that also spontaneously align in the magnetic field. Bicelles were discovered in the 1990s (Sanders and Schwonek 1992) and are shaped as 500-Å diameter *camemberts* or *wheels*, composed of a planar lipid bilayer, generally made of around 6,000 DMPC molecules, surrounded by *tires* of around 2,000 short-chain lipids, usually DHPC. Depending on the ratio between both lipids, as well as temperature, bicelles can also be too small to align in a standard NMR magnetic field. In such a case (for example, 50 DMPC, 200 DHPC, and a 100-Å diameter), they are called *isotropic bicelles* and were described in an earlier section, for use in solution-state NMR (vide supra § 12.2.2. herein).

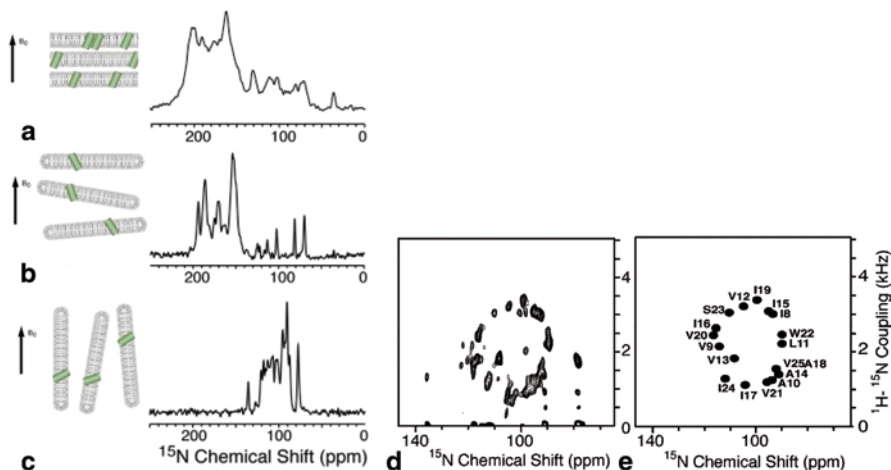


Fig. 12.6 Solid-state NMR spectra of the TM domain of Vpu in lipid bilayers differently aligned: on glass plates (a), in flipped bicelles (b), or in normal bicelles (c–e). a–c are 1D ^{15}N NMR spectra while d is a 2D ^{15}N - ^1H - ^{15}N NMR spectrum allowing for resonance assignment (simulated and shown in (e)) and helix tilt determination (determined here to be approximately 30° with respect to the bilayer normal). (Adapted from De Angelis et al. 2004 with permission from the American Chemical Society. The cartoons on the left are adapted from Dürr et al. 2007a with permission from Elsevier)

12.3.2.2 Illustrations

Recent progress in magic-angle spinning NMR has confined aligned solid-state NMR to mostly low-resolution structure determination, such as peptide orientation in the membrane, either on glass plates (Gong et al. 2004; Michalek et al. 2013) or bicelles (De Angelis et al. 2004; Triba et al. 2006a; Dürr et al. 2007b; Müller et al. 2007; Park et al. 2011a; Cook et al. 2011). Several high-resolution structures of small proteins were determined as well (Ketchem et al. 1993; Opella et al. 1999; Park et al. 2003; De Angelis et al. 2006; Hu et al. 2007; Traaseth et al. 2009; Ahuja et al. 2013), sometimes with the help of other complementary techniques such as solution-state NMR or X-ray crystallography (Warschawski 2013). Figure 12.6 shows typical NMR spectra of proteins reconstituted in lipids on glass plates, *normal* bicelles, or *flipped* bicelles (supplemented with small amount of lanthanides), of which the helix tilt is deduced.

12.3.2.3 Advantages and Drawbacks

Aligned solid-state NMR is more useful for the determination of peptide orientation in the membrane than for *ab initio* MP complete structure determination. Lipids on glass plates can be aligned almost regardless of lipid composition and temperature, at any given orientation in the magnet. Large bicelles can only align with their

bilayer normal perpendicular or parallel to the magnetic field (*flipped* bicelles), and only with special lipid composition and in a specific temperature range. Nevertheless, bicelle samples are better hydrated, easier to prepare, and more convenient in case one needs to change the buffer. In addition, as can be seen on the spectral resolution in Fig. 12.6, the average lipid alignment is better in bicelles than on glass plates.

12.3.3 Magic-Angle Spinning: Liposomes

12.3.3.1 Generalities

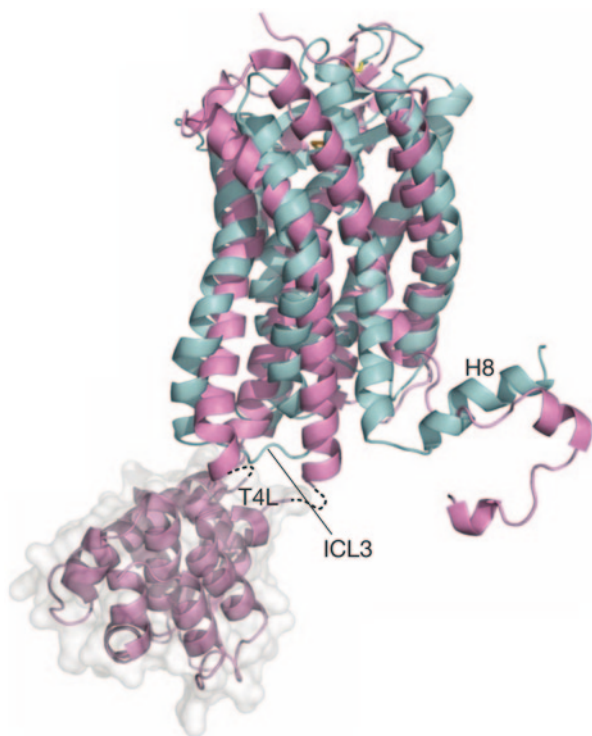
Liposomes are mostly composed of lipids that can be chosen from an incredible variety of charges, headgroups, chain lengths, or insaturations. Spontaneously, lipids and water form heterogeneous multilamellar vesicles of around 1- μm diameter and up to a dozen bilayers. Several MP reconstitution methods are available, usually, but not necessarily, with the help of detergent molecules.

12.3.3.2 Illustrations

Solid-state NMR of MPs in liposomes is almost exclusively performed under magic-angle spinning. Residual linewidths have hampered the structure determination of many MPs, but assignment is on the way for various proteins (Andronesi et al. 2005; Hiller et al. 2005; Lange et al. 2006; Frericks et al. 2006; Etzkorn et al. 2007; Shi et al. 2009; Abdine et al. 2010; Emami et al. 2013). Dynamic information can also be obtained (Ullrich et al. 2011; Yang et al. 2011; Williams et al. 2013). Recent structures include the small protegrin (Mani et al. 2006), influenza M2 channel (Cady et al. 2010), membrane domain of Mer F (Lu et al. 2013), as well as the large seven-helix G protein-coupled chemokine receptor CXCR1 (Park et al. 2012).

The structure of CXCR1 determined by solid-state NMR shares significant similarities with that of CXCR4 determined by X-ray crystallography (Park et al. 2012, Fig. 12.7). Differences mostly reflect the modifications made to the sequence of CXCR4 required for crystallization: In contrast with the NMR sample made of wild-type protein embedded in a liquid crystalline phospholipid bilayer, the crystal is made of a mutant protein, mostly by replacing the third intracellular loop (ICL3) by T4 lysosyme, and by removing the last 33 C-terminal residues. Removing ICL3 rendered CXCR4 incapable of activating G proteins, while CXCR1 in the NMR sample was fully active with respect to both G protein activation and chemokine binding. In addition, the C-terminus of CXCR1 forms a well-defined helix (H8) that aligns along the membrane surface. This helix, as well as the membrane, is absent from the crystal of the mutated CXCR4.

Fig. 12.7 Backbone structural comparison of CXCR1 determined by solid-state NMR (PDB accession 2LNL, in *cyan*) and CXCR4 determined by X-ray crystallography (PDB accession 3ODU, in *pink*). The third intracellular loop (ICL3) of CXCR4 is replaced by T4 lysozyme (T4 L) present in the crystal. The C-terminus of CXCR1 forms a well-defined amphipathic helix (H8), whereas that of CXCR4 is only loosely helical. (Reprinted from Park et al. 2012 by permission from Macmillan)



12.3.3.3 Advantages and Drawbacks

Magic-angle spinning NMR has shown its ability to determine high-resolution 3D structures of proteins. Liposomes are supposed to be the most natural local environment (in terms of size, shape, curvature, thickness, fluidity, lateral pressure, dielectric constant, hydration...) for MPs, where their structure and dynamics are supposed to be native-like, and where they can be studied in a functional state (Warschawski et al. 2011; Park et al. 2012). Lipids, which are the major constituent in the sample, can be chosen to suit the protein, if necessary. On the other hand, liposomes are heterogeneous multilamellar vesicles, where MPs may also experience heterogeneous conformations, slight differences between monomers in an oligomer, or slow and fast motion that, altogether, may broaden the NMR lines and affect the spectral resolution.

12.3.3.4 Future Perspective

Dynamic nuclear polarization (DNP) is a technique that can be combined with solid-state NMR under magic-angle spinning for signal enhancement of up to 120 on frozen samples. It has successfully been applied to MPs in liposomes, such as the

SecYEG translocon, allowing to detect 40 nmol of peptide bound to the translocon (Reggie et al. 2011) or to the M2 proton transporter of Influenza A, allowing the precise positioning of rimantadine bound to the protein (Andreas et al. 2013). Such a progress will undoubtedly allow new applications for solid-state NMR, including structure determination of MPs in complex environment in a near future.

12.3.4 Nanodiscs and Other New Environments for Solid-State NMR

12.3.4.1 Nanodiscs

The smallest nanolipoproteins described in an earlier section of this chapter (*vide supra* § 12.2.4. herein) have a diameter of 100 Å, comparable to small isotropic bicelles. As opposed to bicelles, those nanodiscs, or larger discs dubbed *macrodiscs* of up to 300-Å diameter (Park et al. 2011b), could be precipitated for solid-state NMR, at any given temperature. Since they are monodisperse, proteins in precipitated nanodiscs are expected to be more homogeneous than in liposomes of various sizes and lamellarities. Such samples should therefore provide narrower NMR lines under magic-angle spinning. In some cases, nanodisc samples could also accommodate a higher protein-to-lipid ratio, thus providing more intense NMR lines. Nevertheless, few examples of solid-state NMR studies of MPs have been shown to use nanodiscs so far (Kijac et al. 2007; Mörs et al. 2013).

12.3.4.2 Other Options

In the same line of thought, MPs can sometimes be (micro)crystallized to provide homogeneous, monodisperse samples that are known to give high-resolution solid-state NMR spectra (Castellani et al. 2002). Several MPs have followed this trend usually limited to soluble or fibrillar proteins: Structures of the small HNP1 defensin (Zhang et al. 2010), large complexes of DsbB (Tang et al. 2011, 2013), and the anchor domain of YadA (Shahid et al. 2012) have been determined by solid-state NMR using this procedure (Warschawski 2013). Other less common membrane mimetics for solid-state NMR, including APols, have also been described in a recent review (Warschawski et al. 2011).

12.3.5 Intact or Fragmented Cells Studied by NMR

12.3.5.1 Generalities

Cellular structural biology has known tremendous advances along the 2000s decade. Among the few techniques that can give rise to structural information at the

atomic level, NMR spectroscopy represents today a unique opportunity to work with intact or slightly modified biological samples. Solution-state NMR can give rise to highly resolved NMR signals of proteins in living cells (e.g., Inomata et al. 2009), opening the possibility to determine a complete protein structure *de novo* (Sakakibara et al. 2009). With MPs, the use of solution-state NMR techniques is essentially limited to observing ligand–protein interactions (e.g., Claasen et al. 2005; Assadi-Porter et al. 2008; Potenza et al. 2011), solid-state NMR being by far the most appropriate approach. Since solid-state NMR resolution is not affected by the large size or slow motion of the macromolecule observed, MPs can be studied in large liposomes or even the intact cells where they were produced or fragmented cellular membranes. Not only would the proteins be in their native environment (both the membrane and the soluble partners) but they will also be preserved from any potentially denaturing reconstitution protocol. Here, the main difficulty is to be able to isolate signals coming from the protein of interest in the forest of signals from all other molecules in the sample. ^{31}P solid-state NMR of intact cells may also be a method of choice to follow the kinetics of membrane assembly, the production of lipids, and energetics parameters such as the adenosine triphosphate/adenosine diphosphate (ATP/ADP) ratio.

12.3.5.2 Illustrations

Recent studies have tackled the proteasome and other biomolecules (MPs, polysaccharides, carotenoids, lipids, etc.) of several organisms such as *A. thaliana* (Dick-Pérez et al. 2011), *E. coli* (Renault et al. 2012a; Tardy-Laporte et al. 2013), *C. tepidum* (Kulminskaya et al. 2012), or *S. enterica* (Zandomeneghi et al. 2012). Others use bacterial cells to study individual proteins of interest (over-)expressed in *E. coli*, such as the human LR11 (Fu et al. 2011), the human FK506-binding protein (Reckel et al. 2012), the M2 from influenza virus (Miao et al. 2012), the diacylglycerol kinase (Shi et al. 2012), or the b subunit of the F1Fo ATP synthase (our current work).

Our studies of the b subunit are taking advantage of a mutant of *E. coli* called C43(λ DE3) (Miroux and Walker 1996; Arechaga et al. 2000), known for overproducing *internal membranes* when this protein is over-expressed. Focusing on these internal membranes rather than on whole cells, and using various tricks (low or high temperatures, different pulse schemes, sample preparation, etc.), we manage to reduce background signals (such as those coming from lipids) and discriminate signals arising from proteins. We validate this approach by comparing the spectrum of biological samples with that of model membranes with the same lipid composition as these internal membranes (see Fig. 12.8).

12.3.5.3 Future Perspective

DNP is particularly suited for the detection of small amount of active MPs in intact cells. It is not clear if structure determination is reachable, but attempts have already

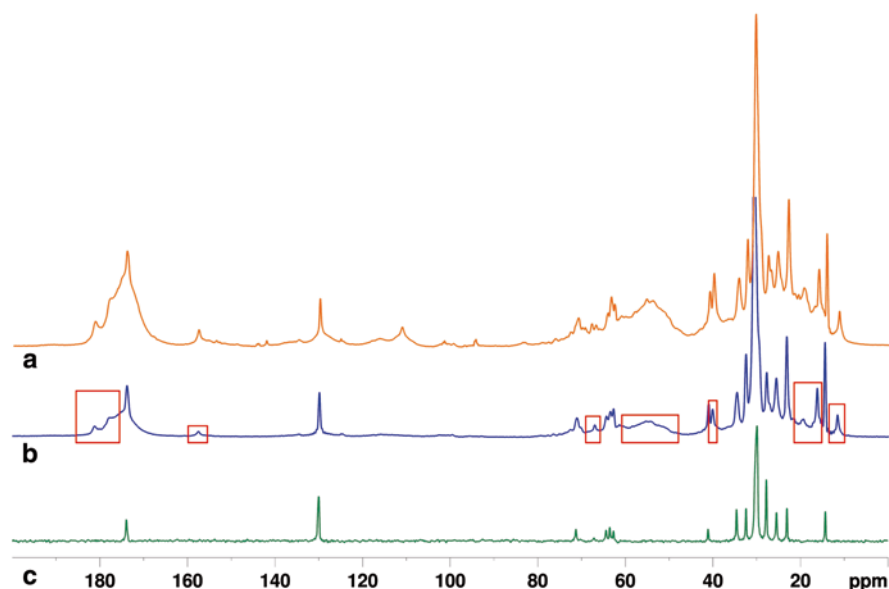


Fig. 12.8 NOE-enhanced ^{13}C solid-state NMR spectra recorded at 700 MHz of samples under magic-angle spinning at 11 kHz. **a** Whole C43(λ DE3) cells over-expressing the b subunit. **b** Purified internal membranes. **c** Model lipid membranes. Protein signals are highlighted by *red squares*

included bacteriorhodopsin (Bajaj et al. 2009), the acetylcholine receptor (Linden et al. 2011), mistic (Jacso et al. 2012), the whole proteasome of *E. coli* (Renault et al. 2012b), or the cell walls of *B. subtilis* (Takahashi et al. 2013).

12.4 Conclusion

This chapter is like a *hitch-hiker guide to the study of membrane proteins by NMR* in the jungle of surfactants and lipids for a clear overview on the best possible membrane mimetics. Detergents are a world of their own, which fascinates well beyond NMR: What would the biochemist do without detergents to solubilize, purify, transfer, renature, reconstitute, manipulate, and basically study MPs? If that was not enough, fellow scientists have gone out of their way to invent new molecules such as amphiphathic polymers and nanoparticles of lipids and proteins, or new techniques such as magic-angle spinning. In the country of lipids, we have described artificial membranes of various sizes and lamellarity, and the quasi black magic that lets lipid self-align, either on glass plates or in the intriguing *bicelles*. Eventually, *native-like* membrane mimetics were replaced by real *native* cell membranes, the graal of MP structural biology. Two routes were suggested: either solid- or solution-state NMR, but always through the arduous path of biochemistry. With such a guide, the

biochemist can take his/her favorite MP on an NMR journey, and hopefully come back home with a picture of its 3D structure.

All those paths have shown to lead to structures, but they have also shown rough passes that require technical improvement. While it is not clear what new route will be opened, we can already suggest directions in which to look for:

- *Isotopic labeling.* In solution-state NMR, selective methyl labeling helps pushing the limit to studying protein–surfactant complexes beyond 100 kDa. In solid-state NMR, sparse labeling is necessary to improve the spectral resolution.
- *Ligand binding.* Both in solution- and solid-state NMR, focusing on the ligand in a protein–ligand complex is both easier and often a niche that is inaccessible to X-ray crystallography. In addition, it can freeze a single protein conformation that would improve the resolution in solid-state NMR.
- *Combination of various techniques.* Solution- and solid-state NMR, X-ray crystallography, small-angle X-ray scattering, electron, atomic-force or light microscopy, structural mass spectrometry, and molecular modeling should be combined in order to close the resolution scale gap between structural and cellular biology, including dynamic processes, of MPs in ever increasingly complex environments, and get a more complete picture of these proteins.

One exciting new development highlighted in our guide is *in-cell* NMR. Sensitivity and resolution issues impose long experiments under conditions that are unfavorable for cell survival. Thereby, in most experiments that we have described here, the cell membranes and proteins were untouched but the cells were most likely dead. With little technical improvement already tested by our colleagues in metabolomics (Gowda et al. 2012; Jordà et al. 2013), one can hope to develop efficient *in vivo* structural and dynamical NMR in the near future.

Acknowledgments This work was supported by the CNRS (UMR 7099), the Université Paris Diderot, the Labex Dynamo (ANR-11-LABX-0011-01), and a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche (to XLW). We thank Eric Guittet and Christina Sizun for help with the NMR, Oana Illoaia for advice in microbiology, and Jean-Luc Popot for proof-reading the manuscript. Bienvenue à Bérénice.

References

- Abdine A, Verhoeven MA, Park KH, Ghazi A, Guittet E, Berrier C, Van Heijenoort C, Warschawski DE (2010) Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J Magn Reson* 204:155–159. doi:10.1016/j.jmr.2010.02.003
- Abdine A, Verhoeven MA, Warschawski DE (2011) Cell-free expression and labeling strategies for a new decade in solid-state NMR. *N Biotechnol* 28:272–276. doi:10.1016/j.nbt.2010.07.014
- Ahuja S, Jahr N, Im SC, Vivekanandan S, Popovych N, Le Clair SV, Huang R, Soong R, Xu J, Yamamoto K, Nanga RP, Bridges A, Waskell L, Ramamoorthy A (2013) A model of the membrane-bound cytochrome b5-cytochrome P450 complex from NMR and mutagenesis data. *J Biol Chem* 288:22080–22095. doi:10.1074/jbc.M112.448225

- Andreas LB, Barnes AB, Corzilius B, Chou JJ, Miller EA, Caporini M, Rosay M, Griffin RG (2013) Dynamic nuclear polarization study of inhibitor binding to the M2(18–60) proton transporter from influenza A. *Biochemistry* 52:2774–2282. doi:10.1021/bi400150x
- Andronesi OC, Becker S, Seidel K, Heise H, Young HS, Baldus M (2005) Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. *J Am Chem Soc* 127:12965–12974
- Arechaga I, Miroux B, Karrasch S, Huijbregts R, de Kruijff B, Runswick MJ, Walker JE (2000) Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase. *FEBS Lett* 482:215–219
- Arora A, Abildgaard F, Bushweller JH, Tamm LK (2001) Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nat Struct Biol* 8:334–338
- Assadi-Porter FM, Tonelli M, Maillet E, Hallenga K, Benard O, Max M, Markley JL (2008) Direct NMR detection of the binding of functional ligands to the human sweet receptor, a heterodimeric family 3 GPCR. *J Am Chem Soc* 130:7212–7213
- Badola P, Sanders CR (1997) *Escherichia coli* diacylglycerol kinase is an evolutionarily optimized membrane enzyme and catalyzes direct phosphoryl transfer. *J Biol Chem* 272:24176–24182
- Bajaj VS, Mak-Jurkauskas ML, Belenky M, Herzfeld J, Griffin RG (2009) Functional and shunt states of bacteriorhodopsin resolved by 250 GHz dynamic nuclear polarization-enhanced solid-state NMR. *Proc Natl Acad Sci U S A* 106:9244–9249. doi:10.1073/pnas.0900908106
- Bayburt TH, Grinkova YV, Sligar SG (2002) Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Lett* 2:853–856
- Bayburt TH, Sligar SG (2010) Membrane protein assembly into nanodiscs. *FEBS Lett* 584:1721–1727
- Bazzacco P, Billon-Denis E, Sharma KS, Catoire LJ, Mary S, Le Bon C, Point E, Banères JL, Durand G, Zito F, Pucci B, Popot JL (2012) Nonionic homopolymeric amphipols: application to membrane protein folding, cell-free synthesis, and solution nuclear magnetic resonance. *Biochemistry* 51:1416–1430
- Beaugrand M, Arnold AA, Hénin J, Warschawski DE, Williamson PTF, Marcotte I (2014) Lipid Concentration and Molar Ratio Boundaries for the Use of Isotropic Bicelles. *Langmuir*, in press
- Berardi MJ, Shih WM, Harrison SC, Chou JJ (2011) Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature* 476:109–113
- Bocharov EV, Pustovalova YE, Pavlov KV, Volynsky PE, Goncharuk MV, Ermolyuk YS, Karpunin DV, Schulga AA, Kirpichnikov MP, Efremov RG, Maslennikov IV, Arseniev AS (2007) Unique dimeric structure of BNip3 transmembrane domain suggests membrane permeabilization as a cell death trigger. *J Biol Chem* 282:16256–16266
- Bocharov EV, Mineev KS, Volynsky PE, Ermolyuk YS, Tkach EN, Sobol AG, Chupin VV, Kirpichnikov MP, Efremov RG, Arseniev AS (2008) Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *J Biol Chem* 283:6950–6956
- Bokoch MP, Zou Y, Rasmussen SG, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi HJ, Thian FS, Kobilka TS, Puglisi JD, Weis WI, Pardo L, Prosser RS, Mueller L, Kobilka BK (2010) Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature* 463:108–112
- Bowie JU (2001) Stabilizing membrane proteins. *Curr Opin Struct Biol* 11:397–340
- Cady SD, Schmidt-Rohr K, Wang J, Soto CS, Degrado WF, Hong M (2010) Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* 463:689–692. doi:10.1038/nature08722
- Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420:98–102
- Catoire LJ, Zoonens M, van Heijenoort C, Giusti F, Popot JL, Guittet E (2009) Inter- and intramolecular contacts in a membrane protein/surfactant complex observed by heteronuclear dipole-to-dipole cross-relaxation. *J Magn Reson* 197:91–95

- Catoire LJ, Zoonens M, van Heijenoort C, Giusti F, Guittet E, Popot JL (2010a) Solution NMR mapping of water-accessible residues in the transmembrane beta-barrel of OmpX. *Eur Biophys J* 39:623–630
- Catoire LJ, Damian M, Giusti F, Martin A, van Heijenoort C, Popot JL, Guittet E, Banères JL (2010b) Structure of a GPCR ligand in its receptor-bound state: leukotriene B4 adopts a highly constrained conformation when associated to human BLT2. *J Am Chem Soc* 132:9049–9057
- Catoire LJ, Damian M, Baaden M, Guittet E, Banères JL (2011) Electrostatically-driven fast association and perdeuteration allow detection of transferred cross-relaxation for G protein-coupled receptor ligands with equilibrium dissociation constants in the high-to- low nanomolar range. *J Biomol NMR* 50:191–195
- Chae PS, Rasmussen SG, Rana RR, Gotfryd K, Chandra R, Goren MA, Kruse AC, Nurva S, Lolland CJ, Pierre Y, Drew D, Popot JL, Picot D, Fox BG, Guan L, Gether U, Byrne B, Kobilka B, Gellman SH (2010) Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* 7:1003–1008
- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* 318:1258–1265
- Chung KY, Kim TH, Manglik A, Alvares R, Kobilka BK, Prosser RS (2012) Role of detergents in conformational exchange of a G protein-coupled receptor. *J Biol Chem* 287:36305–36311
- Claassen B, Axmann M, Meinecke R, Meyer B (2005) Direct observation of ligand binding to membrane proteins in living cells by a saturation transfer double difference (STDD) NMR spectroscopy method shows a significantly higher affinity of integrin alpha(IIB)beta3 in native platelets than in liposomes. *J Am Chem Soc* 127:916–919
- Cook GA, Zhang H, Park SH, Wang Y, Opella SJ (2011) Comparative NMR studies demonstrate profound differences between two viroporins: p7 of HCV and Vpu of HIV-1. *Biochim Biophys Acta* 1808:554–560. doi:10.1016/j.bbame.2010.08.005
- Czerski L, Sanders CR (2000) Functionality of a membrane protein in bicelles. *Anal Biochem* 284:327–333
- Dahmane T, Damian M, Mary S, Popot JL, Banères JL (2009) Amphipol-assisted in vitro folding of G protein-coupled receptors. *Biochemistry* 48:6516–6521
- Dahmane T, Giusti F, Catoire LJ, Popot JL (2011) Sulfonated amphipols: synthesis, properties, and applications. *Biopolymers* 95:811–823
- De Angelis AA, Nevzorov AA, Park SH, Howell SC, Mrse AA, Opella SJ (2004) High-resolution NMR spectroscopy of membrane proteins in aligned bicelles. *J Am Chem Soc* 126:15340–15341
- De Angelis AA, Howell SC, Nevzorov AA, Opella SJ (2006) Structure determination of a membrane protein with two trans-membrane helices in aligned phospholipid bicelles by solid-state NMR spectroscopy. *J Am Chem Soc* 128:12256–12267
- Denisov IG, Grinkova YV, Lazarides AA, Sligar SG (2004) Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *J Am Chem Soc* 126:3477–3478
- Dick-Pérez M, Zhang Y, Hayes J, Salazar A, Zabolina OA, Hong M (2011) Structure and interactions of plant cell-wall polysaccharides by two- and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry* 50:989–1000. doi:10.1021/bi101795q
- Dürr UH, Waskell L, Ramamoorthy A (2007a) The cytochromes P450 and b5 and their reductases-promising targets for structural studies by advanced solid-state NMR spectroscopy. *Biochim Biophys Acta* 1768:3235–3259
- Dürr UH, Yamamoto K, Im SC, Waskell L, Ramamoorthy A (2007b) Solid-state NMR reveals structural and dynamical properties of a membrane-anchored electron-carrier protein, cytochrome b5. *J Am Chem Soc* 129:6670–6671
- Elter S, Raschle T, Arens S, Gelev V, Eitzkorn M, Wagner G (2014) The use of amphipols for NMR structural characterization of 7-TM proteins. *J Membr Biol*, in press
- Emami S, Fan Y, Munro R, Ladizhansky V, Brown LS (2013) Yeast-expressed human membrane protein aquaporin-1 yields excellent resolution of solid-state MAS NMR spectra. *J Biomol NMR* 55:147–155. doi:10.1007/s10858-013-9710-5

- Etzkorn M, Martell S, Andronesi OC, Seidel K, Engelhard M, Baldus M (2007) Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angew Chem Int Ed Engl* 46:459–462
- Etzkorn M, Raschle T, Hagn F, Gelev V, Rice AJ, Walz T, Wagner G (2013) Cell-free expressed bacteriorhodopsin in different soluble membrane mimetics: biophysical properties and NMR accessibility. *Structure* 21:394–401
- Fernández C, Adeishvili K, Wüthrich K (2001) Transverse relaxation-optimized NMR spectroscopy with the outer membrane protein OmpX in dihexanoyl phosphatidylcholine micelles. *Proc Natl Acad Sci U S A* 98:2358–2363
- Frericks HL, Zhou DH, Yap LL, Gennis RB, Rienstra CM (2006) Magic-angle spinning solid-state NMR of a 144 kDa membrane protein complex: *E. coli* cytochrome bo3 oxidase. *J Biomol NMR* 36:55–71
- Fu R, Wang X, Li C, Santiago-Miranda AN, Pielak GJ, Tian F (2011) In situ structural characterization of a recombinant protein in native *Escherichia coli* membranes with solid-state magic-angle-spinning NMR. *J Am Chem Soc* 133:12370–12373. doi:10.1021/ja204062v
- Gabriel NE, Roberts MF (1984) Spontaneous formation of stable unilamellar vesicles. *Biochemistry* 23:4011–4015
- Gautier A, Mott HR, Bostock MJ, Kirkpatrick JP, Nietlispach D (2010) Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nat Struct Mol Biol* 17:768–774
- Giusti F, Rieger J, Catoire LJ, Qian S, Calabrese AN, Watkinson TG, Casiraghi M, Radford SE, Ashcroft AE, Popot JL (2014) Synthesis, characterization and applications of a perdeuterated amphipol. *J Membr Biol*, in press
- Glück JM, Wittlich M, Feuerstein S, Hoffmann S, Willbold D, Koenig BW (2009) Integral membrane proteins in nanodiscs can be studied by solution NMR spectroscopy. *J Am Chem Soc* 131:12060–12061
- Gohon Y, Dahmane T, Ruigrok R, Schuck P, Charvolin D, Rappaport F, Timmins P, Engelman DM, Tribet C, Popot JL, Ebel C (2008) Bacteriorhodopsin/amphipol complexes: structural and functional properties. *Biophys J* 94:3523–3537
- Gong XM, Choi J, Franzin CM, Zhai D, Reed JC, Marassi FM (2004) Conformation of membrane-associated proapoptotic tBid. *J Biol Chem* 279:28954–28960
- Gowda GA, Shanaiah N, Raftery D (2012) Isotope enhanced approaches in metabolomics. *Adv Exp Med Biol* 992:147–164. doi:10.1007/978-94-007-4954-2_8
- Hagn F, Etzkorn M, Raschle T, Wagner G (2013) Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *J Am Chem Soc* 135:1919–1925
- Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola VP, Chien EY, Velasquez J, Kuhn P, Stevens RC (2008) A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor. *Structure* 16:897–905
- Helenius A, Simons K (1975) Solubilization of membranes by detergents. *Biochim Biophys Acta* 415:29–79
- Hiller M, Krabben L, Vinothkumar KR, Castellani F, van Rossum BJ, Kühlbrandt W, Oschkinat H (2005) Solid-state magic-angle spinning NMR of outer membrane protein G from *Escherichia coli*. *Chembiochem* 6:1679–1684
- Hiller S, Garces RG, Malia TJ, Orekhov VY, Colombini M, Wagner G (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science* 321:1206–1210
- Hu J, Asbury T, Achuthan S, Li C, Bertram R, Quine JR, Fu R, Cross TA (2007) Backbone structure of the amantadine-blocked trans-membrane domain M2 proton channel from influenza A virus. *Biophys J* 92:4335–4343
- Hwang PM, Choy WY, Lo EI, Chen L, Forman-Kay JD, Raetz CR, Privé GG, Bishop RE, Kay LE (2002) Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc Natl Acad Sci U S A* 99:13560–13565

- Inomata K, Ohno A, Tochio H, Isogai S, Tenno T, Nakase I, Takeuchi T, Futaki S, Ito Y, Hiroaki H, Shirakawa M (2009) High-resolution multi dimensional NMR spectroscopy of proteins in human cells. *Nature* 458:106–109
- Ito Y, Selenko P (2010) Cellular structural biology. *Curr Opin Struct Biol* 20:640–648
- Jasco T, Franks WT, Rose H, Fink U, Broecker J, Keller S, Oschkinat H, Reif B (2012) Characterization of membrane proteins in isolated native cellular membranes by dynamic nuclear polarization solid-state NMR spectroscopy without purification and reconstitution. *Angew Chem Int Ed Engl* 51:432–435. doi:10.1002/anie.201104987
- Jordà J, Suarez C, Carnicer M, ten Pierick A, Heijnen JJ, van Gulik W, Ferrer P, Albiol J, Wahl A (2013) Glucose-methanol co-utilization in *Pichia pastoris* studied by metabolomics and instantaneous ¹³C flux analysis. *BMC Syst Biol* 7:17. doi:10.1186/1752-0509-7-17
- Kang CB, Li Q (2011) Solution NMR study of integral membrane proteins. *Curr Opin Struct Biol* 15:560–569
- Ketchum RR, Hu W, Cross TA (1993) High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR. *Science* 261:1457–1460
- Kijac AZ, Li Y, Sliagar SG, Rienstra CM (2007) Magic-angle spinning solid-state NMR spectroscopy of nanodisc-embedded human CYP3A4. *Biochemistry* 46:13696–13703
- Kofuku Y, Ueda T, Okude J, Shiraiishi Y, Kondo K, Maeda M, Tsujishita H, Shimada I (2012) Efficacy of the β 2-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nat Commun* 3:1045
- Kulminkskaya NV, Pedersen M, Bjerring M, Underhaug J, Miller M, Frigaard NU, Nielsen JT, Nielsen NC (2012) In situ solid-state NMR spectroscopy of protein in heterogeneous membranes: the baseplate antenna complex of *Chlorobaculum tepidum*. *Angew Chem Int Ed Engl* 51:6891–6895. doi:10.1002/anie.201201160
- Lange A, Giller K, Hornig S, Martin-Eauclaire MF, Pongs O, Becker S, Baldus M (2006) Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature* 440:959–962
- Lau TL, Partridge AW, Ginsberg MH, Ulmer TS (2008) Structure of the integrin beta3 transmembrane segment in phospholipid bicelles and detergent micelles. *Biochemistry* 47:4008–4016
- Lau TL, Kim C, Ginsberg MH, Ulmer TS (2009) The structure of the integrin alphaIb-beta3 transmembrane complex explains integrin transmembrane signalling. *EMBO J* 28:1351–1361
- Lee AG (2004) How lipids affect the activities of integral membrane proteins. *Biochim Biophys Acta* 1666:62–87
- Lee D, Walter KF, Brckner AK, Hilty C, Becker S, Griesinger C (2008) Bilayer in small bicelles revealed by lipid-protein interactions using NMR spectroscopy. *J Am Chem Soc* 130:13822–13823
- Linden AH, Lange S, Franks WT, Akbey U, Specker E, van Rossum BJ, Oschkinat H (2011) Neurotoxin II bound to acetylcholine receptors in native membranes studied by dynamic nuclear polarization NMR. *J Am Chem Soc* 133:19266–19269. doi:10.1021/ja206999c
- Liu JJ, Horst R, Katritch V, Stevens RC, Wüthrich K (2012) Biased signaling pathways in β 2-adrenergic receptor characterized by 19F-NMR. *Science* 335:1106–1110
- Loquet A, Lv G, Giller K, Becker S, Lange A (2011) ¹³C spin dilution for simplified and complete solid-state NMR resonance assignment of insoluble biological assemblies. *J Am Chem Soc* 133:4722–4725. doi:10.1021/ja200066s
- Lu GJ, Tian Y, Vora N, Marassi FM, Opella SJ (2013) The structure of the mercury transporter MerF in phospholipid bilayers: a large conformational rearrangement results from N-terminal truncation. *J Am Chem Soc* 135:9299–9302. doi:10.1021/ja4042115
- MacKenzie KR, Prestegard JH, Engelman DM (1997) A transmembrane helix dimer: structure and implications. *Science* 276:131–133
- Mani R, Tang M, Wu X, Buffy JJ, Waring AJ, Sherman MA, Hong M (2006) Membrane-bound dimer structure of a β -hairpin antimicrobial peptide from rotational-echo double-resonance solid-state NMR. *Biochemistry* 45:8341–8349
- McDonnell PA, Opella SJ (1993) Effect of detergent concentration on multidimensional solution NMR spectra of membrane proteins in micelles. *J Magn Reson B* 102:120–125

- McGregor CL, Chen L, Pomroy NC, Hwang P, Go S, Chakrabarty A, Priv GG (2003) Lipopeptide detergents designed for the structural study of membrane proteins. *Nat Biotechnol* 21:171–176
- Miao Y, Qin H, Fu R, Sharma M, Can TV, Hung I, Luca S, Gor'kov PL, Brey WW, Cross TA (2012) M2 proton channel structural validation from full-length protein samples in synthetic bilayers and *E. coli* membranes. *Angew Chem Int Ed Engl* 51:8383–8386. doi:10.1002/anie.201204666
- Michalek M, Salnikow ES, Werten S, Bechinger B (2013) Membrane interactions of the amphipathic amino terminus of huntingtin. *Biochemistry* 52:847–858. doi:10.1021/bi301325q
- Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260:289–298
- Mörs K, Roos C, Scholz F, Wachtveitl J, Dötsch V, Bernhard F, Glaubitc C (2013) Modified lipid and protein dynamics in nanodiscs. *Biochim Biophys Acta* 1828:1222–1229. doi:10.1016/j.bbame.2012.12.011
- Müller SD, De Angelis AA, Walther TH, Grage SL, Lange C, Opella SJ, Ulrich AS (2007) Structural characterization of the pore forming protein TatAd of the twin-arginine translocase in membranes by solid-state ¹⁵N-NMR. *Biochim Biophys Acta* 1768:3071–3079
- Nakano M, Fukuda M, Kudo T, Miyazaki M, Wada Y, Matsuzaki N, Endo H, Handa T (2009) Static and dynamic properties of phospholipid bilayer nanodiscs. *J Am Chem Soc* 131:8308–8312
- Nath A, Atkins WM, Sligar SG (2007) Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry* 46:2059–2069
- Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, Pan AC, Liu CW, Fung JJ, Bokoch MP, Thian FS, Kobilka TS, Shaw DE, Mueller L, Prosser RS, Kobilka BK (2013) The dynamic process of (2)-adrenergic receptor activation. *Cell* 152:532–542
- Ohashi R, Mu H, Wang X, Yao Q, Chen C (2005) Reverse cholesterol transport and cholesterol efflux in atherosclerosis. *Q J Med* 98:845–856
- Oliver RC, Lipfert J, Fox DA, Lo RH, Doniach S, Columbus L (2013) Dependence of micelle size and shape on detergent alkyl chain length and head group. *PLoS ONE* 8:e62488. doi:10.1371/journal.pone.0062488
- Opella SJ, Marassi FM, Gesell JJ, Valente AP, Kim Y, Oblatt-Montal M, Montal M (1999) Structures of the M2 channel-lining segments from nicotinic acetylcholine and NMDA receptors by NMR spectroscopy. *Nat Struct Biol* 6:374–379
- Park SH, Mrse AA, Nevzorov AA, Mesleh MF, Oblatt-Montal M, Montal M, Opella SJ (2003) Three-dimensional structure of the channel-forming trans-membrane domain of virus protein “u” (Vpu) from HIV-1. *J Mol Biol* 333:409–424
- Park SH, Casagrande F, Das BB, Albrecht L, Chu M, Opella SJ (2011a) Local and global dynamics of the G protein-coupled receptor CXCR1. *Biochemistry* 50:2371–2380. doi:10.1021/bi101568j
- Park SH, Berkamp S, Cook GA, Chan MK, Viadiu H, Opella SJ (2011b) Nanodiscs versus macrodiscs for NMR of membrane proteins. *Biochemistry* 50:8983–8985. doi:10.1021/bi201289c
- Park SH, Das BB, Casagrande F, Tian Y, Nothnagel HJ, Chu M, Kiefer H, Maier K, De Angelis AA, Marassi FM, Opella SJ (2012) Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature* 491:779–783. doi:10.1038/nature11580
- Planchard N, Point E, Dahmane T, Giusti F, Renault M, Le Bon C, Durand G, Milon A, Guittet E, Zoonens M, Popot JL, Catoire LJ (2014) The use of amphipols for solution NMR studies of membrane proteins: advantages and constraints as compared to other solubilizing media. *J Membr Biol*, in press
- Plevin MJ, Boisbouvier J (2012) Isotope-labelling of methyl groups for NMR studies of large proteins. In: Clore M, Potts J (eds) *Recent developments in biomolecular NMR*. Royal Society of Chemistry, London
- Poget SF, Girvin ME (2007) Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better. *Biochim Biophys Acta* 1768:3098–3106
- Poget SF, Cahill SM, Girvin ME (2007) Isotropic bicelles stabilize the functional form of a small multidrug-resistance pump for NMR structural studies. *J Am Chem Soc* 129:2432–2433

- Poget SF, Harris R, Cahill SM, Girvin ME (2010) ^1H , ^{13}C , ^{15}N backbone NMR assignments of the *Staphylococcus aureus* small multidrug-resistance pump (Smr) in a functionally active conformation. *Biomol NMR Assign* 4:139–142
- Popot JL (2010) Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Ann Rev Biochem* 79:737–775
- Popot JL, Althoff T, Bagnard D et al (2011) Amphipols from A to Z. *Ann Rev Biophys* 40:379–408
- Potenza D, Vasile F, Belvisi L, Civera M, Araldi EMV (2011) STD and trNOESY NMR study of receptor-ligand interactions in living cancer cells. *Chembiochem* 12:695–699
- Privé GG (2011) Lipopeptide detergents for membrane protein studies. *Curr Opin Struct Biol* 19:379–385
- Raschle T, Hiller S, Yu TY, Rice AJ, Walz T, Wagner G (2009) Structural and functional characterization of the integral membrane protein VDAC-1 in lipid bilayer nanodiscs. *J Am Chem Soc* 131:17777–17779
- Raschle T, Hiller S, Eitzkorn M, Wagner G (2010) Nonmicellar systems for solution NMR spectroscopy of membrane proteins. *Curr Opin Struct Biol* 20:471–479
- Reckel S, Lopez JJ, Löhr F, Glaubitz C, Dötsch V (2012) In-cell solid-state NMR as a tool to study proteins in large complexes. *Chembiochem* 13:534–537. doi:10.1002/cbic.201100721
- Reggie L, Lopez JJ, Collinson I, Glaubitz C, Lorch M (2011) Dynamic nuclear polarization-enhanced solid-state NMR of a ^{13}C -labeled signal peptide bound to lipid-reconstituted Sec translocon. *J Am Chem Soc* 133:19084–19086. doi:10.1021/ja209378h
- Renault M (2008) Etudes structurales et dynamiques de la protéine membranaire KpOmpA par RMN en phase liquide et solide. Dissertation, Université Paul Sabatier, Toulouse
- Renault M, Tommassen-van Boxtel R, Bos MP, Post JA, Tommassen J, Baldus M (2012a) Cellular solid-state nuclear magnetic resonance spectroscopy. *Proc Natl Acad Sci U S A* 109:4863–4868. doi:10.1073/pnas.1116478109
- Renault M, Pawsey S, Bos MP, Koers EJ, Nand D, Tommassen-van Boxtel R, Rosay M, Tommassen J, Maas WE, Baldus M (2012b) Solid-state NMR spectroscopy on cellular preparations enhanced by dynamic nuclear polarization. *Angew Chem Int Ed Engl* 51:2998–3001. doi:10.1002/anie.201105984; E coli
- Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, Sligar SG (2009) Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Method Enzymol* 464:211–213
- Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, Devree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK (2011) Structure and function of an irreversible agonist-(2) adrenoceptor complex. *Nature* 469:236–240
- Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Wälchli M, Smith BO, Shirakawa M, Güntert P, Ito Y (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* 458:102–105
- Sanders CR, Prestegard JH (1990) Magnetically orientable phospholipid bilayers containing small amounts of a bile salt analogue, CHAPSO. *Biophys J* 58:447–460
- Sanders CR, Schwonek JP (1992) Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR. *Biochemistry* 31:8898–8905
- Sanders CR, Landis GC (1995) Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies. *Biochemistry* 34:4030–4040
- Schnell JR, Chou JJ (2008) Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* 451:591–595
- Selenko P, Wagner G (2007) Looking into live cells with in-cell NMR spectroscopy. *J Struct Biol* 158:244–253
- Shahid SA, Bardiaux B, Franks WT, Krabben L, Habeck M, van Rossum BJ, Linke D (2012) Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. *Nat Methods* 9:1212–1217. doi:10.1038/nmeth.2248

- Shenkarev ZO, Lyukmanova EN, Paramonov AS, Shingarova LN, Chupin VV, Kirpichnikov MP, Blommers MJ, Arseniev AS (2010) Lipid-protein nanodiscs as reference medium in detergent screening for high-resolution NMR studies of integral membrane proteins. *J Am Chem Soc* 132:5628–5629
- Shenkarev ZO, Lyukmanova EN, Butenko IO, Petrovskaya LE, Paramonov AS, Shulepko MA, Nekrasova OV, Kirpichnikov MP, Arseniev AS (2013) Lipid-protein nanodiscs promote in vitro folding of transmembrane domains of multi-helical and multimeric membrane proteins. *Biochim Biophys Acta* 1828:776–784
- Shi L, Ahmed MA, Zhang W, Whited G, Brown LS, Ladizhansky V (2009) Three-dimensional solid-state NMR study of a seven-helical integral membrane proton pump-structural insights. *J Mol Biol* 386:1078–1093
- Shi P, Li D, Chen H, Xiong Y, Wang Y, Tian C (2012) In situ 19F NMR studies of an *E. coli* membrane protein. *Protein Sci* 21:596–600. doi:10.1002/pro.2040
- Takahashi H, Ayala I, Bardet M, De Paëpe G, Simorre JP, Hediger S (2013) Solid-State NMR on bacterial cells: selective cell wall signal enhancement and resolution improvement using dynamic nuclear polarization. *J Am Chem Soc* 135:5105–5110. doi:10.1021/ja312501d
- Tanford C, Reynolds JA (1976) Characterization of membrane proteins in detergent solutions. *Biochim Biophys Acta* 457:133–170
- Tang M, Sperling LJ, Berthold DA, Schwieters CD, Nesbitt AE, Nieuwkoop AJ, Gennis RB, Rienstra CM (2011) High-resolution membrane protein structure by joint calculations with solid-state NMR and X-ray experimental data. *J Biomol NMR* 51:227–233. doi:10.1007/s10858-011-9565-6
- Tang M, Nesbitt AE, Sperling LJ, Berthold DA, Schwieters CD, Gennis RB, Rienstra CM (2013) Structure of the disulfide bond generating membrane protein Dsbb in the lipid bilayer. *J Mol Biol* 425:1670–1682. doi:10.1016/j.jmb.2013.02.009
- Tardy-Laporte C, Arnold AA, Genard B, Gastineau R, Moranchais M, Mouget JL, Tremblay R, Marcotte I (2013) A (2)H solid-state NMR study of the effect of antimicrobial agents on intact *Escherichia coli* without mutating. *Biochim Biophys Acta* 1828:614–622. doi:10.1016/j.bbame.2012.09.011
- Teriete P, Franzin CM, Choi J, Marassi FM (2007) Structure of the Na, K-ATPase regulatory protein FXyD1 in micelles. *Biochemistry* 46:6774–6783
- Thompson AA, Liu JJ, Chun E, Wacker D, Wu H, Cherezov V, Stevens RC (2011) GPCR stabilization using the bicelle-like architecture of mixed sterol-detergent micelles. *Methods* 55:310–317
- Tieleman DP, van der Spoel D, Berendsen HJC (2000) Molecular dynamics simulations of dodecylphosphocholine micelles at three different aggregate sizes: micellar structure and lipid chain relaxation. *J Phys Chem B* 104:6380–6388
- Traaseth NJ, Shi L, Verardi R, Mullen DG, Barany G, Veglia G (2009) Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach. *Proc Natl Acad Sci U S A* 106:10165–10170. doi:10.1073/pnas.0904290106
- Triba MN, Warschawski DE, Devaux PF (2005) Reinvestigation by phosphorus NMR of lipid distribution in bicelles. *Biophys J* 88:1887–1901
- Triba MN, Zoonens M, Popot JL, Devaux PF, Warschawski DE (2006a) Reconstitution and alignment by a magnetic field of a beta-barrel membrane protein in bicelles. *Eur Biophys J* 35:268–275
- Triba MN, Devaux PF, Warschawski DE (2006b) Effects of lipid chain length and unsaturation on bicelles stability. A phosphorus NMR study. *Biophys J* 91:1357–1367
- Tribet C, Audebert R, Popot JL (1996) Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci U S A* 93:15047–15050
- Tzitzilonis C, Eichmann C, Maslennikov I, Choe S, Riek R (2013) Detergent/nanodisc screening for high-resolution NMR studies of an integral membrane protein containing a cytoplasmic domain. *PLoS ONE* 8(1):e54378. doi:10.1371
- Ullrich SJ, Hellmich UA, Ullrich S, Glaubit C (2011) Interfacial enzyme kinetics of a membrane bound kinase analyzed by real-time MAS-NMR. *Nat Chem Biol* 7:263–270. doi:10.1038/nchembio.543

- Vold RR, Prosser RS, Deese AJ (1997) Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides. *J Biomol NMR* 9:329–335
- Wacker D, Fenalti G, Brown MA, Katritch V, Abagyan R, Cherezov V, Stevens RC (2010) Conserved binding mode of human beta2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J Am Chem Soc* 132:11443–11445
- Walsh JP, Bell RM (1986) sn-1,2-Diacylglycerol kinase of *Escherichia coli*. Structural and kinetic analysis of the lipid cofactor dependence. *J Biol Chem* 261:15062–15069
- Warne T, Moukhametzianov R, Baker JG, Nehm R, Edwards PC, Leslie AG, Schertler GF, Tate CG (2011) The structural basis for agonist and partial agonist action on a (1)-adrenergic receptor. *Nature* 469:241–244
- Warschawski DE (2013) Membrane proteins of known structure determined by NMR. <http://www.drorlist.com/nmr/MPNMR.html>. Accessed 30 Aug 2013
- Warschawski DE, Arnold AA, Beaugrand M, Gravel A, Chartrand É, Marcotte I (2011) Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochim Biophys Acta* 1808:1957–1974. doi:10.1016/j.bbamem.2011.03.016
- Whiles JA, Deems R, Vold RR, Dennis EA (2002) Bicelles in structure-function studies of membrane-associated proteins. *Bioorg Chem* 30:431–442
- Williams JK, Zhang Y, Schmidt-Rohr K, Hong M (2013) pH-dependent conformation, dynamics, and aromatic interaction of the gating tryptophan residue of the influenza M2 proton channel from solid-state NMR. *Biophys J* 104:1698–1708. doi:10.1016/j.bpj.2013.02.054
- Wittlich M, Thiagarajan P, Koenig BW, Hartmann R, Willbold D (2010) NMR structure of the transmembrane and cytoplasmic domains of human CD4 in micelles. *Biochim Biophys Acta* 1798:122–127
- Wüthrich K (1986) NMR of proteins and nucleic acids. Wiley, New York
- Yang J, Ma YQ, Page RC, Misra S, Plow EF, Qin J (2009) Structure of an integrin alphaIIb beta3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. *Proc Natl Acad Sci U S A* 106:17729–17734
- Yang J, Aslimovska L, Glaubitz C (2011) Molecular dynamics of proteorhodopsin in lipid bilayers by solid-state NMR. *J Am Chem Soc* 133:4874–4881. doi:10.1021/ja109766n
- Yeh JI, Du S, Tordajada A, Paulo J, Zhang S (2005) Peptergent: peptide detergents that improve stability and functionality of a membrane protein glycerol-3-phosphate dehydrogenase. *Biochemistry* 44:16912–16919
- Yokogawa M, Kobashigawa Y, Yoshida N, Ogura K, Harada K, Inagaki F (2012) NMR analyses of the interaction between the FYVE domain of early endosome antigen 1 (EEA1) and phosphoinositide embedded in a lipid bilayer. *J Biol Chem* 287:34936–34945
- Zandomenighi G, Ilg K, Aebi M, Meier BH (2012) On-cell MAS NMR: physiological clues from living cells. *J Am Chem Soc* 134:17513–17519. doi:10.1021/ja307467p
- Zhang Y, Doherty T, Li J, Lu W, Barinka C, Lubkowski J, Hong M (2010) Resonance assignment and three-dimensional structure determination of a human alpha-defensin, HNP-1, by solid-state NMR. *J Mol Biol* 397:408–422. doi:10.1016/j.jmb.2010.01.030
- Zhao X, Nagai Y, Reeves PJ, Kiley P, Khorana HG, Zhang S (2006) Designer short peptide surfactants stabilize G protein-coupled receptor bovine rhodopsin. *Proc Natl Acad Sci U S A* 103:17707–17712
- Zoonens M, Catoire LJ, Giusti F, Popot JL (2005) NMR study of a membrane protein in detergent-free aqueous solution. *Proc Natl Acad Sci U S A* 102:8893–8898

Chapter 13

Foundations of Biomolecular Simulations: A Critical Introduction to Homology Modeling, Molecular Dynamics Simulations, and Free Energy Calculations of Membrane Proteins

Jérôme Héning, Marc Baaden and Antoine Taly

13.1 Introduction

13.1.1 *Why Study Membrane Proteins with Molecular Modeling?*

During the last decade, the number of structures of membrane proteins (MPs) has exploded. The availability of these structures has a profound impact on at least two research domains: (1) structure–function relationships of these proteins and (2) drug design, as most drugs target MPs. Interestingly, a recent review has illustrated the possibility to perform target-based drug design on G-protein-coupled receptors (GPCRs) with a collection of recent successes (Shoichet and Kobilka 2012). It is noteworthy that new molecular entities were discovered with virtual screening.

Despite this encouraging situation, not all structures of interest are available. More often than with soluble proteins important structural data are missing. In addition, structure–function studies, as well as drug design efforts, require knowledge about structural changes in the course of allosteric transitions (Changeux and Taly 2008). It should be noted that the increase in the number of structures in recent years is associated with the resolution of structures for a few states for an increasing number of MPs (see below).

Molecular modeling can fill in some of these gaps, allowing us to prepare models for proteins of interest by comparative modeling. For example, it is possible to build models for most of the numerous MPs that are targets for the treatment of

M. Baaden (✉) · J. Héning · A. Taly

Laboratory of Theoretical Biochemistry, Institute of Physico-Chemical Biology, French National Centre for Scientific Research, Université Paris Diderot, Paris, France
e-mail: baaden@smplinux.de

J. Héning

e-mail: jerome.henin@ibpc.fr

A. Taly

e-mail: taly@ibpc.fr

schizophrenia (Taly 2013). The models can then provide insight into the dynamics of such structures—although the accessible timescales are limited—as well as into their energetics of binding or of conformational change, however, imposing a high computational cost. In recent years, such theoretical approaches have been successfully used to complement experimental studies. Here, we will try to provide an insight into state-of-the-art methods by illustrating recent studies of an important family of ion channels.

13.1.2 An Example Application: The Ligand-Gated Ion Channels

Neurotransmitters—classically referred to as agonists—are chemical signals involved in the communications between neurons in the brain and between the brain and the body. There are two types of neurotransmitter receptors that bind to neurotransmitters and transmit their signal: GPCRs that mediate a slow metabolic response via second messenger cascades (Conn et al. 2009), and ligand-gated ion channels (LGICs) that mediate a fast response to the neurotransmitter (Lemoine et al. 2012).

LGICs mediate intercellular communication by converting the neurotransmitter released from the nerve ending into a transmembrane (TM) ion flux in the post-synaptic neuron or muscle fiber. They are integral oligomeric MPs that contain an orthosteric binding site for the neurotransmitter—the agonist—and an ion channel that spans the membrane. At its resting state, the channel is closed, and binding of the agonist triggers a rapid conformational change that opens the gate of the ion channel, a process called “gating.” This process, which takes place on a microsecond timescale, represents one of the most rapid conformational changes in oligomeric proteins. Once the channel is opened, cations or anions are able to diffuse through it at rates approaching tens of millions of ions per second. In addition to the well-established neurotransmission, it is now recognized that some LGICs are expressed in nonexcitable cells, suggestive of a wider functional role of these receptors. Moreover, modulation of gating can occur by the binding of endogenous or exogenous modulators at allosteric sites that are topologically distinct from the orthosteric binding sites. These modulators eventually modify the excitatory or inhibitory postsynaptic potentials. The LGICs thus present attractive targets for which more than 150 years of research have been achieved.

13.1.3 Selected Methods

Here we present methods that could typically be part of a pipeline for structure–function analyses of MPs or for drug design, using them as target: (1) homology modeling to design a starting model when an experimental structure is not available (docking is mentioned in that context), (2) normal modes of analyses and MD simulations that allow to explore protein dynamics, which is important for allosteric

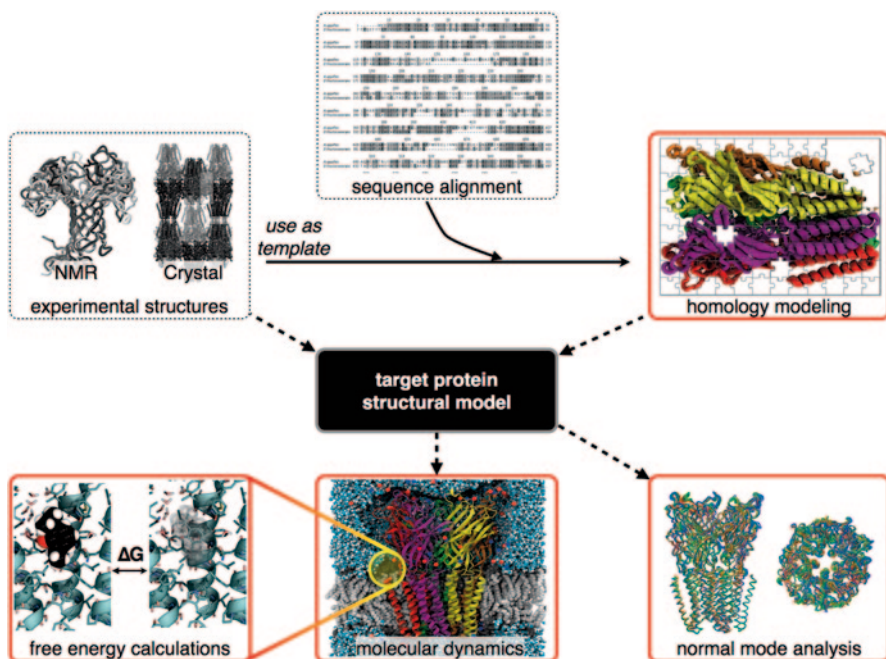


Fig. 13.1 Schematic relationship between the different modeling approaches. A structural model of the target protein is essential and can be obtained either directly from experimental methods when available or via homology modeling approaches based on sequence alignment and available homologous structures. The structural model may serve as basis for normal mode analysis or for molecular dynamics simulations. Free energy calculations expand these latter simulations by quantitatively estimating free energies, for example, of ligand binding. The four *red boxes* correspond to sections described in this chapter

regulation, and (3) free energy calculations enabling quantitative predictions, for example, of binding energies. The relationship between these approaches is illustrated in Fig. 13.1.

The present document only partly covers the literature for the sake of brevity and because it focuses on modeling results subsequently confirmed by experiment.

13.2 Homology Modeling

13.2.1 Introduction

Homology modeling, also called “comparative modeling,” allows to construct a model based on the structure of a homologous protein and has been shown to be useful in many situations (Šali 1995). This is particularly true in the context of structure-based drug design with MP targets (Taly 2013).

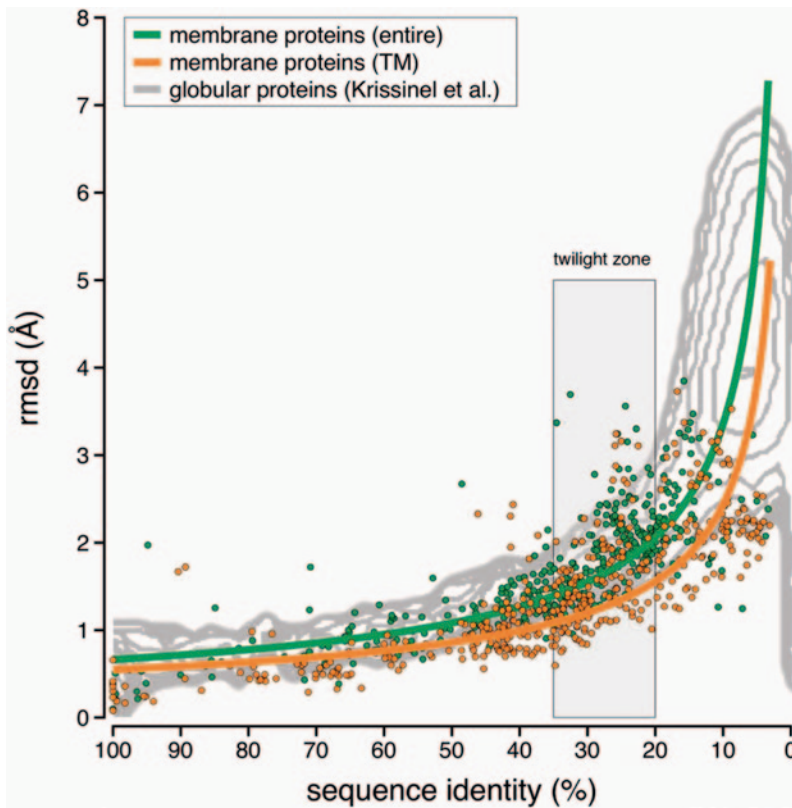


Fig. 13.2 Comparison of sequence identity and root mean square deviation (rmsd) for membrane and globular proteins. The curves fit rmsd and sequence identity values for entire membrane proteins (in green; $y = 17.0x^{-0.70}$, $R^2 = 0.63$) and for TM segments (in orange; $y = 10.8x^{-0.65}$, $R^2 = 0.61$). Both curves are superimposed to data for globular proteins represented as contour maps of the reduced density of probability of obtaining 3D alignments with the corresponding rmsd (Krissinel and Henrick 2004). Transmembrane segments present higher structural similarity (lower rmsd) than globular proteins at low values of sequence identity (<40%). (Reproduced from reference, Olivella et al. 2013, with permission)

The methods rely on the observed correlation between similarity in sequence and in structure (Chothia and Lesk 1986). This trend, initially found with soluble proteins, has recently been confirmed for MPs (Fig. 13.2; Olivella et al. 2013). Therefore, sequence identity is a key criterion for comparing structures and preparing models by comparative modeling. A value of 40% sequence identity is often considered as a conservative threshold above which homology modeling yields robust results, as illustrated below.

In the following sections, we will present three cases of LGICs that illustrate how preparing a model by comparative modeling can increase in complexity, whereby resulting models decrease in quality, when the sequence identity is decreasing.

13.2.2 *P2X Receptors, A Case of Available Template Structures of High Sequence Identity*

13.2.2.1 Structure of P2X Receptors and their Binding Site for ATP

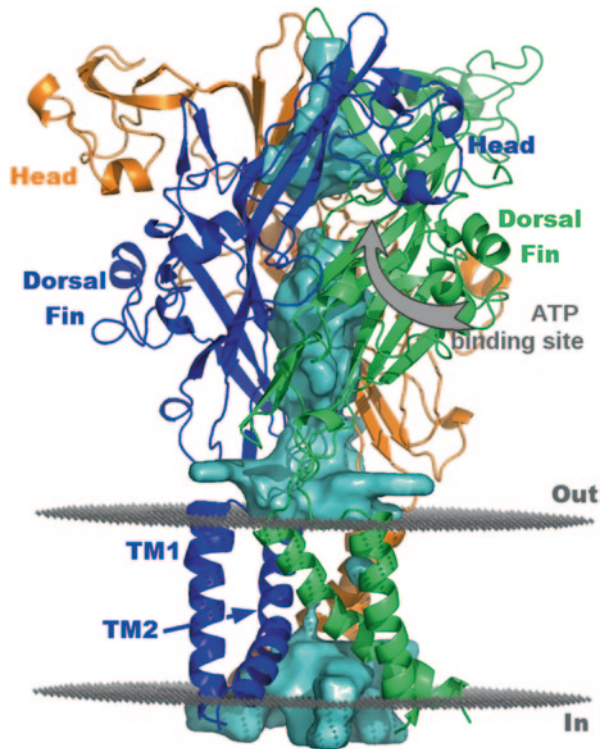
Several consensus amino acid motifs, such as the Walker motif (Walker et al. 1982) and Q-motif (Tanner et al. 2003), may account for adenosine triphosphate (ATP) binding in a wide range of ATP-sensitive proteins. Nevertheless, such consensus motifs for ATP binding are absent in the P2X receptor family. It is known that binding of phosphates of ATP molecules often involves polar and charged residues, and that aromatic residues may bind to adenine and ribose moieties. Therefore, the search of ATP-binding sites started with a focus on conserved positively charged, polar, and aromatic residues (Lemoine et al. 2012). Based on these data, a common ATP-binding site model for the P2X receptors has been proposed. However, the proposed binding model remained speculative because of the lack of a crystal structure with bound ATP. The most direct evidence for the location of the agonist-binding site in P2X receptors before crystal structure determination of a P2X receptor was provided by a study in which the mouse P2X7 residue R125 was ADP-ribosylated, a biochemical process that covalently tethers an ADP-ribose moiety (Corda et al. 2003) to proteins leading to an irreversible channel gating (Adriouch et al. 2008).

The first X-ray structure of the zebrafish, *Danio rerio*, P2X4 receptor was resolved, in 2009, at 3.1 Å (Kawate et al. 2009). In agreement with previous biochemical and biophysical data (Nicke et al. 1998; Jiang et al. 2003; Aschrafi et al. 2004; Barrera et al. 2005), the structure confirmed the trimeric assembly of the ATP-gated P2X receptors. Each subunit of the trimer has two TM alpha helices (TM1 and TM2), two intracellular N and C termini, and a large ectodomain. The two TM helices extend ~28 Å through the membrane, and the extracellular domain (ECD) projects 70 Å above the cell surface. Although the crystal structure of the zebrafish P2X4 receptor was resolved in the absence of ATP, and thus represents a closed channel state, it clearly shows that the above-mentioned conserved residues line a large and deep intersubunit cavity—there are three cavities per trimer—shaped like an open jaw that has been proposed as the putative ATP-binding site (Kawate et al. 2009).

13.2.2.2 Homology Model of the P2X2 Receptor and Experiment-Constrained Docking

The structure presented above was obtained for the *D. rerio* P2X4 receptor but the experiments presented below were performed with P2X2 and P2X1 receptors. The sequence identity between them is relatively high (almost 50% identity), which enables the preparation of accurate models usable for computational methods such as docking that require high-quality models (Fig. 13.3). Two studies that have performed experiment-constrained docking will be presented below.

Fig. 13.3 Model of the rat P2X2 receptor obtained by homology modeling using the zebrafish P2X4 receptor as a template



First, a direct evidence for the orientation of ATP in the binding site was provided by a site-directed affinity-labeling approach (Jiang et al. 2011). In this study, the sulfhydryl-reactive 8-thiocyno-ATP (NCS-ATP), a P2X2 agonist, was used to covalently crosslink the ATP moiety in an affinity-dependent manner to single engineered cysteine mutants. Out of 26 positions tested around the putative binding jaw, only L186 and N140 residues, when individually mutated into cysteines, were specifically labeled by NCS-ATP. This work thus provided direct evidence that the adenine base of NCS-ATP is close to these previously unidentified residues, and allowed us to dock NCS-ATP within the binding site of a P2X2 homology model using the geometrical constraints established by the engineered affinity labeling data (Jiang et al. 2011). Consistent with previous studies, the eight conserved residues were found in close proximity to the docked NCS-ATP.

Second, to study the exact mode of ATP binding, voltage-clamp fluorometry has been used (Lörinczi et al. 2012), in particular to investigate the cysteine-rich head domain of the P2X1R (A118-I125) that projects over the ATP-binding site. Upon substitution with cysteine residues, six of these residues (N120-I125) were specifically labeled by tetramethyl-rhodamine maleimide (TMRM) and showed significant changes in the emission of the fluorescence probe upon application of the agonists ATP and benzoyl-benzoyl-ATP (Bz-ATP). Mutants N120C and G123C showed fast fluorescence decreases with similar kinetics as the current increases. In contrast,

mutants P121C and I125C showed slow fluorescence decreases that appeared to correlate with the current decline during desensitization. Mutant E122C showed a fast fluorescence decrease and slow increase with ATP and Bz-ATP, respectively. Application of the competitive antagonist trinitrophenyl-ATP (TNP-ATP) resulted in large fluorescence changes with the N120C, E122C, and G123C mutants and minor or no changes with the other mutants. Likewise, TNP-ATP-induced changes in control mutants distant from the proposed ATP-binding sites (G115C, R139C, and W164C) were only small.

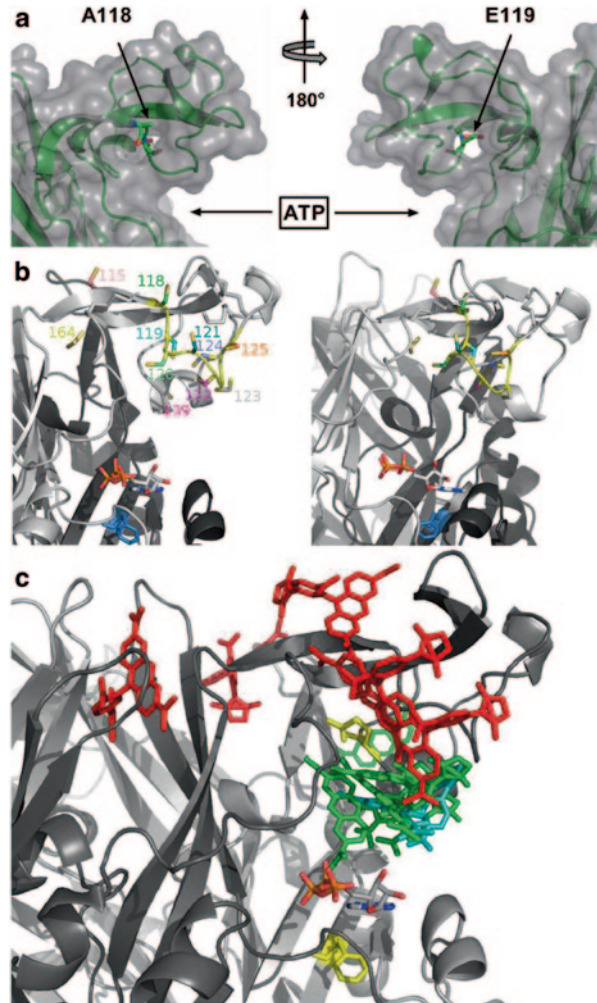
Homology models were generated for each of the P2X1 mutants using the crystal structure of the zebrafish P2X4R receptor as a template. In these models, the stretch between C117 and C126 forms a loop that reaches over the ATP-binding site, and all introduced cysteines except for I19C are predicted to be solvent-accessible. In agreement with the experimental data, docking of TMRM to the individual cysteine mutants revealed at least two poses that should allow the reaction with the side chain for all models except for the solvent inaccessible E119C mutant and the A118C mutant, in which only one pose was identified. To further interpret our data, we measured the distance between the maleimide moiety of the docked TMRM and residue F188 (homologous to residue L186 in the P2X2R) as a marker for the ATP-binding site (see above). This takes into account the flexibility of TMRM that should explore a cone in space once fixed to the receptor, and thus gives a rough idea of possible interactions of TMRM with the ligands. Two groups could be differentiated based on the functional and docking data: (1) mutants N120C, E122C, and G123C with a large TNP fluorescence, fast-decreasing signal with Bz-ATP, and distances below 21 Å and (2) mutants P121C and I125C with small or no TNP-ATP-induced fluorescence changes, slow increasing signals with agonists, and distances above 26 Å. These molecular modeling studies confirm the proposed ATP-binding site and provide evidence that ATP orients in its binding site with the ribose moiety facing the solution (Fig. 13.4).

13.2.3 *The Nicotinic Acetylcholine Receptor: A Borderline Case*

13.2.3.1 Structural Information on Pentameric LGICs

At the turn of the century, both prokaryotic and eukaryotic members had been identified in the family of K^+ and Na^+ voltage-dependent channels (Ito et al. 2004) pointing to the implication of ion channels in primary physiological function far before development of nervous systems in eukaryotes. This observation led to the quest for prokaryotic homologues of pentameric LGICs. Sequence alignments using the signature loop of $\alpha 7$ nicotinic acetylcholine receptor (nAChR) as a starting point identified genes coding for putative LGICs (at the time, 15 in bacteria and one in archaeobacteria; Tasneem et al. 2005; Corringer et al. 2010). Prokaryotic LGICs' structures are simpler than their eukaryotic counterpart: They have an ECD folded as a β -sandwich, but lack the N-terminal helix and the two cysteines that border

Fig. 13.4 Homology and docking models of the P2X1 receptor. **a** Surface representation of the head domain showing the position of the TMRM-inaccessible A118C and E119C residues. **b** Model of the proposed intersubunit ATP-binding site showing the TMRM-accessible cysteine-substituted residues. ATP was docked in the proximity of F188 (*blue*). This constraint was used as the docking was not robust because of the large binding site. **c** Model of the cys-rich P2X1 head domain with docked TMRM and ATP. TMRM docked to residues that were associated with ligand binding and desensitization is presented in *green* and *red*, respectively. TMRM bound to residue 124 is shown in cyan. Residues Q142 and F188 (homologous to P2X2 N140 and L186 identified by affinity labeling) are in *yellow*. (From reference Lörinczi et al. 2012)

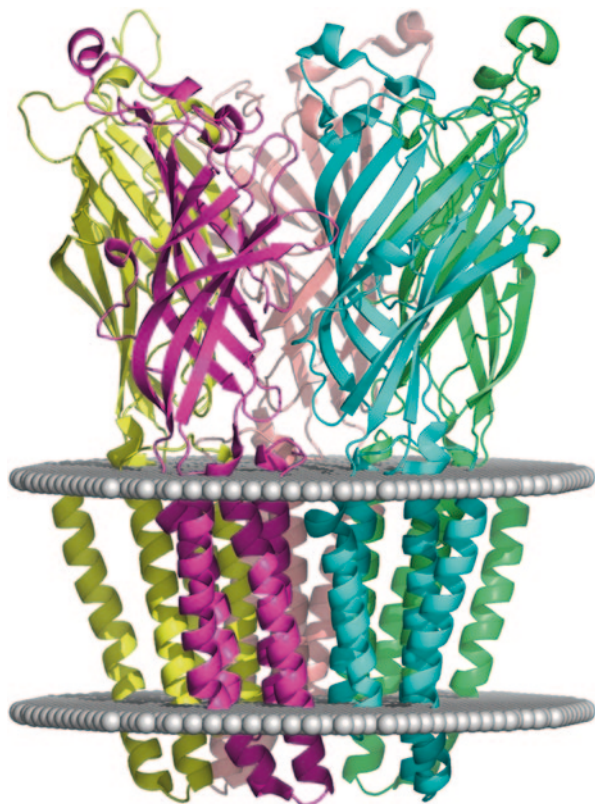


the signature loop; this domain is followed by four TM helices connected by short loops, i.e., without a cytoplasmic domain.

In this particular case, the sequence identity between eukaryotic and prokaryotic pentameric LGICs is low (~20%). Their belonging to the family was tested experimentally: the gene from *Gloeobacter violaceus* (*Gloeobacter* ligand-gated ion channel, GLIC) was cloned and the expressed protein showed a pentameric assembly (Bocquet et al. 2007). It was found to be a cationic ion channel activated by protons for which no desensitization was observed (Bocquet et al. 2007). The X-ray crystallography used crystals grown in acidic conditions that are therefore believed to represent an active state (Fig. 13.5; Bocquet et al. 2009; Hilf and Dutzler 2009).

The relevance of the results obtained on prokaryotic LGICs may be questioned, notably because they share very little sequence identity with eukaryotic LGICs.

Fig. 13.5 X-ray structure of the pentameric LGIC from *Gloeobacter violaceus*. The extracellular domain (ECD) and transmembrane domain (TMD) are indicated



As a consequence, the identification of eukaryotic LGIC ligands acting on their prokaryotic counterparts is an active area of research, and the first results of this quest indeed attenuated these concerns. Taking GLIC (further discussed below) as specific example, both TM and ECDs can be linked to human pentameric ligand-gated ion channels (pLGICs). It has been observed that GLIC proton-elicited ion currents can be regulated by anesthetics acting on gamma-aminobutyric acid (GABA) receptors (halothane, sevoflurane, isoflurane, desflurane, and propofol; Weng et al. 2010). Recently, structures of GLIC were obtained in the presence of the general anesthetics desflurane and propofol (Nury et al. 2011) and of the local anesthetic lidocaine (Hilf et al. 2010) as well as bromoform (Sauguet et al. 2013a). Such allosteric binding sites within the TM domain, also observed for ethanol (Howard et al. 2011), are closely related to human pLGICs and point to GLIC as a good model regarding the TM pharmacology of therapeutically active compounds. GLIC's ECD can be fused to the TM domain of the $\alpha 1$ glycine receptor, forming a functional chimera (Duret et al. 2011).

The ACh-binding protein (AChBP), first described in 2001, was identified from the freshwater mollusk *Lymnaea stagnalis* (Brejc et al. 2001; Sixma and Smit 2003) and has been proposed to participate in a buffering activity, modulating the

concentration of ACh in the mollusk's synapses (Smit et al. 2001). In addition to its similar function, AChBP forms a homopentamer and has significant sequence homology with the ECD of the nAChRs; roughly 30%) and the residues of the binding site are remarkably conserved as reviewed in Taly and Changeux 2008. The structure of the protein was solved by X-ray crystallography (Brejc et al. 2001), which immediately had an enormous impact on the study of structure–function relationships of pentameric LGICs.

Indeed, AChBP is the largest source of atomic resolution structures in the LGIC family offering complexes with known agonists and antagonists of the nAChRs (Hansen et al. 2005; Celie et al. 2004, 2005; Bourne et al. 2005). Therefore, the structure of AChBP has been exploited to derive models of the ECD of nAChRs (Le Novere et al. 2002; Costa et al. 2003; Iorga et al. 2006; Grutter et al. 2004; Henchman et al. 2005; Law et al. 2005a; Xu et al. 2005; Schapira et al. 2002; Sine 2002; Huang et al. 2008a; Toshima et al. 2009; Luttmann et al. 2009).

13.2.4 Building Models of the NMDA Receptor: A Complex Case

13.2.4.1 Structure of iGluRs

At the end of the last century, biophysical and bioinformatics techniques, combined with extensive mutational analyses, had already revealed the characteristic modular architecture of ionotropic glutamate receptors (iGluRs; Armstrong and Gouaux 2000; Hogner et al. 2002). They showed that the extracellular part of these receptors was composed of the amino-terminal domain (ATD; for eukaryotic iGluRs), a domain that is structurally related to the bacterial leucine-binding protein (Jin et al. 2003), and the ligand-binding domain (LBD), which shares sequence homologies with bacterial periplasmic ligand-binding proteins (Mayer 2005; Mayer et al. 2006). The LBD is formed by the association of two domains, D1 and D2, which are composed mainly of the S1 and S2 peptide segments, respectively. These techniques had also allowed the generation of computer-assisted structural models of crucial elements involved in ligand binding.

In 2009, the first X-ray structure of an intact iGluR, the rat GluA2 receptor, was solved at 3.6 Å (Fig. 13.6; Sobolevsky et al. 2009). Because the structure was solved with a bound antagonist lodged in the clamshell of each LBD, this study provided further evidence that the agonist-binding site/competitive antagonist-binding site is located within and not between subunits. The TM segments, in which the channel lies, share a conserved global architecture and fourfold symmetry with that of tetrameric voltage-gated ion channels, but with inverted topology (Karakas et al. 2009). However, the full-length GluA2 structure revealed numerous unexpected features not anticipated from previous works. First, an unexpected crossover of subunits forming dimer pairs in the ATD and LBD layers was described. Second, multiple conformations are adopted for the ATD/LBD and the LBD/ion-channel-

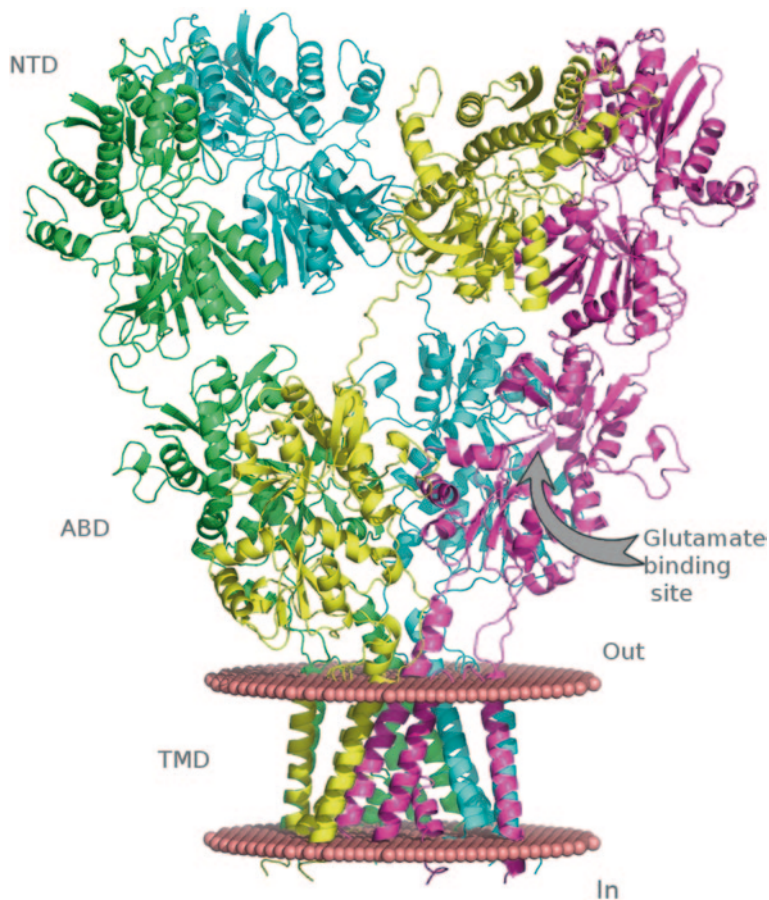


Fig. 13.6 X-ray structure of the AMPA receptor. The N-terminal domain (*NTD*), agonist-binding domain (*ABD*), and trans-membrane domain (*TMD*) are indicated

connecting linkers, leading to the amazing observation that different conformations are possible for identical subunits. Third, molecular symmetry switches from two- to fourfold between the ion channel and the ECD, respectively. Fourth, no packing contacts are observed between the ATD and the LBD on the central axis of symmetry of the tetramer assemblies, leading to an enormous chalice-shaped hole.

13.2.4.2 Complications for the Preparation of NMDA Models

Preparing a model of a *N*-methyl-D-aspartate (NMDA) receptor involves using the structure of the α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor presented above. The case of the NMDA receptor modeling is compared

with the other cases mentioned above for the following reasons: (1) the structure of the template (AMPA receptor) is incomplete, i.e., loops are missing in the structure, (2) the domains are separated by linkers that may differ between AMPA and NMDA receptors, and (3) the symmetry of AMPA receptors is complex and it is uncertain whether this feature is conserved in NMDA receptors. Many important compounds bind at the interface between domains, so their study is complex at present.

13.3 Normal Mode Analysis

13.3.1 Introduction

Normal mode analysis (NMA) approximates the surface of the conformational landscape and provides a decomposition of the movements into discrete modes. It takes advantage of a simplified but physically meaningful representation of the interaction between the atoms. Both the quadratic approximation of the landscape and the crude representation of the protein lead to the loss of the fine details of the dynamics. However, this method leads to a time-independent equation that can be solved in closed form analytically (at variance to MD), and therefore allows studying slow (biologically relevant) and collective conformational transitions.

On a database of proteins for which experimental structures of different conformations were known, it has been shown that only a handful of the very first modes (usually one or two) are usually enough to describe the transitions (Krebs et al. 2002).

13.3.2 Coarse-Grained Representation of the Protein

Coarse-grained models are simplified representations of the protein. This approximation is made at the expense of precision and the description of fine details such as local movements of loops. However, coarse-grained models have the double advantage of: (1) making calculations faster and (2) being often more robust, allowing to analyze models of medium resolutions as those presented below.

NMA of MPs regularly employs the elastic network model, which is based on simple springs connecting close pairs of atoms in the native structure (Tirion 1996). This model has been further simplified by Hinsen (1998) who showed that it was possible to use C-alpha atoms only. NMA using this approximation was shown to yield a fair description of protein flexibility (Bahar 1999; Bahar and Rader 2005) and it is the same as the one used by Krebs et al. 2002 in their survey of all conformational transitions documented in the protein databank (PDB).

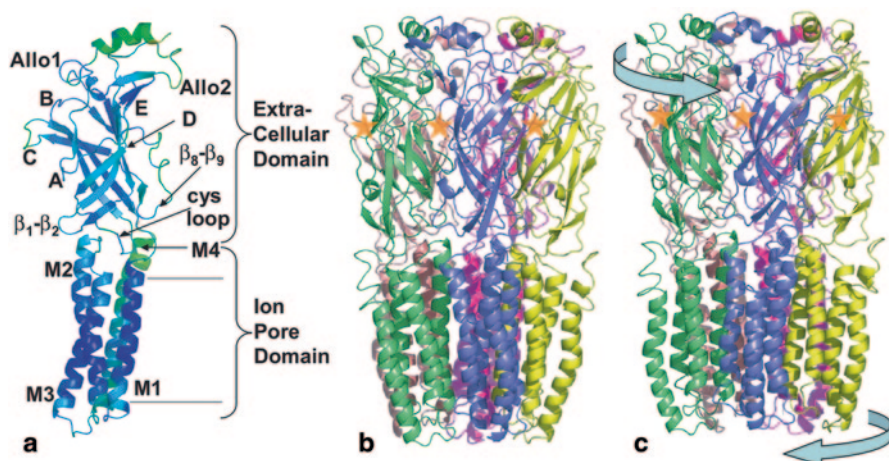


Fig. 13.7 Model structures and comparison. **a** Structure of one subunit of the present $\alpha 7$ nAChR model. The rmsd between the present and the previous model(s), computed for each atom, is displayed on the structure through a color coding from $\approx 2 \text{ \AA}$ (dark blue) to $\approx 10 \text{ \AA}$ (orange). **b** Present $\alpha 7$ nAChR model viewed from the membrane plane. A different color is used for each subunit. **c** Open-pore model obtained after exploration of the twist mode and energy minimization (see main text). The comparison of band C demonstrates the quaternary twist motion between the structures (following the arrows). Orange stars were added to help localize the binding sites. (From reference Taly et al. 2006)

13.3.3 Experimental Observations Confirming Conformational Changes Predicted from NMA

13.3.3.1 nAChR: From NMA to X-Ray Crystallography

The application of NMA to a model of alpha-7 nAChR, based on the then available AChBP and electron microscopy *Torpedo* structures (Taly et al. 2005, 2006; Taly 2007), shows that the lower-frequency mode corresponds to a global quaternary twist motion of the protein, resulting from a tilt of each subunit that causes anti-clockwise motion in the upper part of the nAChR pentamer (Fig. 13.7). The twist mode accounts for key features of nicotinic receptor channel opening and closing by agonists and antagonists and of naturally occurring mutations of the nAChRs, causing autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE) and congenital myasthenia and altering gating properties (Sine and Engel 2006; Taly et al 2006). This mode is stabilized in the resting conformation following binding of the antagonist cobratoxin (Konstantakaki et al. 2007; Samson and Levitt 2008; Yi et al. 2008).

The atomic mechanism of channel opening may be seen using the X-ray structures of the bacterial nAChRs, *Erwinia chrysanthemi* (*Erwinia* ligand-gated ion channel, ELIC), and *G. violaceus* (GLIC) stabilized in closed versus open conformation (Bocquet et al. 2009; Hilf and Dutzler 2009, 2008; Bocquet et al. 2007).

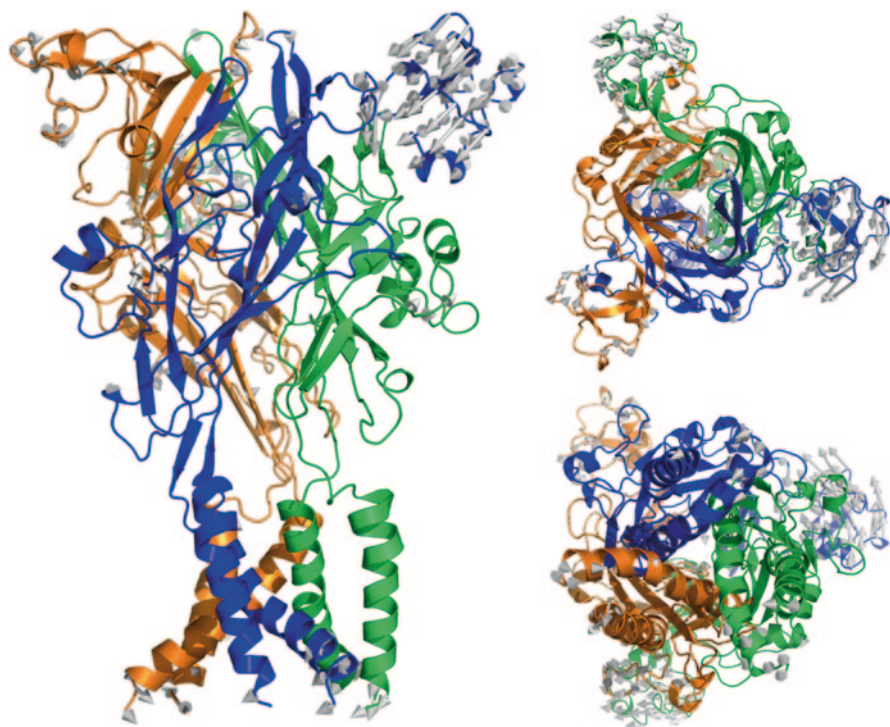


Fig. 13.8 The P2X2 receptor model is shown in cartoon representation and colored by subunits. Three views are presented: lateral to the membrane plane (*left*), top view from the extracellular side (*right-top*), and bottom view from the intracellular side (*right-bottom*). Arrows represent protein displacement captured by mode 10. Movements inferior to 1 Å are omitted for clarity

Despite low ELIC/GLIC sequence identity (18%), their common core structure undergoes a quaternary twist closely related to that described above. This twist motion contributes to at least 29% of the closed-to-open transition, with each subunit undergoing tertiary deformations (Hilf and Dutzler 2008).

13.3.3.2 P2X: From NMA to Receptor Engineering

We used NMA to reveal inherent motions of the rP2X2 receptor homology model based on the closed-state zP2X4 structure using the elastic network model. Because no clear pore opening was observed in any of the tested modes, we decided to perturb the modes by the presence of ATP. This approach approximates changes that occur on protein dynamics upon ligand binding (Taly et al. 2006; Ming and Wall 2005). We computed NMA in the presence of ATP. Interestingly, mode 10, which corresponds to an asymmetric motion of the three heads out of the membrane plane (Fig. 13.8), displayed the largest frequency difference, suggesting that the presence of ATP substantially modified the energy needed to explore this mode.

These calculations thus suggest that the domains surrounding the ATP-binding site experience significant mobility that may be related to ATP binding, and potentially to P2X gating.

In order to validate the model, we analyzed histidine residues His120 and His213 that are known to form a zinc-potentiating site (Tittle and Hume 2008; Nagaya et al. 2005; Tittle et al. 2007). Interestingly, the distance separating α carbon atoms from residues His120 and His213 that were initially found to be too far apart to create a Zn^{2+} -coordinating site shortened, following mode 10 exploration from 15.5 to 12.2 Å. More importantly, we succeeded in forming a Zn^{2+} -binding site through these histidine residues, and the distance and angle measured between Zn^{2+} and the coordinating atom NE2 of the histidine residues were close to those obtained from the analysis of (111) crystal structures of Zn^{2+} -binding proteins available in the PDB (Alberts et al. 1998). Therefore, these results suggest that mode 10 is able to form the natural Zn^{2+} -potentiating site in the rP2X2 receptor by bringing residues His120 and His213 together.

Although previous studies had shown that zinc potentiates ATP currents, there was no direct evidence supporting the hypothesis that close apposition of residues His120 and His213 gates the ion channel. We thus asked whether zinc could open the ion channel, by its own, in the absence of ATP. In human embryonic kidney (HEK)-293 cells expressing the wild-type (WT) rP2X2 receptor, extracellular zinc failed to produce detectable currents as assayed by whole-cell and single-channel patch-clamp electrophysiology. We decided to introduce the T339 S mutation, which is known to confer to the channel significant spontaneous openings in the absence of ATP (Cao et al. 2007), with the hypothesis that the mutation would be able to reduce the energy barrier for the receptor to reach the open state. Outstandingly, Zn^{2+} produced significant currents that were concentration dependent with maximal current representing about 4% of that evoked by ATP. These results thus demonstrate that coordination of Zn^{2+} ions by the pair of histidines His120 and His213 controls gating of the T339 S mutant channel.

To further explore the dynamic motion of the ATP-binding site during gating, we decided to transfer the natural zinc-potentiating site to another place of the jaw with the hypothesis that the divalent cation would also bridge the engineered histidines in the activated state(s). To select all possible pairwise positions, we computed a matrix representing the relative movement of the protein between initial and final models explored in mode 10, and selected pairs of residues in which their relative movement measured from their respective C_α atoms was greater than 2.5 Å, and the distance separating these C_α atoms was less than 13 Å in the final model, compatible with a putative zinc-coordinating site (Alberts et al. 1998). Interestingly, we found that two previously unidentified regions of the jaw satisfied these geometrical constraints, suggesting that these loops may get close enough together (Jiang et al. 2012).

To experimentally test this prediction, we first abolished the natural zinc-potentiating site in the rP2X2 receptor channel as reported previously (Nagaya et al. 2005). Introduction of the double mutant S116H/T170H* in the T339S background enabled zinc to gate the ion channel in the absence of ATP, as revealed by both whole-cell and single-channel recordings (Jiang et al. 2012).

13.4 MD Simulations

13.4.1 Introduction

MPs in their native environment are excessively difficult to model as they pose many challenges for an accurate representation of their properties. Some phenomena may be of quantum nature, for example, involving photons, chemical reactions, or electrochemical processes, and other phenomena may span timescales that cannot be reached by any currently available modeling approach. The appropriate representation and model hence depend on the question that is to be answered, and sometimes multiple scales of representation need to be combined to capture all relevant properties (Baaden and Lavery 2007). Particle-based modeling at an atomistic level using MD simulations in a lipid bilayer environment has nevertheless proven itself as a rather generic approach to describe structure–function relationships of such complex biological systems as membranes and channels (Laurent et al. 2013). In this section, we investigate the study of a single ion channel in an environment mimicking physiological conditions, using a fully flexible atomistic force field representation. For this purpose, we describe the example of the pLGIC (GLIC) from *G. violaceus* that was introduced in the previous section.

13.4.2 *The Pentameric Ligand-Gated Ion Channel GLIC as an Example*

The availability of several crystal structures of GLIC, both for an open and for a locally closed form, where the central pore does not conduct ions, combined with extensive functional characterization by electrophysiology, provides a fertile background to conduct MD simulations that can be confronted with experimental observations. At the same time, many open questions remain concerning channel gating, permeation, and anesthetics binding. Molecular-level insight into these functional aspects can be gained from MD simulations, ideally leading to experimentally testable hypotheses or providing answers to specific questions raised by experiments.

In the following subsection, we will describe how to setup such MD simulations. Many observations are not specific to the GLIC system, but would equally apply to the setup of another MP model.

13.4.3 *Setting Up MD Simulations of MP Systems*

13.4.3.1 Origin and Quality of the Initial MP Model

Reliable MD simulations of MPs require sufficiently accurate starting models. Experimentally derived structures from crystallography or nuclear magnetic resonance

(NMR) spectroscopy are among the most common sources. Based on a study of three outer MPs, it has been shown that NMR-based simulations feature globally similar conformational dynamics with respect to X-ray-based simulations, but eventually with higher mobility (Cox et al. 2007).

When the starting model is of lower or uncertain accuracy, such as with electron microscopy (EM)-derived or homology models, short MD simulations can be used as a first-quality assessment (Law et al. 2005a; Ivetac and Sansom 2008). Medium-quality, but not wrong, models may be improved by means of MD simulations (Fan and Mark 2004; Kannan and Zacharias 2010; Mirjalili and Feig 2013). However, refining protein homology models by MD simulations in general remains a challenging task with pitfalls (Raval et al. 2012).

Beyond these general observations, depending on the protein family of interest, the modeling of an initial structure may sometimes be facilitated. As described above, the Cys-loop receptor family is a benign case, even though sequence identity may be relatively low among individual homologues. All available structures resolved by crystallography so far are quite similar, providing a good basis for refinement by MD simulations (Chen and Brooks 2007). Furthermore, the regular secondary structure elements and overall topology are known to a good extent. The TM region in particular is structurally more conserved than the rest of the proteins and globular proteins (Fig. 13.2). Numerous experimental data can be used to validate or improve the models, such as knowledge about residues facing the ligand-binding sites or the ion channel pore or residues of two helices facing each other (Mnatsakanyan and Jansen 2013).

13.4.3.2 System Composition

A fundamental choice for the setup of an MD simulation is which molecular species to include in the model. Common ingredients comprise lipids, proteins, nucleic acids, sugars, ions, water, and other small molecules (alcohols, anesthetics, etc.). For membrane-embedded receptors such as GLIC, whose hydrophobic TM domain is stabilized by the membrane environment, a lipid bilayer has to be included in a realistic model. Typical biological membranes are themselves complex ensembles composed of lipids, carbohydrates, proteins, and cholesterol in the case of eukaryotic cells (van Meer et al. 2008). Furthermore, it is established that lipid “building blocks” have the potential to generate up to tens of thousands of different molecular species (Shevchenko and Simons 2010), but more lipidomics research is needed to complement our limited knowledge (Brown 2012). In MD simulations, such a complexity is very difficult to reproduce, although valuable attempts are now being made (Piggot et al. 2013). The composition of the membrane is often simplified by using a single type of phospholipid. This approach was also adopted for GLIC by embedding it in a fully hydrated phosphatidylcholine membrane (Fig. 13.9a).

For a permeating ion channel, it is furthermore important to include electrolytes in the model. Typically, on the order of 150 mM of NaCl or KCl ions are added to mimic physiologic electrolyte concentration. Higher concentrations may be used to enhance the probability of processes such as ion permeation.

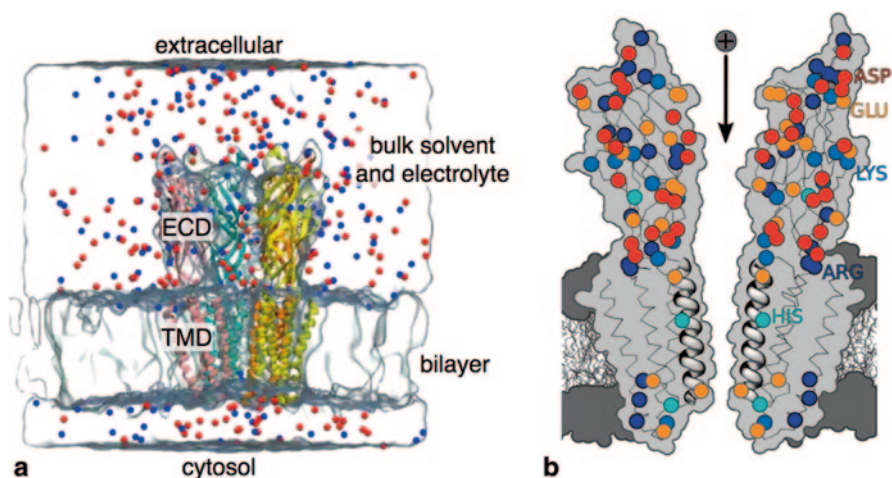


Fig. 13.9 *Panel a* depicts a typical simulation system where the GLIC ion channel shown in colored cartoon representation is embedded in a fully hydrated lipid bilayer. The extracellular (*ECD*) and transmembrane (*TMD*) domains of GLIC are annotated. Lipid and water molecules are not shown for clarity. Electrolyte is shown as cation (*blue*) and anion (*red*) spheres. *Panel b* provides a cut through GLIC's channel highlighting all ionizable side chains

In some of the studies described below, ligands such as general anesthetics were included in the model in order to study binding to the channel and modulation of its functional properties related to permeation and gating.

13.4.3.3 Assigning Hydrogens

Experimental crystallographic structures of MPs at medium resolution lack hydrogen atoms. However, in a standard MD simulation, all hydrogen atoms need to be explicitly defined, and dynamic protonation or deprotonation cannot be modeled, as no breaking or reforming of covalent bonds is possible. Hence, an important choice concerns the assignment of a protonation state for all ionizable residues and the selection of tautomers as in the case of neutral histidine residues.

The protonation state is difficult to predict and depends on a residue's local environment as well as on the protonation states of surrounding residues. This interdependence may lead to a combinatorial problem when many ionizable residues are present, such as in the GLIC pentamer with 81×5 residues (Fig. 13.9b). Standard pK_a values measured in bulk cannot be applied to buried protein residues, which is particularly critical for MPs with an environment that largely differs from aqueous solution.

We previously assessed the pK_a predictions of several widely used programs and web services (Laurent et al. 2013) yielding broadly varying pK_a shift predictions for

GLIC. This observation strongly emphasizes that accurate assignment of hydrogen atoms remains a critical and delicate step in setting up an MD simulation.

13.4.3.4 Filling Voids

MPs are often porous structures with channels, pockets, and cavities. Unless a high-resolution structure is available, it remains unknown whether and how such voids are filled. Starting a simulation with too little or too many molecules in such voids can quickly lead to artifacts and destabilize the canonical state of the model system. Hence, this question should be addressed in an initial equilibration phase, for example, by restraining the protein to let the solvent enter any accessible spaces, or by using specific approaches such as grand canonical ensemble Monte Carlo simulations (Resat and Mezei 1996).

13.4.3.5 Small System Size Effects on Concentrations

MD simulations of membrane systems are computationally expensive and hence models are typically designed as small as possible, focusing on the immediate membrane environment of the channel of interest and minimizing the bulk solvent part. This introduces a bias in the calculation of concentrations because there are ambiguities in defining them at the microscopic level. Let us consider pH as one of many possible examples. GLIC is active at acidic pH 4.6, but even at this low pH, a simulation box 50 times the size of the one typically used would be needed to observe a single H^+ ion in solution. In other words, the number of bulk water molecules is too low to extrapolate to macroscopic dimensions, which renders it delicate to exactly match experimental concentrations. This particular limitation implies that the effect of pH is solely represented by the chosen fixed protonation state of the protein.

In a membrane system, the precise assessment of concentrations may further be dynamically affected by partitioning between solution, membrane, and protein, as is the case for anesthetics or alcohols (Laurent et al. 2013).

13.4.4 Producing MD Trajectories

13.4.4.1 The Number One Limitations: Timescale, Sampling, and Statistics

For a $\sim 150,000$ atom system, such as the GLIC model described here, currently accessible simulation timescales are typically limited to the 1–10- μs timescale. This includes running simulation replica and control simulations, and hence represents an important limitation, as a simulation should be run at least ten times longer than the slowest timescale of interest (Zuckerman 2011). But even if this is possible,

sufficient sampling needs to be achieved so that the system may escape from any local minimum it may be trapped in. Furthermore, sufficient statistics on the event of interest need to be acquired. For an ion channel, this may concern the permeation of a single ion that should be observed several times before drawing conclusions from such simulations. These issues are discussed in more detail in Laurent et al. (2013).

Previous studies on the convergence of MP simulations have shown that flexible loops need longer simulations to fully converge, whereas well-structured regions, such as the barrels of outer MPs, are well sampled even in relatively short simulations (Faraldo-Gomez et al. 2004). Convergence has subsequently been discussed by Grossfield et al. (2007) and Romo and Grossfield (2011). MD simulations of GPCRs are another important source of inspiration for such considerations (Grossfield 2011).

13.4.4.2 Working Around Limitations

Specific simulation methods may be employed to work around the limitations described in the previous section. For example, to mimic biological conditions of ion permeation, an external electrical field may be applied to the simulation in various ways, typically using voltages somewhat above experimental values to accelerate the permeation process. An alternative approach is to explicitly pull ions through the channel by applying a biasing force (also see the next section on free energy calculations). A large variety of methods exist to circumvent limitations of plain MD simulations, each such method implying a specific trade-off, such as choosing a specific coordinate along which to accelerate transitions, in return.

13.4.5 Examples of Simulation-Derived Results

13.4.5.1 Confronting Static and Dynamic Views of Water and Ion Binding

A high-resolution crystal structure at 2.4 Å of GLIC's open state provided unprecedented insight into pore hydration, where two water pentagons are formed, as well as the existence of ion-binding sites in the ECD of the channel (Sauguet et al. 2013b). These observations under crystallographic conditions were compared to MD simulations of GLIC at body temperature in a bilayer environment. Overall, the ion-binding sites predicted by the MD simulation were in very good agreement with the various ions observed experimentally. Average MD occupancies reached from 1.0 to 99.9%, providing a nuanced dynamic view of ion coordination.

The presence of water pentagons in an ion channel pore had not been observed previously, raising the question of possible functional implications. The simulations revealed that the pentagon positions are well occupied on average, but not instantaneously (Fig. 13.10a). As will be described in the next section, water molecules at the pentagon positions are critical to help a permeating ion overcome bottlenecks within the channel pore when crossing the membrane.

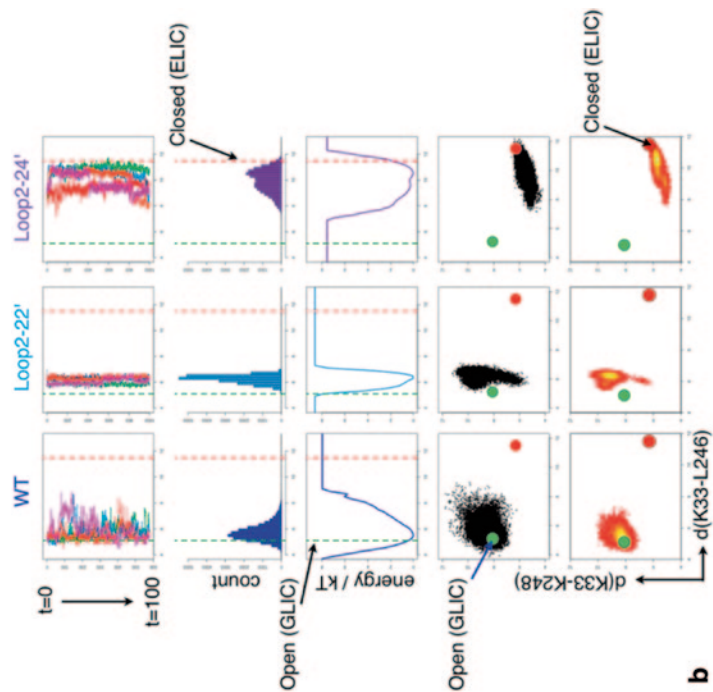


Fig. 13.10 *Panel a* highlights GLIC's central pore in the transmembrane domain, lined by the so-called M2 helices. Two water pentagons, P1 and P2, described in the text are indicated. The *left* view is an instantaneous snapshot highlighting the water structure in the pore. On the *right*, average water densities are shown. *Panel b* depicts two mutant channels described in the text, Loop2-22' and Loop2-24'. They are compared to wild-type open channel and to a closed homologue (ELIC). The top line provides the time series for each of the five subunits of a selected distance, K33-L246. The second row is a histogram of these distance series. The histogram is transformed to relative energies in row 3. Rows 4 and 5 indicate the distribution of two selected distances relative to open GLIC (*green dot*) and closed ELIC (*red dot*) channels. The main location of the distributions coincides with the phenotypes of these mutants

13.4.5.2 Predicting Critical Residues from Mechanistic Hypotheses on Ion Permeation

Using a set of four simulations where a sodium ion is dragged through GLIC's pore, we were able to observe a continuous hydration of the ion all along its path (Sauguet et al. 2013b). This permeation mechanism is very different from other cation channel families, where a permeating ion has to lose at least part of its hydration sphere. A closer analysis reveals two hydrophobic barriers in GLIC's pore where water coordination becomes anisotropic. At precisely these positions, the previously mentioned water pentagons help to maintain the cation hydration, backed by two pore-lining residues, S6' and T2'. Electrophysiology experiments on S6' and T2' mutants inspired by these observations have confirmed the critical role of these amino acids, in particular the very conserved serine. The tested S6'G and S6'V mutants lead to dramatic effects in MD simulations, notably dewetting of the pore between the 2' and 9' levels, which explains their loss-of-function phenotype.

13.4.5.3 Understanding the Molecular Determinants of Given Phenotypes

When investigating GLIC's gating mechanism, we studied a permeable A13'F mutant with paradoxical properties, as its crystal structure determined at 3.15 Å showed a pore too narrow to conduct ions (Nury et al. 2010). Although the mutant itself was not simulated, the observed extensive flexibility of the top of the wild-type pore-lining M2 helix in immediate vicinity of the 13' side chain offers a possible explanation. The mutant crystal structure may be trapped in a closed state where M2 flexibility is blocked. In a membrane environment, it will recover its enhanced flexibility and enter a conformational equilibrium between an open and a closed pore. This would explain the observation of a functional channel phenotype by electrophysiology.

Subsequently, a locally closed form of GLIC was characterized, its hallmark being a local closure of the top of the M2 helices, fully in line with this hypothesis on the A13'F mutant (Prevost et al. 2012). Cross-linking the M2–M3 loop with K33 on loop 2 triggered the local closure of GLIC, producing two particularly interesting phenotypes. The loop 2-L22' mutant is the only functional cross-linked mutant and, correspondingly, MD simulations show that it behaves very similarly to the wild type in its open form (Fig. 13.10b). The loop 2-K24' mutant on the other hand is not functional and requires reduction of the disulfide bridge to recover a current. Its properties observed in MD simulations closely match characteristics of the ELIC channel, a supposedly closed state, in line with the observed phenotype.

13.4.5.4 Exploring Conformational Transitions

In 2010, we studied GLIC's gating mechanism in a 1- μ s MD simulation where the switch from an open to a closed conformation was triggered by a pH jump

(Nury et al. 2010). Despite the large amount of computational resources corresponding to approximately 10 months of calculations on a supercomputer in 2009, i.e., tens of years on a desktop machine, only two protomers had fully undergone the transition to a closed state at the end of the simulation. This observation suggested that a much longer simulation was required to achieve a complete gating transition. In a collaborative effort, we are currently analyzing a more than 20- μ s dataset featuring several gating events to address this question more fully. As an important feature, we observe that dewetting of the channel pore occurs early in this process and may be a key event in preventing ion permeation. The actual conformational change is more difficult to pinpoint, as the simulation data are intrinsically noisy. However, important aspects such as the twist motion mentioned above are clearly observed and confirm previous hypotheses.

Another aspect of the conformational changes related to gating GLIC's channel is the previously mentioned motion of the top of the M2 helix, documented by the crystallographic structures of the locally closed form (Prevost et al. 2012). Interestingly, MD simulations suggest that this state is spontaneously accessible from the open form.

13.5 Free Energy Simulations on Mps

This section is somewhat technical, which is not so much intentional as unavoidable, given the highly technical field it deals with. For a much more complete, yet accessible, introduction to the statistical physics of biomolecules, the reader is referred to a book dedicated to the topic (Zuckerman 2010).

13.5.1 Free Energy Simulations: Why and How?

13.5.1.1 Free Energy: a Descriptor of Equilibrium Probability

Our best tool to describe biomolecular processes quantitatively is statistical thermodynamics, which starts with a microscopic picture and yields macroscopic quantities that lend themselves to experimental measurement. As will be seen in the examples provided in this section, free energy can be an invaluable link between the atomic details of a simulation, far beyond experimental reach, and quantities that make sense from an experimental perspective.

Under constant temperature conditions, thermodynamic equilibrium is governed by free energy—in other words, perhaps more accurate, free energy is the quantity we use to describe equilibrium. Nonequilibrium processes have their own dedicated field of thermodynamics, which we will not delve into here. This may seem ironic, as sustained departure from equilibrium is a defining property of living organisms. While equilibrium thermodynamics cannot fully capture the function of living

systems such as cells, it proves powerful in describing life's elementary sub-processes at the molecular level. Moreover, life's steady-state nonequilibrium regime is often close enough to equilibrium to be captured as "equilibrium plus a small perturbation," e.g., a linear response model.

Consider an enzyme able to switch between conformations C (closed) and O (open). This equilibrium is captured quantitatively by a law of mass action with equilibrium constant K , and undergraduate chemistry states that:

$$K = \exp\left(\frac{-\Delta G_0}{RT}\right) \text{ or equivalently, } -\Delta G_0 = RT \ln(K), \quad (13.1)$$

where ΔG_0 is the standard Gibbs free energy difference for the transformation. This seemingly unimpressive relationship actually contains a tremendous amount of physical insight.

Suppose that ΔG_0 is equal to three times the thermal energy, $3 \times RT \sim 7.5$ kJ/mol, meaning that K is about 1/20. The consequence is that, in a macroscopic sample at equilibrium, the concentration of enzymes in the closed state is roughly 20-fold that of enzymes in the open state. Let us now translate this description into microscopic terms: The concentration ratio means that for any given protein molecule, at any point in time, the probability of state C is 20 times that of state O; concentration is the macroscopic reflection of microscopic probability.

Yet in microscopic terms, proteins are dynamic and can obviously occupy many more than two conformations. We must therefore differentiate microscopic states (*microstates*) from the *macrostates* "open" and "closed" that we define and consider to be of biophysical relevance. In statistical physics, this conformational equilibrium will therefore be described as an ensemble of many microstates, belonging to either macrostate. The probability of a macrostate is just the sum of probabilities of its constituent microstates. The set of all possible microstates is called *configuration space*. Molecular simulations describe microstates, and can be seen as tools to explore configuration space. Additive changes in the standard free energy difference correspond to multiplicative changes in probability (a change of RT for a factor of e). The exponential relationship between *free energies* of *macrostates* and their probabilities is a consequence of the Boltzmann distribution, which is an exponential relationship between the *energy* of *microstates* and their probabilities. This form of relationship—probabilities as exponentials of energies, and its inverse, free energies as logarithms of probabilities—is pervasive in the statistical physics of systems at constant temperature.

13.5.1.2 Free Energy Profiles are Not Free Energies

We have described a conformational switch as a discrete process, but some processes are better described as continuous, and measured by a reaction coordinate.

Consider for example the position z of a molecule diffusing across a TM channel. The equilibrium statistics of this variable are captured by its probability density $\rho(z)$, such that the elementary probability at position z is $dp = \rho(z)dz$.

Analogous to Eq. (13.1), a free energy profile is defined as a function of z , as:

$$A(z) = -RT\ln(\rho(z)) + C. \quad (13.2)$$

It is called either free energy profile or *potential of mean force* (PMF), as it corresponds to the reversible work exerted on the system when traveling along z , and can be calculated as the work of a mean force acting on the coordinate (see estimators in the next section). The PMF is defined up to an arbitrary additive constant C : adding a constant to the PMF is equivalent to multiplying the probability density by a constant ($\exp(-C/RT)$), which does not change its physical interpretation. As a result, *any single value of the PMF is meaningless*: only its variations are significant.

For any value z_0 , the quantity $A(z_0)$ characterizes a state defined by $z=z_0$, in other words, the ($z=z_0$) isosurface of the system's configuration space. This is intuitively understood as "the free energy of the permeating molecule at position z_0 in the channel." The physical meaning of such a state is actually not that simple, as it is neither a single microstate, nor a true macrostate: Geometrically speaking, it is an infinitely thin slice of configuration space, orthogonal to the z direction. As a result, its probability is zero (as with any single value of a continuous random variable).

By contrast, physically meaningful macrostates can be defined in terms of coordinate z , in the form of finite ranges of the variable, with nonzero probability. In our example, these could be one side of the membrane, different regions of the pore, and the other side of the membrane. While these mathematical definitions are arbitrary, they take physical significance if they correspond to free energy basins, separated by barriers. In kinetic terms, the states are meaningful if the system takes less time to evolve within a state than it does to hop between states. Its long-time dynamics should therefore be a succession of long periods of fluctuation within basins, punctuated by rare barrier-crossing events.

However, such a clear kinetic description is often challenged by the complexity of actual biomolecular processes, which often exhibit exceedingly complex energy landscapes, with many basins, small and large, short- and long-lived, linked by many possible transition pathways, occurring on a broad range of timescales. Protein folding is a classic example. In such cases, the dimension reduction problem, i.e., coming up with a few variables that describe the process of interest, becomes challenging, if not entirely intractable.

Summarizing the relationships given by Eqs. (13.1) and (13.2): Free energy is to probability what a free energy profile is to probability density (that is, its logarithm). Just as probability density is not a probability, a free energy profile is not free energy. Yet one can be derived from the other by integrating the probability density over a finite range.

13.5.1.3 From Theory to Computation 1: Approximate Methods

Because of the complexity of protein energy landscapes, it is, generally speaking, a difficult task to explore and analyze them with exact methods, which will be described in detail below. However, numerous projects benefit from a rapid evaluation of the free energy (such as virtual screening approaches or post-processing of molecular dynamics trajectories as in the molecular mechanics Poisson–Boltzmann surface area, MM/PBSA, method). Approximate methods are therefore widely used as a computationally cheaper—but also less accurate—alternative. Docking is an example of approximate methods. It nevertheless requires structures/models of high quality. This is illustrated by the use of AChBP in docking studies of nAChR ligands (Turabekova et al. 2010; Gao et al. 2003; Abin-Carriquiry et al. 2010; Slavov et al. 2010; Utsintong et al. 2009; Babakhani et al. 2009; Ulens et al. 2009; Artali et al. 2005; Sine et al. 2004; Konstantakaki et al. 2007). It was, for example, shown that alpha-toxin is able to bind only when the binding site is open (Konstantakaki et al. 2007). Interestingly, similar observations have also been made with nAChR models (Haddadian et al. 2008; Huang et al. 2008b).

Describing the wealth of approximate methods is beyond the scope of this manuscript and we will concentrate on exact free energy approaches.

13.5.1.4 From Theory to Computation 2: Exact Estimators

In practice, a simulation of a membrane system yields a trajectory (set of configurations) sampled according to a well-defined probability distribution. In the process, energies and often forces are calculated. Free energy estimators are expressions that can be evaluated numerically, starting from those known quantities, and yield an approximate value of a free energy difference or PMF.

Estimators have different qualities: computational cost, bias (systematic error), and speed of convergence as sampling increases (or equivalently, variance of the estimate for a given amount of sampling).

Histograms

Once the states of interest are sampled in the simulated trajectories, probability estimates may be obtained by simple counting (in the case of reaction coordinates, binning into histograms). In practice, we would run a simulation of our conformation-switching enzyme and count the ratio of occurrences of states C and O, or collect a histogram of the z positions of the permeating molecule in its channel. The logarithmic relationship given by Eqs. (13.1) and (13.2) then provides a straightforward way to estimate the free energy.

While conceptually and computationally simple, this method is not used very often, except for the fastest-relaxing systems, due to its large statistical uncertainty. In general, interesting degrees of freedom need to diffuse and cross barriers to relax.

Therefore, they are typically slow relaxing in atomic-scale simulations, and the histograms, or state occupancies, converge slowly. This issue can be overcome through the application of artificially modified simulation potentials.

Umbrella Sampling

In umbrella sampling as it is most commonly practiced, a number of simulations are run concurrently, each with a modified potential energy that confines the reaction coordinate to oscillate in a narrow range (a *window*), so that the combined sampling of all simulations covers the full pathway. This makes diffusion and barrier crossing along the reaction coordinate unnecessary. The partial, overlapping histograms from all windows are combined in a statistically effective way, yielding a well-converged, complete histogram from which the free energy profile is then computed.

The main limitation of such a stratified approach is the relaxation of degrees of freedom other than the reaction coordinate (orthogonal degrees of freedom). Convergence within each window is only reached when all other degrees of freedom have been well sampled, which the confining potential (along the reaction coordinate) does not aid, but more likely hampers. When monitoring only the reaction coordinate, relaxation of orthogonal degrees of freedom may be difficult to estimate.

Exponential Formula (Free Energy Perturbation)

Consider two states, A and B, defined by their respective potential energy functions V_A and V_B . For each configuration, one may calculate the energy difference $V_{AB} = V_B - V_A$. How can we calculate the free energy difference $F_{AB} = F_B - F_A$?

A fascinating consequence of the exponential relationships outlined above is this expression for the free energy difference:

$$\exp\left(\frac{-\Delta F_{AB}}{RT}\right) = \left\langle \exp\left(\frac{-\Delta V_{AB}}{RT}\right) \right\rangle_A. \quad (13.3)$$

This states that the difference in free energy between A and B is the exponential average of the difference in potential energy between these two states. The average is weighted according to the probability distribution of state A. Swapping the labels A and B, one obtains an equivalent formula where B is the reference state. This estimator has been widely used, particularly in the context of “alchemical” free energy simulations (see the next subsection on binding for an example).

The weakness of this estimator lies in the properties of the exponential average, where large positive values of ΔV give nearly zero contribution, while negative values carry enormous weight. The average is dominated by a few large values, which occur as rare events in the simulation, slowing down convergence. Variants of this estimator with much improved convergence behavior are covered in detail in Chipot and Pohorille 2007.

Thermodynamic Integration

Taking the definition of a free energy profile, and differentiating it with respect to z , we obtain:

$$\frac{dA}{dz} = \left\langle \frac{dV}{dz} - RT \frac{d \ln |J|}{dz} \right\rangle_z. \quad (13.4)$$

Note the two terms in the average, one corresponding to physical interactions, and the other purely geometric in origin. dV/dz is the opposite of the force acting on the coordinate z , so the free energy derivative appears as a *mean force*, justifying the term “potential of mean force” for the free energy profile.

In general, z is not a Cartesian coordinate; it is defined through a coordinate transform for which J is the Jacobian matrix. The second term in the average accounts for the fact that the very geometry of the reaction coordinate can influence its probability distribution. Suppose that the reaction coordinate is the distance between two molecules: If they diffuse randomly in the absence of any force between them, they will tend to diffuse away from each other, and large distances will be much more probable than small ones. The intermolecular distance d is obtained when one molecule lies on a sphere of radius d centered on the other one, so that the probability density grows as the surface area of that sphere, i.e., as the square distance. Suppose now that the molecules are two integral MPs: Since they lie on a two-dimensional surface, a given distance is defined by a circle in the membrane plane, and the probability distribution is then proportional to d . The second term in the average of Eq. (13.4) can be understood as a form of centrifugal force, due to the tendency of particles to diffuse away from one another for purely geometric reasons.

13.5.1.5 Sampling and the Error-Bar Problem

All estimators listed in the previous section are mathematically exact; they are not based on any approximation whatsoever. So where is the catch? They are all asymptotic expressions involving infinite sums. Those are approximated by finite sums in our numerical estimates, which then suffer from finite sampling error.

Any free energy estimate from simulations will therefore combine two forms of error:

1. Systematic modeling error due to the physical model such as molecular mechanics, force field parameters, components of the simulated system, etc.
2. Stochastic error due to finite sampling

Both types of error are difficult to estimate. The specific problem with finite sampling error is the temptation to resort to simplistic statistical formulae that do not yield reliable upper bounds on the error in simulations of complex, slow systems such as MP models. The variance on a numerical estimate of an average is the

intrinsic variance of the sampled quantity, divided by the number of statistically independent (uncorrelated) samples taken.

When simulating MPs on the atomic scale, obtaining mathematically perfect sampling is unlikely, as some degrees of freedom are just too slow (the protein will not unfold and refold within a simulation), therefore one can only hope to achieve *local convergence*, within the region of conformation space that is of biological interest. That is often achievable, but requires care and skill.

Significant finite sampling errors come from regions of configuration space that are physically important, but were not sampled in the simulation. This incomplete sampling would lead to underestimating the variance of the averaged quantity, and potentially its correlation time as well, thus overestimating the number of independent data points. In summary, the main source of statistical error is also a source of error on the error estimate itself, with a tendency to underestimate the error!

Let us pause here for a second and add that this issue is by no means limited to simulations: wet laboratory experiments may also suffer from biases (such as defects in the materials and instruments) that persist across replicated measurements, and thus are not accounted for by the classic type of error bars. We must then argue that no error bar can be taken at face value, and that one should always question the kinds of error that it reflects, and those it does not.

In short, we cannot count on a simulation to tell us about the information it missed. The only way out of this predicament is to build a physical intuition, a general understanding of the molecules and process under study. Only monitoring of the simulations with a trained eye can ensure that important configurations are visited, and detect cases where sampling is hindered by slow transitions.

13.5.2 *Binding: Interaction of General Anesthetics with GLIC*

13.5.2.1 **The Question: Sites of General Anesthetic Action**

While general anesthesia has been used clinically for more than one and a half century, its detailed molecular mechanisms have only recently started to be unraveled, thanks to current structural and biophysical investigation methods. Among the molecular targets of general anesthetics are synaptic ionotropic receptors of the pLGIC family, including the nAChR and GABA type A (GABA_A) receptors. Some of these channels are cation-selective and participate in excitatory synapses (e.g., nAChR), while others are anion-selective and found at inhibitory synapses (e.g., GABA_A). One of the mysteries still surrounding general anesthetic action on these receptors is that anesthetics tend to inhibit excitatory channels, while they potentiate inhibitory channels. How can a single drug exert opposite modulating effects on structurally related targets? The consensus hypothesis is that of allosteric regulation of channel opening through one or more binding sites located in the TM domain of pLGICs.

As discussed previously, a crystallographic structure of the bacterial GLIC channel bound to the anesthetics propofol and desflurane was resolved (Nury et al. 2011).

The anesthetic-bound structures revealed a binding site within the TM segment of each subunit, closer to the extracellular side and the interface between the TM and ECD.

13.5.2.2 Technical Approach and Pitfalls, Part I: Flooding

In search of potential anesthetic binding sites on pLGICs, we investigated binding of propofol and isoflurane to both the nAChR and GLIC channels through undirected “flooding” simulations, wherein anesthetics were introduced in the solvent surrounding the receptor models, and left to diffuse and bind spontaneously (Branigan et al. 2010). That study predicted several likely binding sites, including the intrasubunit site later revealed by the co-crystal structures.

A more unexpected result was the rapid and persistent binding of anesthetics at the hydrophobic constriction forming the gate within the TM pore. Binding to such a site would result in pore block, in contrast to the allosteric mechanisms envisioned before. As a result, we postulated a mechanism resting on the counteracting effects of several modulatory sites: binding to allosteric sites would cause potentiation of channel opening, while binding to the pore would block the channel. The overall modulatory effect would then result from the fine balance of these two effects on each particular channel, explaining how closely related channels may undergo opposite modulation by the same molecule.

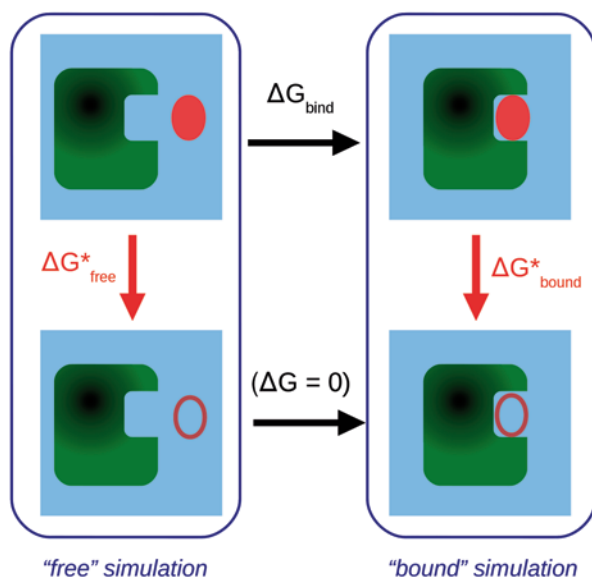
Flooding simulations have the advantage of being free from researcher bias. They may be regarded as a (very expensive) variant of docking calculations, including full receptor flexibility as well as explicit solvent at atomic resolution. While much more detailed than docking, flooding exhibits much less favorable convergence properties, as spontaneous diffusion of the ligand is far less effective than the stochastic search algorithms used in docking. Therefore, unless it is performed over impractical and unaffordable simulation times, its results carry significant statistical uncertainty. One way to ascertain convergence is to observe multiple binding and unbinding events, which in this case happened in some but not all of the sites.

Thus, our flooding simulations indicated that anesthetic binding to the pore was possible and apparently fast, but remained at a qualitative level. To turn this result into a computational prediction, more quantitative data were needed in the form of predicted binding affinities for the two sites of interest: the allosteric site and the TM pore, just above the hydrophobic gate (LeBard et al. 2012). Our competing sites theory would predict that anesthetics bind to the GLIC pore with higher affinity than for the intrasubunit, allosteric site, resulting in net inhibition of channel function.

13.5.2.3 Technical Approach and Pitfalls, Part II: The Way of Alchemy

To predict the affinity for each site, we calculated binding free energies through alchemical free energy calculations, using the thermodynamic cycle illustrated in Fig. 13.11. This technique is grounded in a simple property of free energy: It is a

Fig. 13.11 Schematic illustrating the thermodynamic cycle used for the calculation of a binding free energy through “double decoupling.” A protein (*green*) and its ligand (*red*) are embedded in a solvent (*blue*). The physical process of interest is binding of the ligand to the protein (*top*). In the alchemical simulations, the ligand becomes decoupled from its environment (*red outline*)



state function, that is, its variation depends only on the initial and final states A and B considered, and not on the pathway actually taken when going from A to B. In the context of a numerical simulation, this opens the door to choosing pathways based on their numerical convenience and efficacy, rather than physical realism. What does matter, of course, is that the end points are physically realistic.

In the case of binding, instead of having the ligand enter and exit the binding cavity through space (a generally unknown pathway, potentially featuring barriers and other sampling difficulties), one may decouple the ligand from its environment by artificially turning off the relevant intermolecular interactions (red arrows in Fig. 13.11). The decoupled ligand is then effectively in vacuum, and invisible to the receptor. The same process is repeated in solvent. In each case, the “decoupling free energy” ΔG^* is computed using either the exponential formula or thermodynamic integration (Sect. 13.5.1.4), or variants of these methods. The closed thermodynamic cycle represents an alternate pathway to binding: decoupling the ligand from water, followed by recoupling to the receptor. This yields the following expression for the binding free energy:

$$\Delta G_{\text{bind}} = \Delta G_{\text{free}}^* - \Delta G_{\text{bound}}^* \quad (13.5)$$

A very notable use of this technique on MPs is a calculation of free energy penalties for inserting polar and nonpolar amino-acid residues into the membrane through the bacterial translocon (Gumbart et al. 2011). That work used two thermodynamic cycles, describing insertion into either the lipid phase or the translocon protein SecY. It demonstrated that the seemingly incompatible computational and experimental results on membrane insertion free energies could be reconciled if the reference

state was taken to be the cavity inside SecY. The underlying hypothesis is that insertion of the nascent polypeptide into the translocon is energized by the peptide synthesis itself, and thus has no bearing on the final thermodynamics of insertion.

13.5.2.4 Results: High Pore Affinity, with Dependence on Structural Model

As outlined in the description of MD simulation setup above, setting up a MP simulation entails carefully choosing many parameters that may affect the outcome of the simulation. The main challenge with GLIC is that its pore tends to spontaneously switch to a closed state under conditions where it is expected to remain open. For this reason, we performed the affinity calculations both in an unrestrained, but closed state, and using artificial restraints to keep the pore open (LeBard et al. 2012).

The predicted affinity of propofol for the pore is close to 1 μM regardless of conditions, and its predicted affinity for the allosteric site is 10 μM . Isoflurane proves more sensitive to the pore conformation: Its affinity for the open pore is lesser than the closed pore (3 vs. 600 μM), while binding to the allosteric site is less favorable, around 3 mM. Experimental IC₅₀ is 25 μM for propofol, and 60 μM for isoflurane. Therefore, the affinity of propofol for the pore site seems somewhat overestimated, while the range of predictions for isoflurane contains the experimental value, but is very broad. This underlines the fact that obtaining well-converged free energy predictions from large, sensitive protein models is extremely challenging.

However, our prediction of pore blocking has recently received experimental backing in the form of co-crystal structures of the ELIC channel with bromoform (Spurny et al. 2013). The structures of Spurny et al. clearly show evidence of general anesthetic binding to a pLGIC pore.

13.5.3 Assembly: Lateral Interactions of TM Helices

13.5.3.1 The Question: Driving Force of Lateral Interactions in Glycophorin A

Large classes of integral MPs are bundles of TM α -helices: The tertiary structure of these bundles is held together by lateral interactions between helices. A relatively simple model for such lateral interactions is afforded by bitopic proteins, whose TM domains consist of a single helix. Among those, glycophorin A (GpA) is a classic model for intra-membrane helix dimerization: It has a strong tendency to dimerize in lipids, and even in detergents such as sodium dodecyl sulfate (SDS). An NMR structure of the dimer in detergent micelles was determined in 1997 (MacKenzie et al. 1997), providing a finer structural view of the mode of association of individual TM helices.

MD simulations are based on the calculation of forces between atoms; as such, they are particularly well suited to complement a structural description of molecular

processes with a physical sense of the interactions at play. When simulating the dissociation of GpA, the strength of the interaction is a problem, as dissociation is an exceedingly rare event on the simulation timescale. For that reason, we chose to facilitate the sampling of GpA dissociation by introducing an adaptive bias in our simulation (Hénin et al. 2005).

13.5.3.2 Technical Approach: Adaptive Biasing Force

A slow process such as GpA dissociation can often be described by a reaction coordinate whose dynamics is hampered by free energy barriers. In principle, we can calculate these barriers: Thermodynamic integration (see section 13.5.1.4 above) states that the slope (gradient) of the free energy profile at a given point along a coordinate is the average force felt by this coordinate. This gives a method to calculate the free energy profile, provided that we can generate a trajectory where this coordinate is well sampled at equilibrium: This is particularly difficult if there are high free energy barriers along the way.

To obtain such a trajectory in an affordable time, one may apply an external bias so that the free energy barriers are easier to overcome. If the biasing potential is exactly the opposite of the free energy profile, the barriers can even be completely erased. However, defining such a bias requires knowledge of the very free energy profile that we want to calculate, resulting in a catch-22.

One possible solution is to start calculating an estimate of the local free energy gradient, and apply a biasing force that cancels out that current estimate. As the simulation progresses, more data are accumulated and the estimate of the gradient is refined; meanwhile, the biasing force is updated continuously. This algorithm is known as the adaptive biasing force, or ABF (Darve and Pohorille 2001). ABF works close to equilibrium, as the biasing force varies slowly for an initial period, and then remains nearly constant. Its main advantage is to require mainly a choice of reaction coordinates, with few other tunable parameters.

This work on GpA relied on our implementation of ABF in the *Nanoscale Molecular Dynamics* (NAMD) program (Hénin and Chipot 2004; Hénin et al. 2010). The aim was to model a bilayer environment, but slow-relaxing phospholipids would have slowed down an already slow process, so we chose a faster-relaxing model: a liquid alkane slab surrounded by water.

13.5.3.3 Results: Barrier-Less Association with Two Interaction Regimes

The ABF simulations yielded the free energy profile represented in Fig. 13.12. It is fairly simple, and shows a single, deep well around the dimer distance, and a slope decaying into a long-distance plateau without showing a barrier or other salient features. The lack of a barrier suggests that dimerization is not an activated process.

Part of the magic of simulations is that everything is based on an explicit, detailed physical description. In the framework of thermodynamic integration, the

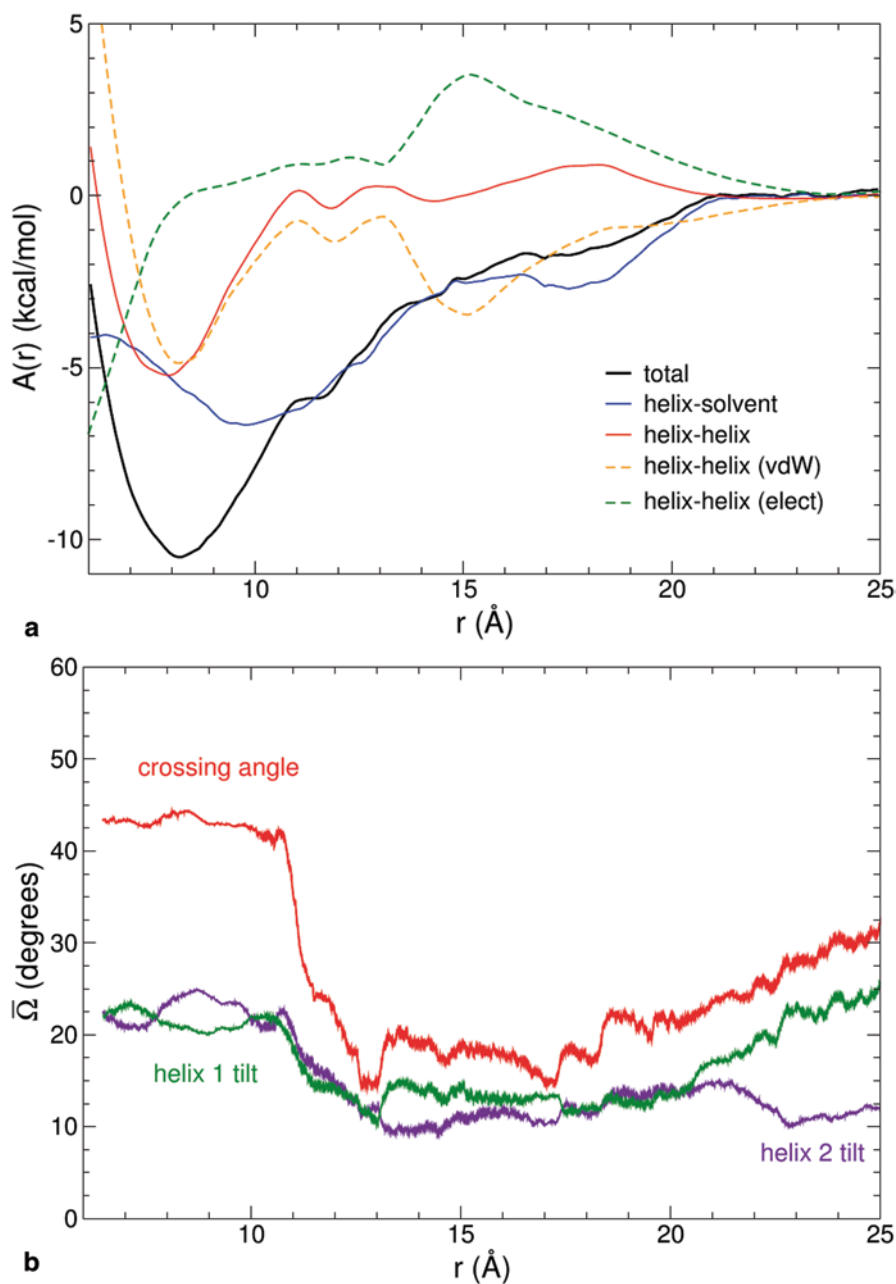


Fig. 13.12 *Left:* Free energy profile for dimerization of the GpA TM segment, and its decomposition into additive terms. *Right:* Average dimer crossing angle and individual helix tilt as a function of inter-helical separation

“mean force” may be decomposed into force components of different physical nature, and integrated separately to yield components of the free energy profile. An important caveat: Although this decomposition is mathematically unique, it does not fully separate the various physical interactions (that is not possible), so that each term still depends on the presence of all other interactions. Still, this decomposition is informative on a qualitative level.

The free energy plotted in Fig. 13.12 is decomposed into a protein–protein and a protein–solvent term. The protein–protein term displays a well at short separations (under 11 Å) that largely accounts for the well in the overall PMF (black line). Beyond 11 Å, that component is nearly flat. In contrast, the protein–solvent term (blue line) is close to the overall PMF above 11 Å. Thus, two interaction regimes are visible: A long-range regime dominated by solvent-mediated interactions, and a short-range regime dominated by direct protein–protein interactions. In turn, the protein–protein component can be separated into electrostatic and van der Waals interactions. The electrostatic component (green dashed line) shows a classic dipole–dipole repulsion between the parallel helices at large distances. Perhaps surprisingly, this repulsion is absent from the total PMF, as it is compensated, mostly by the other protein–protein forces, and partly by the solvent-mediated terms.

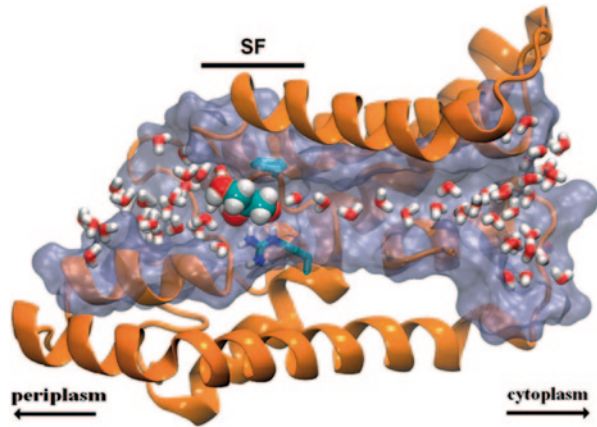
The two interaction regimes also have structural manifestations: The crossing angle of the dimer remains native in the short-range regime (around 45°), and decreases sharply as the helices move apart, indicating that the association is now loose enough that the helices are free to tilt independently.

Based on the PMF, a dimerization free energy can be calculated. In our conditions, we find a dissociation free energy of 11.1 kcal/mol. However, comparison with experiment is very difficult, due to the lack of a common, well-defined reference state for experiments performed in nonhomogeneous environments such as micelle solutions. Fortunately, this reference state problem vanishes if one considers relative rather than absolute affinities, for example, the change in affinity due to a point mutation: Comparison with experiment is then much more straightforward. Such a relative affinity may be predicted using another type of free energy simulation, “alchemical transformations,” presented in a subsequent section 13.5.2.3. We have calculated the change in dimerization free energy caused by two leucine-to-alanine mutations, L75A and L76A. For L75A, the prediction is 1.0 ± 0.6 kcal/mol (experimental 1.1 ± 0.1) and for L76A, 1.4 ± 0.5 kcal/mol (experimental 1.1 ± 0.1 ; Fleming et al. 1997). This accurate prediction suggests that both the simplified model environment and the simulation techniques faithfully describe the physics of GpA dimerization.

13.5.4 Permeation: Transport of Glycerol Through GlpF

Permeation through membrane channels is a key phenomenon in membrane biophysics, and a prime target for computational biophysicists: It is one of the problems most amenable to free energy simulations. Among the growing literature, we

Fig. 13.13 Cartoon rendering of a monomeric GlpF channel, with pore-lining residues as a *blue* surface, selectivity filter residues (Phe and Arg) and pore water as licorice, and glycerol molecule as van der Waals spheres. Note the change in water polarity at the NPA motif in the middle of the pore



should mention one pioneering study of membrane channel transport by free energy simulation: that of the transport of potassium ions by crystallographically sited activation (Kcsa), by Bernèche and Roux (2001). The system is arguably simpler than glycerol in the glycerol uptake facilitator (GlpF), in the sense that potassium ions have no internal degrees of freedom, yet it was made more complex by the interactions between potassium ions present in the selectivity filter simultaneously, resulting in concerted motion of two or three ions through the pore. Bernèche and Roux addressed this challenge by explicitly sampling pairs of ion coordinates within the umbrella sampling scheme (described in the section on estimators above), obtaining two-dimensional free energy maps on which they could describe possible conduction pathways.

13.5.4.1 The Question: Mode of Glycerol Diffusion in Passive Transport

GlpF is a water- and glycerol-conducting channel (aquaglyceroporin) found in the inner membrane of *Escherichia coli*. It adopts a classic aquaporin structure, a tetrameric assembly of TM α -helix bundles, forming four identical, monomeric pores. Each pore starts with a periplasmic vestibule, followed by a narrower section containing the selectivity filter (SF). Two half-membrane-spanning helices join at the Asn-Pro-Ala (NPA) region, critical for the orientational tuning of water that prevents the channel from conducting protons (Fig. 13.13).

The question we sought to address here was apparently simple: How can an aquaporin conduct glycerol? The behavior of water along the pore was understood, but how glycerol too could make its way along the particular pore of GlpF was a mystery. MD simulations are again an appropriate tool to study such a dynamic process in its minute, atomic detail.

The expected slow diffusion of glycerol along the pore, as well as the energetic barriers (particularly at the highly constricted selectivity filter), makes spontaneous sampling of the channel by glycerol unlikely. The first computational study of

glycerol transport through GlpF (Jensen et al. 2002) consisted in nonequilibrium pulling simulations (steered molecular dynamics or SMD), followed by reconstruction of the equilibrium free energy profile from the nonequilibrium work based on the Jarzynski equality (Jarzynski 1997). Such biased simulations have two aims. First, they allow for sampling of the full permeation process in a simulation, from which a precise, atomic picture can be extracted. Second, the free energy profile contains information on the distribution of glycerol along the channel and any bottlenecks, enthalpic or entropic. These, in turn, are key to understanding glycerol conduction.

Nonequilibrium pulling may prevent glycerol from relaxing and finding the optimal diffusion pathway. To obtain a less biased picture, we enhanced glycerol diffusion through GlpF in a near-equilibrium fashion, using the ABF approach described in the section 13.5.3.2 above (Hénin et al. 2008).

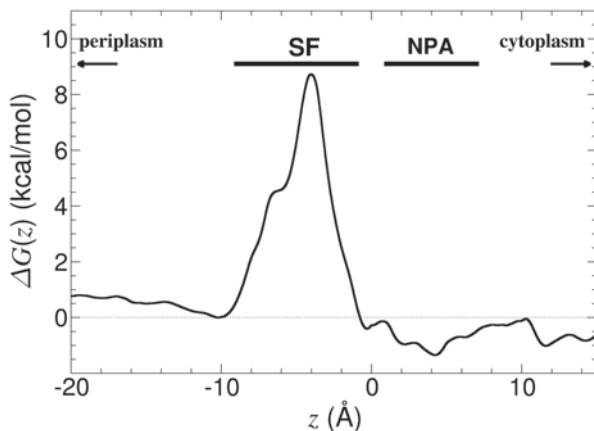
13.5.4.2 Technical Approach and Pitfalls

The tetramer of GlpF was modeled with one glycerol molecule in each monomeric pore. The ABF method was used to enhance diffusion of each glycerol molecule separately, yielding four virtually independent simulations. Convergence of the calculation proved slower than expected for such a small molecule following a seemingly simple and well-defined pathway.

Defining the reaction coordinate unambiguously proved difficult, due to the flexibility of pore-lining residues, including side chains forming the selectivity filter. The initial definition was the vertical position relative to the center of the complete GlpF tetramer. Given the resolution of the PMF and the peaked shape of the free energy barrier, vertical fluctuation on the ångström scale is sufficient to introduce a large local error in the computed free energy profile, hence in the adaptive bias, preventing efficient barrier crossing. The solution was to switch to a local coordinate, defined directly with respect to the position of selectivity filter residues. This is a general problem when describing interactions of a molecule with a “soft” partner such as the surface of a protein, that of a membrane, or a pore lined with flexible residues. Our intuitive representation of the process combines both the global position of the moving particle and its interactions with local features—it covers a broad range but includes local detail. A mathematically defined coordinate, however, will be either local (and limited in range) or global (and lacking resolution). Alternatives to this unappealing choice are more involved, and include stitching together local coordinates, or using an explicitly multi-resolution approach.

In addition to the problem of defining the translocation coordinate precisely, convergence of the measured free energy profile was slowed down by the slow relaxation of parameters other than that coordinate. Such so-called orthogonal degrees of freedom were mostly the reorientation and conformation changes of glycerol, which are fast in bulk water, but drastically slowed down within the confined environment of the pore.

Fig. 13.14 Free energy profile for glycerol conduction through GlpF. *SF* selectivity filter



13.5.4.3 Results: Smooth Free Energy Profile, Coupled Degrees of Freedom

The free energy profile previously computed from nonequilibrium steered simulations was rough: It had several minima separated by large barriers; the minima broadly matched the positions of crystallographic glycerol molecules that were used as starting coordinates for the pulling simulations. In contrast, our ABF simulations yielded the profile depicted in Fig. 13.14, mostly smooth and featuring a single barrier located at the tightly constricted selectivity filter.

Since the ABF data did not reflect crystallographic positions while SMD data did, it is tempting to conclude that the SMD profile is more accurate than the ABF one from the experimental perspective. However, the crystal structure describes a frozen channel where glycerol molecules are bound to occupy lowest-energy positions: these may not be minima in the room-temperature free energy profile. In the SMD simulation, one-way pulling from the crystal structure may have produced spurious minima due to the contribution of irreversible work.

The ABF simulations reveal that several degrees of freedom other than the vertical position come into play during permeation, and are mutually correlated.

Correlations between progress coordinate z and other coordinates are illustrated in Fig. 13.15.

Glycerol conformation: in bulk solution, glycerol adopts two main conformations (*anti-anti* and *gauche-anti*). In the selectivity filter, these conformations become very unlikely, while the otherwise disfavored *gauche-gauche* form becomes strongly favored.

Glycerol orientation: glycerol tumbles freely in the wider sections of the pore, as evidenced by the zero average dipole moment along z . But it is confined to a specific orientation when crossing the SF, hence a loss of entropy and a free energy penalty. In addition, the inversion of the local electric field at the NPA motif leads to a reversal of the dipole moment of glycerol. The total dipole moment fluctuates as the environment becomes more or less polarizing.

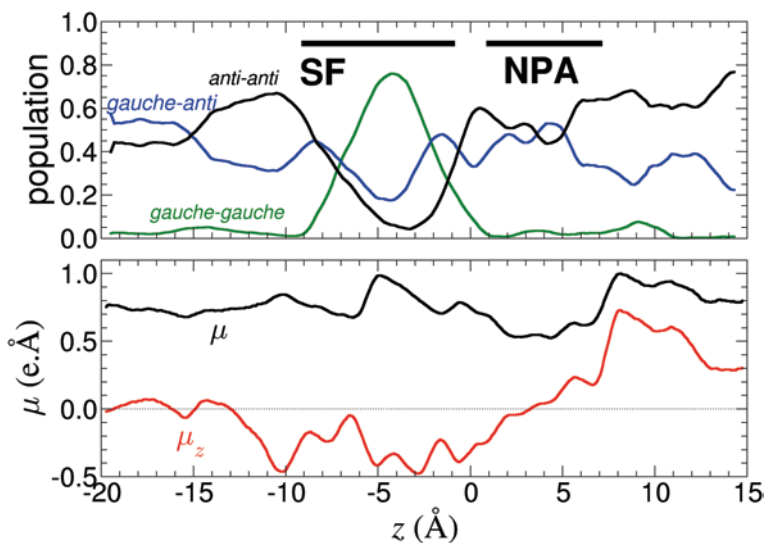


Fig. 13.15 Conformational and orientational “induced fit” of glycerol to the GlpF pore environment. *Up*: Populations of conformational states as a function of the coordinate z : *gauche-gauche* (green), *gauche-anti* (blue), and *anti-anti* (black). Conformations are labeled based on the state of the two O–C–O torsional angles. *Down*: Average molecular dipole moment of glycerol (black) and its projection along the z -direction (red)

The process of glycerol permeation through GlpF shows no timescale separation: Vertical diffusion, rotation, and conformational transitions occur on the timescale of tens of nanoseconds, violating the assumptions of many computational methods designed to accelerate slow processes.

13.5.5 Final Thoughts on Free Energy Simulations

In conclusion, we hope to have made it clear that free energy simulations are powerful techniques to turn molecular modeling into a predictive, quantitative tool. Yet they are by no means mature, black box methods: Caution and expertise are required in their design, execution, and interpretation. Their accuracy is fundamentally dependent on two factors: the accuracy of the underlying model and the statistical convergence of the numerical estimators.

13.6 Concluding Remarks

Molecular modeling is nowadays widely used to complement experimental studies on MPs and provides atomic scale insight into the mechanisms and dynamics of these systems. Selected examples in this review illustrate how these computational

methods can establish a dialogue with experiment. This may be seen as a premise to the next big step in structural biology: integration of all available and computable data to establish the most accurate picture of the structure, function, and dynamics of the molecular machinery of the cell and in particular MPs.

Acknowledgments M.B. thanks the French Agency for Research for funding (Grant ANR-07-CIS7-003-01). This work was supported by the “Initiative d’Excellence” program from the French State (Grant “DYNAMO”, ANR-11-LABX-0011). Benoist Laurent and Samuel Murail are acknowledged for contribution to several figures.

References

- Abin-Carriquiry JA, Zunini MP, Cassels BK, Wonnacott S, Dajas F (2010) In silico characterization of cytisinoids docked into an acetylcholine binding protein. *Bioorg Med Chem Lett* 20(12):3683–3687
- Adriouch S, Bannas P, Schwarz N, Fliegert R, Guse AH, Seman M, Haag F, Koch-Nolte F (2008) ADP-ribosylation at R125 gates the P2X7 ion channel by presenting a covalent ligand to its nucleotide binding site. *FASEB J* 22(3):861–869
- Alberts IL, Nadassy K, Wodak SJ (1998) Analysis of zinc binding sites in protein crystal structures. *Protein Sci* 7:1700–1716
- Armstrong N, Gouaux E (2000) Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron* 28:165–181
- Artali R, Bombieri G, Meneghetti F (2005) Docking of 6-chloropyridazin-3-yl derivatives active on nicotinic acetylcholine receptors into molluscan acetylcholine binding protein (AChBP). *Farmaco* 60(4):313–320
- Aschrafi A, Sadtler S, Niculescu C, Rettinger J, Schmalzing G (2004) Trimeric architecture of homomeric P2X2 and heteromeric P2X1+2 receptor subtypes. *J Mol Biol* 342:333–343
- Baaden M, Lavery R (2007) There’s plenty of room in the middle: multi-scale modelling of biological systems. In: de Brevern AG (ed) *Recent advances in protein engineering*. Research signpost, India, pp 173–195
- Babakhani A, Talley TT, Taylor P, McCammon JA (2009) A virtual screening study of the acetylcholine binding protein using a relaxed-complex approach. *Comput Biol Chem* 33(2):160–170
- Bahar I (1999) Dynamics of proteins and biomolecular complexes: inferring functional motions from structure. *Rev Chem Eng* 15:319–347
- Bahar I, Rader AJ (2005) Coarse-grained normal mode analysis in structural biology. *Curr Opin Struct Biol* 15:586–592
- Barrera NP, Ormond SJ, Henderson RM, Murrell-Lagnado RD, Edwardson JM (2005) Atomic force microscopy imaging demonstrates that P2X2 receptors are trimers but that P2X6 receptor subunits do not oligomerize. *J Biol Chem* 280:10759–10765
- Bernèche S, Roux B (2001) Energetics of ion conduction through the K⁺ channel. *Nature* 414:73–77
- Bocquet N, Prado deCL, Cartaud J, Neyton J, Le Poupon C, Taly A, Grutter T, Changeux J-P, Corringer P-J (2007) A prokaryotic proton-gated ion channel from the nicotinic acetylcholine receptor family. *Nature* 445(7123):116–119
- Bocquet N, Nury H, Baaden M, Le Poupon C, Changeux JP, Delarue M, Corringer PJ (2009) X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* 457(7225):111–114
- Bourne Y, Talley TT, Hansen SB, Taylor P, Marchot P (2005) Crystal structure of a Cbtx-AChBP complex reveals essential interactions between snake alpha-neurotoxins and nicotinic receptors. *EMBO J* 24(8):1512–1522

- Brannigan G, LeBard DN, Hénin J, Eckenhoff RG, Klein ML (2010) Multiple binding sites for the general anesthetic isoflurane identified in the nicotinic acetylcholine receptor transmembrane domain. *Proc Natl Acad Sci U S A* 107:14122–14127
- Brejč K, van Dijk WJ, Klaassen RV, Schuurmans M, van Der OJ, Smit AB, Sixma TK (2001) Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* 411(6835):269–276
- Brown HA (2012) Lipidomics: when apocrypha becomes canonical. *Curr Opin Chem Biol* 16(1–2):221–226. doi:10.1016/j.cbpa.2012.02.003
- Cao L, Young MT, Broomhead HE, Fountain SJ, North RA (2007) Thr339-to-serine substitution in rat P2X2 receptor second transmembrane domain causes constitutive opening and indicates a gating role for Lys308. *J Neurosci* 27:12916–12923
- Celie PH, van Rossum-Fikkert SE, van Dijk WJ, Brejč K, Smit AB, Sixma TK (2004) Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* 41(6):907–914
- Celie PH, Kasheverov IE, Mordvintsev DY, Hogg RC, van Nierop P, van Elk R, van Rossum-Fikkert SE, Zhmak MN, Bertrand D, Tsetlin V, Sixma TK, Smit AB (2005) Crystal structure of nicotinic acetylcholine receptor homolog AChBP in complex with an alpha-conotoxin PnIA variant. *Nat Struct Mol Biol* 12(7):582–588
- Changeux JP, Taly A (2008) Nicotinic receptors, allosteric proteins and medicine. *Trends Mol Med* 14(3):93–102
- Chen J, Brooks CL 3rd (2007) Can molecular dynamics simulations provide high-resolution refinement of protein structure? *Proteins* 67(4):922–930
- Chipot C, Pohorille A (eds) (2007) *Free energy calculations: theory and applications in chemistry and biology*. Springer, New York
- Chothia C, Lesk AM (1986) The relation between the divergence of sequence and structure in proteins. *EMBO J* 5(4):823–826
- Conn PJ, Christopoulos A, Lindsley CW (2009) Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. *Nat Rev Drug Discov* 8(1):41–54. doi:10.1038/nrd2760
- Corda D, Di Girolamo M (2003) Functional aspects of protein mono-ADP-ribosylation. *EMBO J* 22(9):1953–1958
- Corringer PJ, Baaden M, Bocquet N, Delarue M, Dufresne V, Nury H, Prevost M, Van Renterghem C (2010) Atomic structure and dynamics of pentameric ligand-gated ion channels: new insight from bacterial homologues. *J Physiol* 588(4):565–572
- Costa V, Nistri A, Cavalli A, Carloni P (2003) A structural model of agonist binding to the alpha-3beta4 neuronal nicotinic receptor. *Br J Pharmacol* 140(5):921–931
- Cox K, Bond PJ, Grottes A, Baaden M, Sansom MSP (2007) Outer membrane proteins: comparing X-Ray and NMR Structures by MD simulations in lipid bilayers. *Eur Biophys J* 37(2):131–141
- Darve E, Pohorille A (2001) Calculating free energies using average force. *J Chem Phys* 115:9169–9183
- Duret G, Van Renterghem C, Weng Y, Prevost M, Moraga-Cid G, Huon C, Sonner JM, Corringer PJ (2011) Functional prokaryotic-eukaryotic chimera from the pentameric ligand-gated ion channel family. *Proc Natl Acad Sci U S A* 108:12143–12148
- Fan H, Mark AE (2004) Refinement of homology-based protein structures by molecular dynamics simulation techniques. *Protein Sci* 13(1):211–220
- Faraldo-Gomez JD, Forrest LR, Baaden M, Bond PJ, Domene C, Patargias G, Cuthbertson J, Sansom MSP (2004) Conformational sampling and dynamics of membrane proteins from 10-nano-second computer simulations. *Proteins* 57(4):783–791
- Fleming KG, Ackerman AL, Engelman DM (1997) The effect of point mutations on the free energy of transmembrane alpha-helix dimerization. *J Mol Biol* 272:266–275
- Gao F, Bern N, Little A, Wang HL, Hansen SB, Talley TT, Taylor P, Sine SM (2003) Curari-form antagonists bind in different orientations to acetylcholine-binding protein. *J Biol Chem* 278(25):23020–23026

- Grossfield A (2011) Recent progress in the study of G protein-coupled receptors with molecular dynamics computer simulations. *Biochim Biophys Acta* 1808(7):1868–1878. doi:10.1016/j.bbame.2011.03.010
- Grossfield A, Feller SE, Pitman MC (2007) Convergence of molecular dynamics simulations of membrane proteins. *Proteins* 67(1):31–40. doi:10.1002/prot.21308
- Grutter T, Le Novère N, Changeux JP (2004) Rational understanding of nicotinic receptors drug binding. *Curr Top Med Chem* 4(6):645–650
- Gumbart J, Chipot C, Schulten K (2011) Free-energy cost for translocon-assisted insertion of membrane proteins. *Proc Natl Acad Sci U S A* 108:3596–3601
- Haddadian EJ, Cheng MH, Coalson RD, Xu Y, Tang P (2008) In silico models for the human α 4 β 2 nicotinic acetylcholine receptor. *J Phys Chem B* 112(44):13981–13990
- Hansen SB, Sulzenbacher G, Huxford T, Marchot P, Taylor P, Bourne Y (2005) Structures of *Aplysia* AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *EMBO J* 24(20):3635–3646
- Henchman RH, Wang HL, Sine SM, Taylor P, McCammon JA (2005) Ligand-induced conformational change in the α 7 nicotinic receptor ligand binding domain. *Biophys J* 88(4):2564–2576
- Hénin J, Chipot C (2004) Overcoming free energy barriers using unconstrained molecular dynamics simulations. *J Chem Phys* 121:2904–2914
- Hénin J, Pohorille A, Chipot C (2005) Insights into the recognition and association of transmembrane α -helices The free energy of α -helix dimerization in glycophorin A. *J Am Chem Soc* 127:8478–8484
- Hénin J, Tajkhorshid E, Schulten K, Chipot C (2008) Diffusion of glycerol through *Escherichia coli* aquaglyceroporin GlpF. *Biophys J* 94:832–839
- Hénin J, Fiorin G, Chipot C, Klein ML (2010) Exploring multidimensional free energy landscapes using time-dependent biases on collective variables. *J Chem Theory Comput* 6:35–47
- Hilf RJ, Dutzler R (2008) X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* 452(7185):375–379
- Hilf RJ, Dutzler R (2009) Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* 457(7225):115–118
- Hilf RJ, Bertozzi C, Zimmermann I, Reiter A, Trauner D, Dutzler R (2010) Structural basis of open channel block in a prokaryotic pentameric ligand-gated ion channel. *Nat Struct Mol Biol* 17(11):1330–1336
- Hinsen K (1998) Analysis of domain motions by approximate normal mode calculations. *Proteins* 33(3):417–429
- Hogner A, Kastrop JS, Jin R, Liljefors T, Mayer ML, Egebjerg J, Larsen IK, Gouaux E (2002) Structural basis for AMPA receptor activation and ligand selectivity: crystal structures of five agonist complexes with the GluR2 ligand-binding core. *J Mol Biol* 322:93
- Howard RJ, Murail S, Ondricek KE, Corringer PJ, Lindahl E, Trudell JR, Harris RA (2011) Structural basis for alcohol modulation of a pentameric ligand-gated ion channel. *Proc Natl Acad Sci U S A* 108:12149–12154
- Huang X, Zheng F, Stokes C, Papke RL, Zhan CG (2008a) Modeling binding modes of α 7 nicotinic acetylcholine receptor with ligands: the roles of Gln117 and other residues of the receptor in agonist binding. *J Med Chem* 51(20):6293–6302
- Huang X, Zheng F, Zhan CG (2008b) Modeling differential binding of α 4 β 2 nicotinic acetylcholine receptor with agonists and antagonists. *J Am Chem Soc* 130(49):16691–16696
- Iorga B, Herlem D, Barre E, Guillou C (2006) Acetylcholine nicotinic receptors: finding the putative binding site of allosteric modulators using the “blind docking” approach. *J Mol Model* 12(3):366–372
- Ito M, Xu H, Guffanti AA, Wei Y, Zvi L, Clapham DE, Krulwich TA (2004) The voltage-gated Na^+ channel NaVBP has a role in motility, chemotaxis, and pH homeostasis of an alkaliphilic *Bacillus*. *Proc Natl Acad Sci U S A* 101(29):10566–10571
- Ivetac A, Sansom MS (2008) Molecular dynamics simulations and membrane protein structure quality. *Eur Biophys J* 37(4):403–409

- Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78:2690–2693
- Jensen MØ, Park S, Tajkhorshid E, Schulten K (2002) Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc Natl Acad Sci U S A* 99:6731–6736
- Jiang LH, Kim M, Spelta V, Bo X, Surprenant A, North RA (2003) Subunit arrangement in P2X receptors. *J Neurosci* 23:8903
- Jiang R, Lemoine D, Martz A, Taly A, Gonin S, Prado de Carvalho L, Specht A, Grutter T (2011) Agonist trapped in ATP-binding sites of the P2X2 receptor. *Proc Natl Acad Sci U S A* 108(22):9066–9071
- Jiang R, Taly A, Lemoine D, Martz A, Cunrath O, Grutter T (2012) Tightening of the ATP-binding sites induces the opening of P2X receptor channels. *EMBO J* 31(9):2134–2143
- Jin R, Banke TG, Mayer ML, Traynelis S, Gouaux E (2003) Structural basis for partial agonist action at ionotropic glutamate receptors. *Nat Neurosci* 6:803
- Kannan S, Zacharias M (2010) Application of biasing-potential replica-exchange simulations for loop modeling and refinement of proteins in explicit solvent. *Proteins* 78(13):2809–2819. doi:10.1002/prot.22796
- Karakas E, Simorowski N, Furukawa H (2009) Structure of the zinc-bound amino-terminal domain of the NMDA receptor NR2B subunit. *EMBO J* 28(24):3910–3920. doi:10.1038/emboj.2009.338
- Kawate T, Michel JC, Birdsong WT, Gouaux E (2009) Crystal structure of the ATP-gated P2X(4) ion channel in the closed state. *Nature* 460:592
- Konstantakaki M, Changeux J, Taly A (2007) Docking of long chain alpha-cobratoxin suggests a basal state conformation of the nicotinic receptor. *Biochem Biophys Res Commun* 359(3):413–418
- Krebs WG, Alexandrov V, Wilson CA, Echols N, Yu H, Gerstein M (2002) Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic. *Proteins* 48(4):682–695
- Krissinel E, Henrick K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr.* 60(Pt 12 Pt 1):2256–68
- Laurent B, Murail S, Da Silva F, Corringer PJ, Baaden M (2013) Modeling complex biological systems: from solution chemistry to membranes and channels. *Pure Appl Chem* 85:1–13.
- Law RJ, Capener C, Baaden M, Bond PJ, Campbell J, Patargias G, Arinaminpathy Y, Sansom MSP (2005a) Membrane protein structure quality in molecular dynamics simulation. *J Mol Graph Model* 24(2):157–165
- Law RJ, Henchman RH, McCammon JA (2005b) A gating mechanism proposed from a simulation of a human alpha7 nicotinic acetylcholine receptor. *Proc Natl Acad Sci U S A* 102(19):6813–6818
- Le Novère N, Grutter T, Changeux JP (2002) Models of the extracellular domain of the nicotinic receptors and of agonist- and Ca²⁺-binding sites. *Proc Natl Acad Sci U S A* 99(5):3210–3215
- LeBard DN, Hémin J, Eckenhoff RG, Klein ML, Brannigan G (2012) General anesthetics predicted to block the GLIC Pore with micromolar affinity. *PLoS Comput Biol* 8:e1002532
- Lemoine D, Jiang R, Taly A, Chataigneau T, Specht A, Grutter T (2012) Ligand-gated Ion channels: new insights into neurological disorders and ligand recognition. *Chem Rev* 112(12):6285–6318
- Lörinczi É, Bhargava Y, Marino SF, Taly A, Kaczmarek-Hájek K, Barrantes-Freer A, Dutertre S, Grutter T, Rettinger J, Nicke A (2012) Involvement of the cysteine-rich head domain in activation and desensitization of the P2X1 receptor. *Proc Natl Acad Sci U S A* 109(28):11396–11401
- Luttmann E, Ludwig J, Hoffle-Maas A, Samochocki M, Maelicke A, Fels G (2009) Structural model for the binding sites of allosterically potentiating ligands on nicotinic acetylcholine receptors. *ChemMedChem* 4(11):1874–1882
- MacKenzie KR, Prestegard JH, Engelman DM (1997) A transmembrane helix dimer: structure and implications. *Science* 276:131–133
- Mayer ML (2005) Crystal structures of the GluR5 and GluR6 ligand binding cores: molecular mechanisms underlying kainate receptor selectivity. *Neuron* 45(4):539–552

- Mayer ML, Ghosal A, Dolman NP, Jane DE (2006) Crystal structures of the kainate receptor GluR5 ligand binding core dimer with novel GluR5-selective antagonists. *J Neurosci* 26(11):2852–2861
- Ming D, Wall ME (2005) Allostery in a coarse-grained model of protein dynamics. *Phys Rev Lett* 95(19):198103
- Mirjalili V, Feig M (2013) Protein structure refinement through structure selection and averaging from molecular dynamics ensembles. *J Chem Theory Comput* 9(2):1294–1303
- Mnatsakanyan N, Jansen M (2013) Experimental determination of the vertical alignment between the second and third transmembrane segments of muscle nicotinic acetylcholine receptors. *J Neurochem* 125(6):843–854
- Nagaya N, Tittle RK, Saar N, Dellal SS, Hume RI (2005) An intersubunit zinc binding site in rat P2X2 receptors. *J Biol Chem* 280(28):25982–25993
- Nicke A, Bäumert HG, Rettinger J, Eichele A, Lambrecht G, Mutschler E, Schmalzing G (1998) P2X1 and P2X3 receptors form stable trimers: a novel structural motif of ligand-gated ion channels. *EMBO J* 17(11):3016–3028
- Nury H, Poitevin F, Van Renterghem C, Changeux JP, Corringer PJ, Delarue M, Baaden M (2010) One-microsecond molecular dynamics simulation of channel gating in a nicotinic receptor homologue. *Proc Natl Acad Sci U S A* 107:6275–6280
- Nury H, Van Renterghem C, Weng Y, Tran A, Baaden M, Dufresne V, Changeux JP, Sonner JM, Delarue M, Corringer PJ (2011) X-ray structures of general anaesthetics bound to a pentameric ligand-gated ion channel. *Nature* 469(7330):428–431
- Olivella M, Gonzalez A, Pardo L, Deupi X (2013) Relation between sequence and structure in membrane proteins. *Bioinformatics* 29(13):1589–1592. doi:10.1093/bioinformatics/btt249
- Piggot TJ, Holdbrook DA, Khalid S (2013) Conformational dynamics and membrane interactions of the *E. coli* outer membrane protein FecA: a molecular dynamics simulation study. *Biochim Biophys Acta* 1828(2):284–293. doi:10.1016/j.bbame.2012.08.021
- Prevost M, Sauguet L, Nury H, Van Renterghem C, Huon C, Poitevin F, Baaden M, Delarue M, Corringer PJ (2012) A novel locally closed conformation of a bacterial pentameric proton-gated Ion channel. *Nat Struct Mol Biol* 19(6):642–649. doi:10.1038/nsmb.2307
- Raval A, Piana S, Eastwood MP, Dror RO, Shaw DE (2012) Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins* 80(8):2071–2079. doi:10.1002/prot.24098
- Resat H, Mezei M (1996) Grand canonical ensemble Monte Carlo simulation of the dCpG/proflavine crystal hydrate. *Biophys J* 71(3):1179–1190
- Romo TD, Grossfield A (2011) Block covariance overlap method and convergence in molecular dynamics simulation. *J Chem Theory Comput* 7(8):2464–2472. doi:10.1021/ct2002754
- Šali A (1995) Modelling mutations and homologous proteins. *Curr Opin Biotechnol* 6(4):437–451
- Samson AO, Levitt M (2008) Inhibition mechanism of the acetylcholine receptor by alpha-neurotoxins as revealed by normal-mode dynamics. *Biochemistry* 47:4065–4070
- Sauguet L, Howard RJ, Malherbe L, Lee US, Corringer PJ, Harris RA, Delarue M (2013a) Structural basis for potentiation by alcohols and anaesthetics in a ligand-gated ion channel. *Nat Commun* 4:1697. doi:10.1038/ncomms2682
- Sauguet L, Poitevin F, Murail S, Van Renterghem C, Moraga-Cid G, Malherbe L, Thompson AW, Koehl P, Corringer PJ, Baaden M, Delarue M (2013b) Structural basis for ion permeation mechanism in pentameric ligand-gated ion channels. *EMBO J* 32(5):728–741. doi:10.1038/emboj.2013.17
- Schapira M, Abagyan R, Totrov M (2002) Structural model of nicotinic acetylcholine receptor iso-types bound to acetylcholine and nicotine. *BMC Struct Biol* 2:1. doi:10.1186/1472-6807-2-1
- Shevchenko A, Simons K (2010) Lipidomics: coming to grips with lipid diversity. *Nat Rev Mol Cell Biol* 11(8):593–598. doi:10.1038/nrm2934
- Shoichet BK, Kobilka BK (2012) Structure-based drug screening for G-protein-coupled receptors. *Trends Pharmacol Sci* 33(5):268–272
- Sine SM (2002) The nicotinic receptor ligand binding domain. *J Neurobiol* 53(4):431–446

- Sine SM, Engel AG (2006) Recent advances in Cys-loop receptor structure and function. *Nature* 440:448–455
- Sine SM, Wang HL, Gao F (2004) Toward atomic-scale understanding of ligand recognition in the muscle nicotinic receptor. *Curr Med Chem* 11(5):559–567
- Sixma TK, Smit AB (2003) Acetylcholine binding protein (AChBP): a secreted glial protein that provides a high-resolution model for the extracellular domain of pentameric ligand-gated ion channels. *Annu Rev Biophys Biomol Struct* 32:311–334
- Slavov SH, Radzvilovits M, LeFrancois S, Stoyanova-Slavova IB, Soti F, Kem WR, Katritzky AR (2010) A computational study of the binding of 3-(arylidene) anabaseines to two major brain nicotinic acetylcholine receptors and to the acetylcholine binding protein. *Eur J Med Chem* 45(6):2433–2446
- Smit AB, Syed NI, Schaap D, van Minnen J, Klumperman J, Kits KS, Lodder H, van der Schors RC, van Elk R, Sorgedragter B, Brejc K, Sixma TK, Geraerts WP (2001) A glia-derived acetylcholine-binding protein that modulates synaptic transmission. *Nature* 411(6835):261–268
- Sobolevsky AI, Rosconi MP, Gouaux E (2009) X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature* 462(7274):745–756
- Spurny R, Billen B, Howard RJ, Brams M, Debaveye S, Price KL, Weston DA, Strelkov SV, Tytgat J, Bertrand S, Bertrand D, Lummis SC, Ulens C. (2013) Multisite binding of a general anesthetic to the prokaryotic pentameric *Erwinia chrysanthemi* ligand-gated ion channel (ELIC). *J Biol Chem*. 288(12):8355-64
- Taly A (2007) Opened by a twist: a gating mechanism for the nicotinic acetylcholine receptor. *Eur Biophys J* 36(8):911–918
- Taly A (2013) Novel approaches to drug design for the treatment of schizophrenia. *Expert Opin Drug Discov* 8(10):1285–1296. doi:10.1517/17460441.2013.821108
- Taly A, Changeux JP (2008) Functional organization and conformational dynamics of the nicotinic receptor: a plausible structural interpretation of myasthenic mutations. *Ann N Y Acad Sci* 1132:42–52
- Taly A, Delarue M, Grutter T, Nilges M, Le NN, Corringer PJ, Changeux JP (2005) Normal mode analysis suggests a quaternary twist model for the nicotinic receptor gating mechanism. *Biophys J* 88(6):3954–3965
- Taly A, Corringer PJ, Grutter T, Prado deCL, Karplus M, Changeux JP (2006) Implications of the quaternary twist allosteric model for the physiology and pathology of nicotinic acetylcholine receptors. *Proc Natl Acad Sci U S A* 103(45):16965–16970
- Tanner NK, Cordin O, Banroques J, Doère M, Linder P (2003) The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol Cell* 11(1):127–138
- Tasneem A, Iyer LM, Jakobsson E, Aravind L (2005) Identification of the prokaryotic ligand-gated ion channels and their implications for the mechanisms and origins of animal Cys-loop ion channels. *Genome Biol* 6(1):R4
- Tirion MM (1996) Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys Rev Lett* 77:1905–1908
- Tittle RK, Hume RI (2008) Opposite effects of zinc on human and rat P2X2 receptors. *J Neurosci* 28(44):11131–11140
- Tittle RK, Power JM, Hume RI (2007) A histidine scan to probe the flexibility of the rat P2X2 receptor zinc-binding site. *J Biol Chem* 282(27):19526–19533
- Toshima K, Kanaoka S, Yamada A, Tarumoto K, Akamatsu M, Sattelle DB, Matsuda K (2009) Combined roles of loops C and D in the interactions of a neonicotinoid insecticide imidacloprid with the $\alpha 4\beta 2$ nicotinic acetylcholine receptor. *Neuropharmacology* 56(1):264–272
- Turabekova MA, Rasulev BF, Dzhakhangirov FN, Leszczynska D, Leszczynski J (2010) Aconitium and Delphinium alkaloids of curare-like activity. QSAR analysis and molecular docking of alkaloids into AChBP. *Eur J Med Chem* 45(9):3885–3894
- Ulens C, Akdemir A, Jongejan A, van Elk R, Edink E, Bertrand S, Perrakis A, Leurs R, Smit AB, Sixma TK, Bertrand D and de Esch IJP (2009) The use of acetylcholine binding protein in the search for novel $\alpha 7$ -nicotinic receptor ligands. In silico docking, pharmacological screening and X-ray analysis. *J. Med. Chem.* 52(8):2372-83

- Utsintong M, Talley TT, Taylor PW, Olson AJ, Vajragupta O (2009) Virtual screening against alpha-cobratoxin. *J Biomol Screen* 14(9):1109–1118. doi:10.1177/1087057109344617
- van Meer G, Voelker DR, Feigenson GW (2008) Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* 9(2):112–124. doi:10.1038/nrm2330
- Walker JE, Eberle A, Gay NJ, Runswick MJ, Saraste M (1982) Conservation of structure in proton-translocating ATPases of *Escherichia coli* and mitochondria. *EMBO J* 10(4):203–206
- Weng Y, Yang L, Corringier PJ, Sonner JM (2010) Anesthetic sensitivity of the *Gloeobacter violaceus* proton-gated ion channel. *Anesth Analg* 110(1):59–63
- Xu Y, Barrantes FJ, Luo X, Chen K, Shen J, Jiang H (2005) Conformational dynamics of the nicotinic acetylcholine receptor channel: a 35-ns molecular dynamics simulation study. *J Am Chem Soc* 127(4):1291–1299
- Yi M, Tjong H, Zhou HX (2008) Spontaneous conformational change and toxin binding in alpha7 acetylcholine receptor: insight into channel activation and inhibition. *Proc Natl Acad Sci U S A* 105(24):8280–8285. doi:10.1073/pnas.0710530105
- Zuckerman DM (2010) Statistical physics of biomolecules: an introduction. CRC Press, Salem
- Zuckerman DM (2011) Equilibrium sampling in biomolecular simulations. *Annu Rev Biophys* 40:41–62

Chapter 14

Structural Studies of TSPO, a Mitochondrial Membrane Protein

Jean-Jacques Lacapere, Soria Iatmanen-Harbi, Lucile Senicourt, Olivier Lequin, Piotr Tekely, Rudra N. Purusottam, Petra Hellwig, Sebastien Kriegel, Stephanie Ravaud, Céline Juillan-Binard, Eva Pebay-Peyroula and Vassilios Papadopoulos

14.1 Introduction

The 18-kDa translocator protein (TSPO) is a membrane protein (MP) that was previously named peripheral-type benzodiazepine receptor (PBR) because of the binding of diazepam, a well-known benzodiazepine, which was initially observed in the kidney (Papadopoulos et al. 2006). Numerous studies show that this MP is involved in various physiological functions such as the transport of cholesterol, which is a rate-limiting step in the synthesis of steroids and bile salts (Lacapere and Papadopoulos 2003). TSPO transports cholesterol through the external mitochondrial membrane and transfers it to the inner membrane with the assistance of the outer mitochondrial membrane voltage-dependent anion channel (VDAC) and ATPase family AAA domain-containing protein 3 (ATAD3 A), which is an integral MP of the inner mitochondrial membrane crossing the outer

J.-J. Lacapere (✉) · S. Iatmanen-Harbi · L. Senicourt · P. Tekely · R. N. Purusottam
BioMolecules Laboratory, UMR-CNRS 7203, Université Pierre et Marie Curie and Ecole
Normale Supérieure Paris, France
e-mail: jean-jacques.lacapere@upmc.fr

O. Lequin
BioMolecules Laboratory, UMR-CNRS 7203, Ecole Normale Supérieure,
Université Pierre et Marie Curie, Paris, France

P. Hellwig · S. Kriegel
Laboratory of Vibrational Spectroscopy and Electrochemistry of Biomolecules,
UMR-CNRS 7177, Université de Strasbourg, Strasbourg, France

S. Ravaud · C. Juillan-Binard · E. Pebay-Peyroula
Institute of Structural Biology, French National Centre for Scientific Research,
Centre for Nuclear Studies, Université Grenoble Alpes, Grenoble, France

V. Papadopoulos
The Research Institute of the McGill University Health Center, Department of Medicine,
McGill University, Montreal, QC, Canada

I. Mus-Veteau (ed.), *Membrane Proteins Production for Structural Analysis*,
DOI 10.1007/978-1-4939-0662-8_14, © Springer Science+Business Media New York 2014

membrane (Rone et al. 2012; Papadopoulos and Miller 2012). TSPO associates within different protein complexes that perform different physiological functions (Papadopoulos et al. 2006; Rone et al. 2012; Issop et al. 2013). TSPO is highly conserved from bacteria to humans (Fan et al. 2012), but the ancestral role of TSPO is unknown because bacteria do not have cholesterol and do not synthesize steroids. Recent work shows that bacterial TSPO can rapidly catalyze porphyrin degradation (Ginter et al. 2013), which is consistent with previous reports that TSPO can bind porphyrins (Verma et al. 1987; Wendler et al. 2003; Yeliseev and Kaplan 1999)

Few structural studies have been performed on TSPO, in contrast to the numerous physiological and physiopathological studies. Hydrophobicity plots show that all members of the TSPO family have five conserved hydrophobic domains (Fan et al. 2012). Initial analysis suggested a hemi-membrane insertion of the protein because of the small number of amino acids that were predicted to be involved in hydrophobic domains (Bernassau et al. 1993). Further topological analysis using epitope insertion clearly showed that TSPO has a five-transmembrane (TM) structure (Joseph-Liauzin et al. 1998). Synthetic TSPO peptides encompassing the five putative TM domains were studied by $^1\text{H-NMR}$ (nuclear magnetic resonance) and that revealed helical structures (Murail et al. 2008). The five helix-fold structure of TSPO was confirmed by circular dichroism (CD) of the entire protein (Jamin and Lacapere 2007; Murail et al. 2008) and also was observed in two-dimensional (2D) crystals of bacterial TSPO (Korkhov et al. 2010). The first atomic model of TSPO was based on the only available MP structure at the time, the G protein-coupled receptor (GPCR) family member protein bacteriorhodopsin (Bernassau et al. 1993). It is now well established that TSPO belongs to another MP family (Punta et al. 2012), but no structure of any member of this protein family has been obtained. Thus, the determination of the atomic structure of TSPO remains a challenge.

Because of its low natural abundance and association with various protein complexes (Papadopoulos et al. 2006; Rone et al. 2012; Issop et al. 2013), purification of large amounts of TSPO from native cells is a very challenging task. Therefore, overexpression of recombinant TSPO in heterologous cells was developed to produce the large amounts of protein required for structural and functional studies (Sprengel et al. 1989; Parola et al. 1991; Riond et al. 1991; Garnier et al. 1994; Joseph-Liauzin et al. 1998). There are currently many different approaches to obtain the 3D structures of MPs (Lacapere et al. 2007; Lacapere 2010). In this chapter, we present an overview of the structural results obtained for TSPO. A crucial step in structure determination is the stabilization of protein conformation. This can be achieved by different strategies such as protein engineering (Tate and Schertler 2009) and by adding specific ligands. TSPO has different ligands endowed with different functions, including ligands that are transported such as cholesterol or protoporphyrin, and ligands that activate or inhibit transport such as the isoquinoline PK 11195 (Papadopoulos et al. 2006).

14.2 Production, Purification, and Characterization of Recombinant TSPO

14.2.1 Overexpression System

Among the various overexpression systems for producing large quantities of recombinant proteins, bacteria are an easy and inexpensive system when the intracellular abundance of recombinant protein is low. However, heterologous expression systems in bacteria often result in recombinant protein accumulation in inclusion bodies (IBs). This can be prevented by the expression of a fusion protein that directs the insertion of the recombinant MP into the bacterial membrane. Early work on TSPO overexpression was performed using a construct fused to maltose-binding protein (Garnier et al. 1994). In some cases, the formation of IBs can be advantageous because the overexpressed protein will not affect bacterial growth (Mouillac and Baneres 2010).

14.2.2 Plasmid Choice

Expression plasmids that contain a poly-histidine tag are very efficient for protein purification using immobilized metal affinity chromatography (IMAC). The presence of a proteolytic cleavage site between the protein and the tag enables the tag to be removed so that it will not interfere with functional protein domains. The location of the tag at the N- or C-terminus is important, because the latter enables purification of full-length recombinant protein. However, the C-terminus of TSPO is involved in cholesterol binding; thus, a tag in this region alters TSPO function (Li and Papadopoulos 1998).

The first TSPO complementary DNA (cDNA) used for the expression of recombinant protein was that of rat (Sprengel et al. 1989), which was followed by other mammalian TSPO cDNAs such as that of bovine (Parola et al. 1991), human (Riond et al. 1991), and mouse (Garnier et al. 1994). Most of the data presented in this chapter were obtained with mouse TSPO (Fig. 14.1). It has to be reminded that there is a very high degree of identity between all mammalian TSPOs (Fan et al. 2012). Nonmammalian TSPOs, such as those of bacteria (Yeliseev and Kaplan 1995; Ginter et al. 2013) and plants (Guillaumot et al. 2009), also have been overexpressed.

14.2.3 Protein Extraction

When an overexpressed protein is incorporated into IBs, it can be extracted using strong denaturing agents such as urea or guanidinium chloride or strong detergents like sodium dodecyl sulfate (SDS). Protein extraction with SDS is an advantage

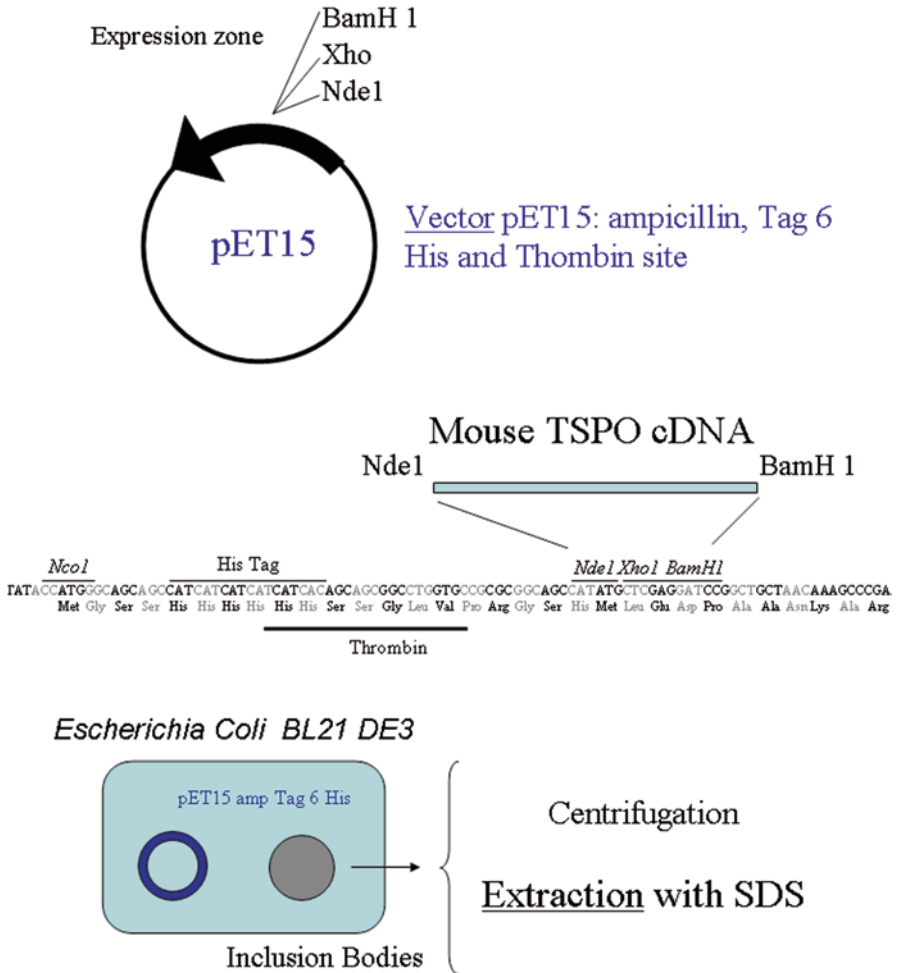


Fig. 14.1 *TSPO* overexpression. Mouse *TSPO* cDNA was inserted in a *pET15* vector containing T7 promoter, ampicillin resistance, six histidine tags and a thrombin cleavage site. Plasmid was transfected into *Escherichia coli BL21 DE3* bacteria. IPTG induction of *TSPO* overexpression leads to accumulation of *TSPO* in inclusion bodies, which were collected by centrifugation after bacterial disruption

for MPs, which are often purified with detergent. However, SDS also is used by biochemists for denaturation of proteins in SDS-PAGE.

IBs are storage repositories for waste in bacteria, and they contain high levels of nucleotides. Protein extraction with SDS yields a viscous solution that can be treated with benzonase to increase the fluidity of the solution. An absorption spectrum of this solution shows a peak at approximately 260 nm that is characteristic of nucleotides (Fig. 14.2).

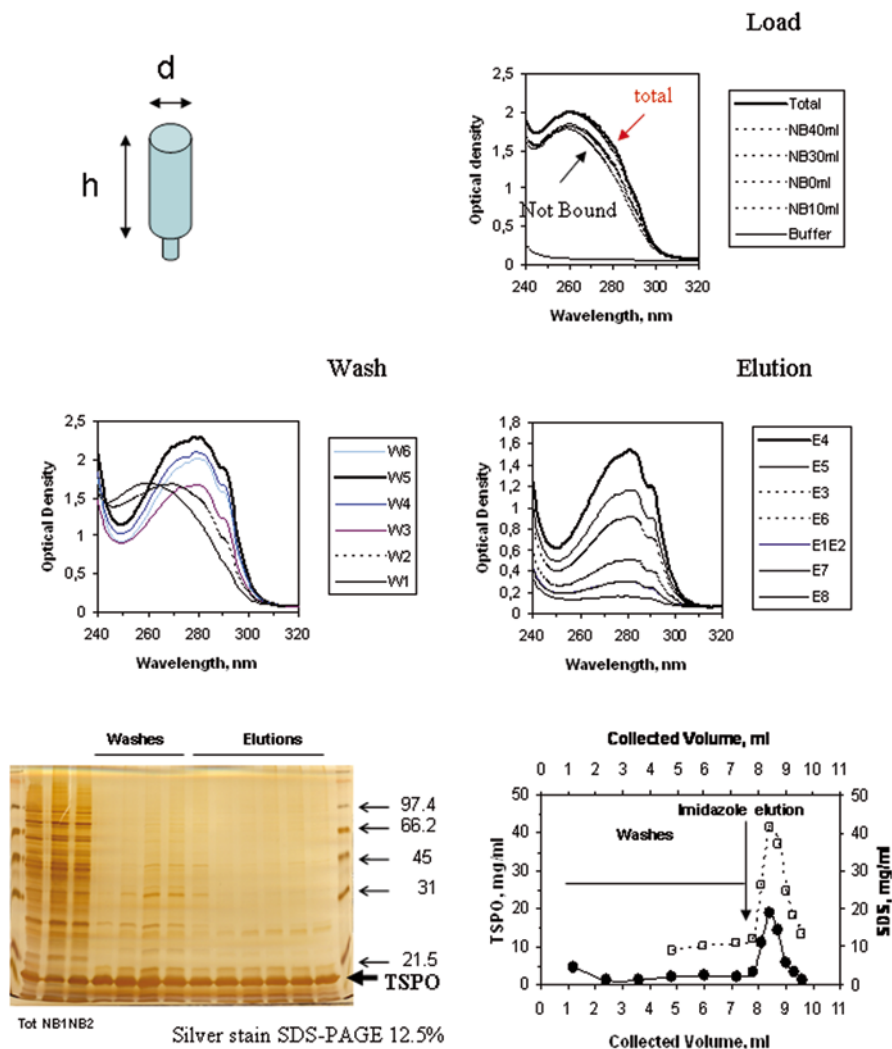


Fig. 14.2 TSPO purification. Solution of SDS-solubilized inclusion bodies from *E. coli* was layered on top of a column filled with NiNTA-agarose resin. Absorption spectra of loaded material (*total*) and collected fractions passing through (*not bound*) were first recorded (*load*). Solution containing SDS (CMC) and low-concentration imidazole (5 mM) were used to remove nontightly bound material, and spectra of the successive wash fractions were recorded (*wash*). Solution containing high-concentration imidazole (250 mM) was used to elute the bound TSPO, and the spectra and the successive collected fractions were recorded (*elution*). Silver-stained gel of the fractions (*load*, *wash*, and *elution*, 0.5 μg per well) shows that the collected fractions were highly pure with a major band below 21.5 kDa corresponding to TSPO. Purification chromatogram was obtained by converting optical densities of collected fractions into TSPO concentration using an extinction coefficient of $3.88 \text{ mg}^{-1} \text{ mL cm}^{-1}$. This coefficient was calculated for the recombinant mouse TSPO (sequence including the added amino acids from the plasmid) using the ProtParam tool of the ExpAsy server (Gasteiger et al. 2005)

14.2.4 *TSPO Purification*

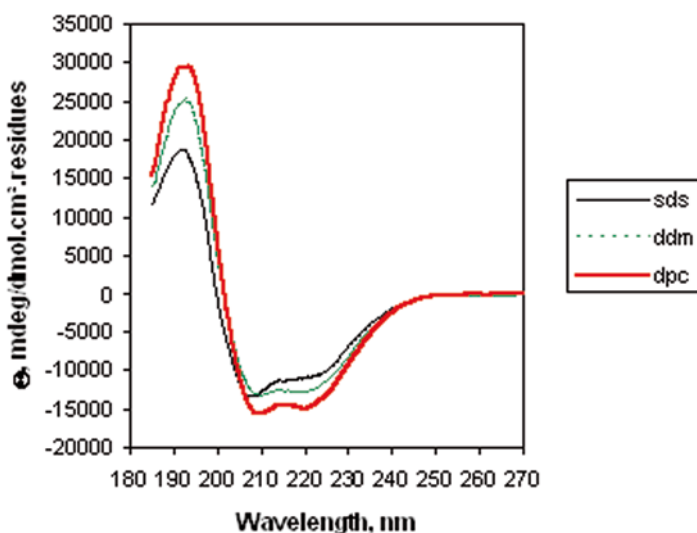
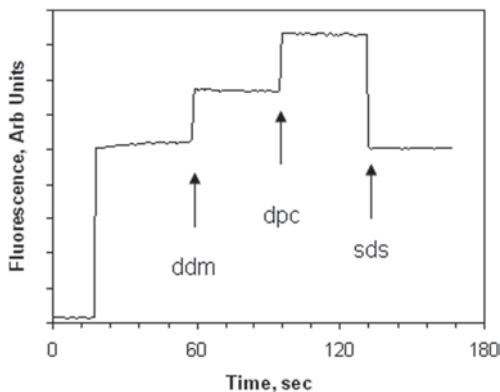
The solubilized extract from bacterial IBs was purified by IMAC using homemade columns that permit adaptation of the amount of resin used in the column to the amount of protein obtained from the bacterial culture. The geometry of the column (Fig. 14.2) is important to produce highly concentrated fractions. A small volume on a thin column will encounter protein all along the length of the column, and thereby yield a concentrated fraction. It is possible to follow the loading of the column by recording the absorption spectra of fractions of column eluates. Similarly, recording the spectra of the washing fractions revealed the change from nucleotide-rich fractions to protein-rich fractions (Fig. 14.2), which was observed by the change in peaks of maximal absorption. Ultimately, the fractions eluted with imidazole exhibit $OD_{280/250}$ absorbance ratios that are characteristic of protein. The presence of TSPO can be assessed by the silver-stained gel, which shows the unique band of TSPO in the elution fractions (Fig. 14.2). The protein content of each pure fraction can be determined by measuring the OD_{280} absorbance, which can be converted into TSPO concentration using an extinction coefficient of $3.88 \text{ mg}^{-1} \text{ mL cm}^{-1}$.

14.2.5 *Protein Characterization*

After MP has been purified, several parameters have to be measured before starting structural studies. The detergent content of each fraction can be measured as previously described (Ostuni et al. 2010) using dye spectroscopy (Fig. 14.2). In the washed fractions, the measured detergent corresponds to the SDS added (1 mg mL^{-1}) in the washing buffer, whereas the detergent concentration in the eluted fractions increases as the protein becomes more concentrated in the eluted fractions. The detergent associated with the protein can be measured in the eluted fractions and used to calculate a SDS:TSPO ratio of approximately 2 (w/w). Pooled fractions of purified TSPO can be analyzed using various biophysical techniques to check structural and functional characteristics. The intrinsic fluorescence of TSPO is highly dependent on its detergent environment (Fig. 14.3). Addition of dodecylmaltoside (DDM) induces a 40% increase in fluorescence, and a subsequent addition of dodecylphosphocholine (DPC) induces a further 40% increase in fluorescence. These changes in fluorescence are completely reversible by SDS at a sufficient concentration (Fig. 14.3). Intrinsic fluorescence is mostly due to the tryptophan environment, which changes from one detergent to another.

We suspected a change in TSPO secondary structure that could be evaluated by CD. Indeed, TSPO displays different spectra depending on the detergent (Fig. 14.4), which was previously described (Jamin and Lacapere 2007). Spectral deconvolution (Sreerama and Woody 2004) enabled an estimation of the secondary structure of TSPO in different detergents (Table in Fig. 14.4). The percentage of α -helix structure is increased when TSPO is in DDM or DPC compared to that in SDS. The presence of the high-affinity drug ligand PK 11195 also has an important effect

Fig. 14.3 TSPO intrinsic fluorescence. Signal was recorded with excitation and emission wavelengths of 290 and 330 nm, respectively. TSPO (0.5 μ M) was added in a buffer (Hepes pH 7.5) containing SDS (CMC). At arrows, DDM (2 CMC) was added first, followed by the addition of DPC (4 CMC), and finally an excess of SDS (5 CMC) was added



Detergent	% α -helix	% β -strand	% turn	% unordered
SDS	36 \pm 4	14 \pm 4	21 \pm 6	28 \pm 7
DDM	40 \pm 5	15 \pm 4	19 \pm 4	26 \pm 5
DPC	45 \pm 7	13 \pm 9	18 \pm 5	24 \pm 9

Fig. 14.4 TSPO circular dichroism spectra and secondary structure analysis. *Upper panel* shows spectra of TSPO (0.1 mg mL⁻¹) recorded with a 200- μ L cell containing phosphate buffer supplemented by 2 CMC of SDS, DDM, or DPC. *Lower panel* shows table of secondary structure analysis gained from deconvolution of the various CD spectra. (CDPro software, Sreerama and Woody 2004)

on TSPO secondary structure as described previously (Murail et al. 2008). If an increase in helical content is clearly observed using CD, and an impressive spreading is clearly observed in 2D NMR spectra (see below), nothing is observed using intrinsic fluorescence. This can be due to a compensatory effect of the fluorescence increase and decrease of the numerous tryptophans of TSPO. The β -sheet content is very difficult to estimate by CD, whereas it is better resolved by infrared spectroscopy (Goormaghtigh et al. 2009). The infrared absorption spectra of TSPO in DDM, SDS, and DPC show that the protein structure is dominated by a high content of α -helices found at $1,654\text{ cm}^{-1}$ (Fig. 14.5a–c). This is in agreement with CD data (Fig. 14.4). The fitting curves enable full secondary structure analysis to be performed (Table I in Fig. 14.5). The β -sheet and β -turn content was fitted at $1,620$, $1,630$, and $1,676\text{ cm}^{-1}$ and shows that it is far more affected by the change of detergent than α -helix content. In conclusion, TSPO may have distinct conformations of secondary structure in different detergents.

Large conformational changes in MPs can be studied by recording the exchange kinetics of hydrogen and deuterium (H/D) at the level of the amide proton in the mid-infrared spectral range (Vigano et al. 2004; Hielscher et al. 2011; Neehaul et al. 2013). Three different types of exchangeable residues can be distinguished. (1) Fast-exchange residues correspond to those that are located at the protein surface or within a channel. (2) Slow-exchange residues correspond to those that are less accessible and/or located in the hydrophobic part of the protein. (3) Nonexchangeable residues are typically found in the hydrophobic core of MPs. Amide hydrogen exchange of TSPO in DDM (Fig. 14.5d, e) with or without PK 11195 shows that proton accessibility is reduced by the presence of the ligand (Table II, Fig. 14.5). Nonexchangeable residues increase from 18 to 25%, whereas those in the hydrophobic part of the protein remain constant at 16–18% (slow-exchange residues), and fast-exchange residues greatly decrease from 70 to 57%. Exchange rates are significantly increased in the presence of PK 11195, which suggests that ligand binding generates a more compact structure for TSPO. These data suggest that PK 11195 binding induces a conformational change in TSPO, which is consistent with data gained from CD and NMR studies (Murail et al. 2008).

The measurement of stoichiometry and ligand-binding affinity with solubilized protein is a difficult task, particularly when using radioactive ligands to identify bound and free ligand. Isothermal titration calorimetry (ITC) is a good alternative because it measures the heat generated or absorbed when a ligand binds. We performed ITC experiments with TSPO solubilized in different detergents and observed that PK 11195 binding to DPC-solubilized TSPO generates heat with an apparent affinity of $0.6\text{ }\mu\text{M}$ (Fig. 14.6a). However, the stoichiometry is high (approximately 8–10), which suggests multiple binding sites on the protein or nonspecific binding in the detergent surrounding TSPO. No measurable heat was either generated or absorbed when TSPO was solubilized in SDS, which suggests that PK 11195 cannot bind to TSPO in this detergent. Similarly, no change in heat was observed after adding PK 11195 when TSPO was solubilized in DDM (Fig. 14.6b).

To characterize the folding of TSPO in different detergents, trypsin digestion was performed. Mouse recombinant TSPO contains 11 arginines and 2 lysines that

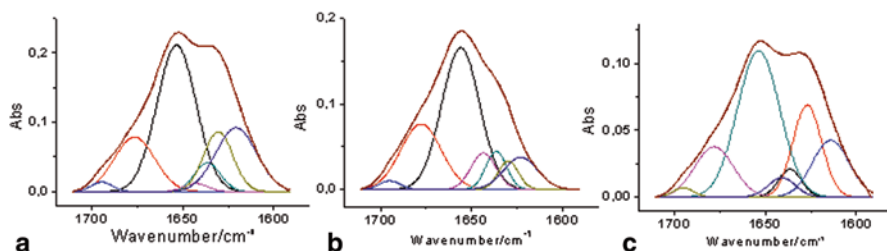


Table I

Detergent	% α -helix	% β -strand	% turn	% random/water
DDM	44 \pm 7	29 \pm 5	17 \pm 5	10 \pm 7
SDS	51 \pm 7	26 \pm 5	16 \pm 5	7 \pm 5
DPC	47 \pm 7	15 \pm 5	22 \pm 5	16 \pm 7

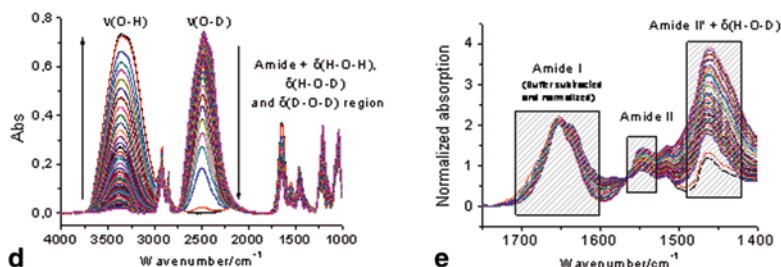


Table II

sample	a_1 %	τ_1 min	a_2 %	τ_2 min	a_3 %
DDM	70	5.5	16	53.9	18
DDM+PK 11195	57	8.3	18.3	70.8	24.8

Fig. 14.5 TSPO infrared spectra (a–c), secondary structure analysis of TSPO in DDM (a), in DPC (b), and in SDS (c). The α -helix structure is found at $1,654\text{ cm}^{-1}$; the β -sheet and β -turn are observed at $1,620$, $1,630$, and $1,676\text{ cm}^{-1}$. *Table I* summarizes values of secondary structure gained from deconvolution of spectra recorded for TSPO in the presence of different detergents. Typical time course of a $^1\text{H}/^2\text{H}$ (hydrogen/deuterium, H/D) exchange on TSPO (d). The initial spectra are characterized by an intense absorption band at $1,650\text{ cm}^{-1}$, called amide I band, which includes $\nu(\text{C}=\text{O})$; 70–85%) and $\nu(\text{C}-\text{N})$; 15–20%) vibrational modes. The amide II band is localized near $1,550\text{ cm}^{-1}$. On H/D exchange, the amide II band intensity decreases and a new band, called amide II', appears near $1,450\text{ cm}^{-1}$ (e). Data after buffer subtraction are used for kinetic analysis. *Table II* summarizes amplitudes and rates of H/D in the presence and absence of PK 11195

are the major cleavage sites for trypsin. All these residues are located in loops connecting TM domains and are therefore putatively accessible to proteolytic cleavage. However, Fig. 14.7 shows that TSPO in SDS is not digested after 120 min at 37°C with a trypsin to TSPO ratio of 1:30 (w/w), whereas TSPO in DPC or DDM is fully digested under similar experimental conditions. This indicates that TSPO loops are

Fig. 14.6 TSPO ligand binding in detergent was measured using isothermal titration calorimetry. TSPO (2 μ M) in DPC (a) or DDM (b) buffer was titrated with a solution of PK 11195 (0.2 mM)

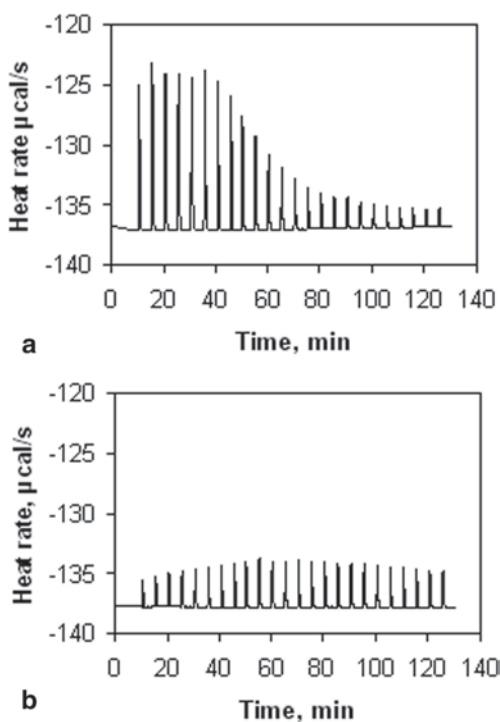
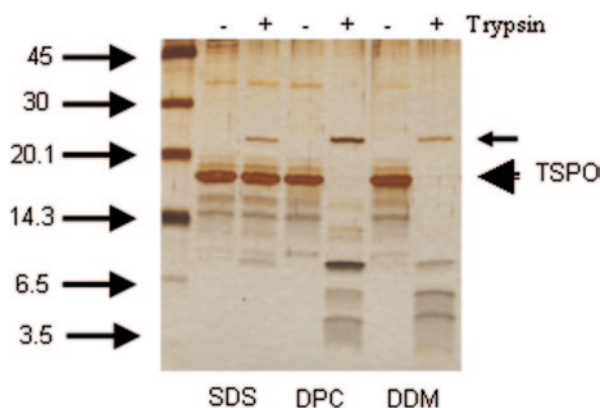


Fig. 14.7 TSPO trypsin digestion. TSPO was resuspended in 10 mM MOPS–Tris buffer pH 7 in the presence of CMC of SDS, DDM, or DPC, and incubated for 2 h at 37°C in the presence of trypsin at a TSPO-to-trypsin ratio of 30 (w/w)



not as accessible in the presence of SDS detergent and confirms that the global folding of TSPO is different in different detergents, as previously shown by CD and fluorescence studies. The binding of PK 11195 occurs in between cytosolic loops (Li and Papadopoulos 1998), and so the loops have to be accessible for binding. We showed that PK 11195 does not bind to TSPO in SDS, whereas it does bind to TSPO in DPC, which is in agreement with the trypsin digestion experiments.

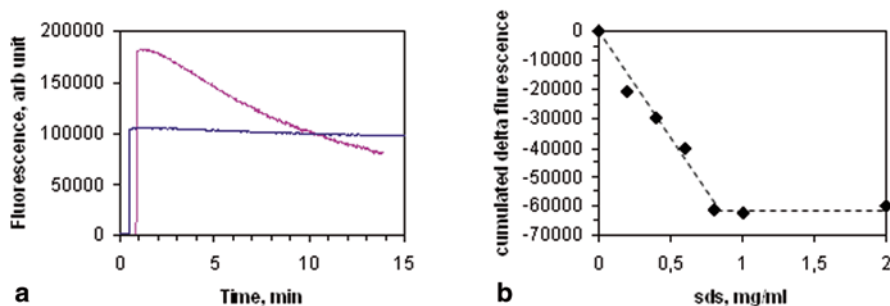


Fig. 14.8 Stability of TSPO in SDS. **a** Kinetics of change of intrinsic fluorescence of TSPO (excitation and emission wavelength set at 290 and 330 nm, respectively) recorded in a buffer in the absence (*upper curve*) or presence of SDS (CMC). **b** Titration of initial intrinsic fluorescence change as a function of SDS content of the buffer

14.2.6 Protein Stabilization

MPs in their native environment are surrounded by lipids and cofactors that are essential to maintain their active conformation. During protein purification, the detergent concentration above the critical micellar concentration (CMC) is used to avoid MP aggregation. Structural contacts between protein and lipids can be destabilized if the protein is extracted from membrane or can be absent if the protein is extracted from IBs (Baneres et al. 2011). Exposure to detergent can lead to protein denaturation. Intrinsic fluorescence can be used to check the stability of TSPO. Addition of SDS-purified TSPO to a buffer without detergent yields a characteristic protein fluorescence that slowly decreases with time (Fig. 14.8a). However, addition of SDS-purified TSPO to a buffer that contains a CMC of SDS yields very stable protein fluorescence. Titration of the SDS concentration required to achieve this protein stabilization gives the CMC value of SDS (Fig. 14.8b). Addition of SDS-purified TSPO to DPC-containing buffer yields a high protein fluorescence that remains relatively stable over time.

Among the different strategies for stabilizing an MP (Tate and Schertler 2009; Baneres et al. 2011, Chaps. 1, 7, 8, 9 in this volume), a simple one is to reincorporate detergent-solubilized protein into liposomes (Rigaud et al. 1995). TSPO was incubated with detergent and lipids, and the detergent was removed by the addition of Bio-Beads (Lacapere et al. 2001; Ostuni et al. 2010; Teboul et al. 2012). Formation of proteoliposomes can be followed using light scattering because small protein–detergent complexes do not diffuse light, whereas proteoliposomes are large objects that diffuse light (Ostuni et al. 2010). Reincorporation of TSPO into a lipid environment induces a large increase in intrinsic fluorescence (Teboul et al. 2012), which indicates that the tryptophan environment is changed (Fig. 14.9a, b). The emission spectra of TSPO show a blue shift, which indicates another reorganization of the tryptophan environment (Fig. 14.9a). We monitored changes in the protein secondary structure by recording CD spectra at different times during TSPO reconstitution

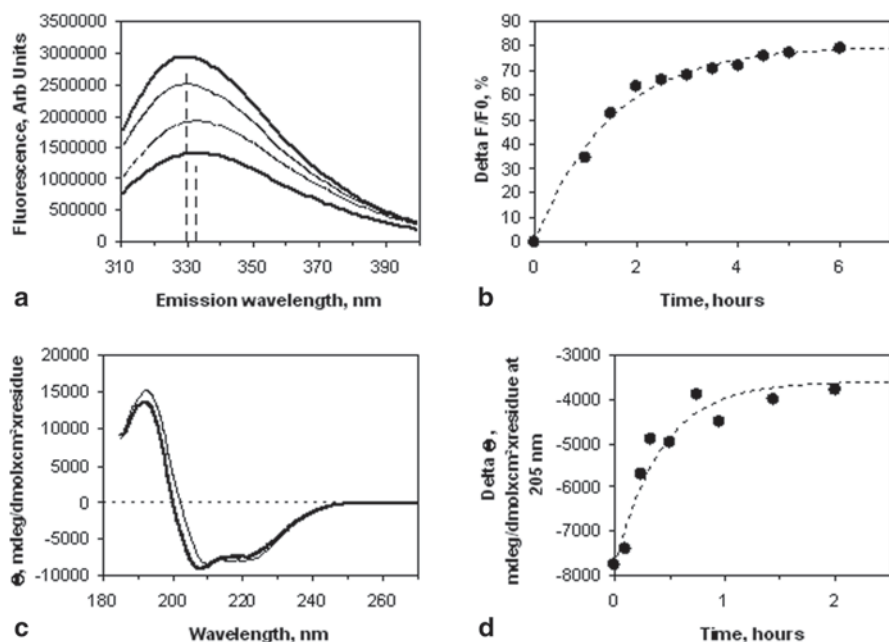


Fig. 14.9 Reconstitution of TSPO into proteoliposomes. **a** Intrinsic fluorescence spectra of TSPO (excitation wavelength set at 290 nm) recorded at different times of incubation of a mixture of TSPO, SDS, and lipids in the presence of Bio-Beads. *Vertical dotted lines* depict blue shift of the maxima of spectra. **b** Time course of intrinsic fluorescence changes plotted as a ratio of fluorescence increase to initial fluorescence. **c** Circular dichroism spectra of TSPO recorded at the beginning and the end of the early steps of reconstitution. **d** Time course of circular dichroism increase observed at 205 nm

(Fig. 14.9c). The spectra show both a shift and a decrease of the 205-nm minimum (Fig. 14.9d). The precise interpretation of these results is difficult, but they do reflect a change in the TSPO environment. The most accurate method to monitor the formation of proteoliposomes is electron microscopy (EM) imaging.

The reconstitution of TSPO into proteoliposomes stabilizes the protein structure and enables the recovery of high-affinity ligand binding (Lacapere et al. 2001; Teboul et al. 2012). PK 11195 binds to reconstituted TSPO proteoliposome with a nanomolar affinity, which is similar to that measured *in vivo* (Ostuni et al. 2009). However, proteoliposomes are not suitable for many structural studies. Ternary complexes of protein–lipid–detergent are better samples for crystallization and solution NMR. We observed that it was impossible to form ternary complexes of SDS-purified TSPO by adding lipids. This can be understood from solubilization experiments, which show that SDS tends to stick to the MP and strongly remove lipids (Kragh-Hansen et al. 1998). By contrast, we observed that addition of lipids to DDM-solubilized TSPO induces a large increase in intrinsic fluorescence for a

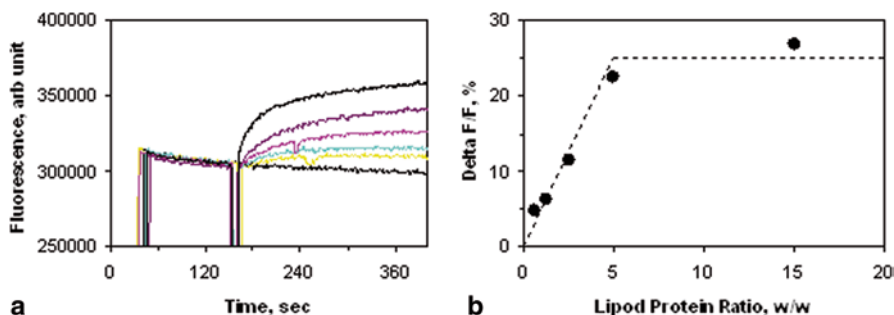


Fig. 14.10 Stability of TSPO in DDM. **a** Kinetics of intrinsic fluorescence change of TSPO (excitation and emission wavelength set at 290 and 330 nm, respectively) recorded in a DDM (CMC)-containing buffer. Repeated experiments were performed with addition of increasing amounts of lipids (DMPC:DMPE, 9:1), giving a lipid-to-protein ratio (LPR) ranging from 0 to 15 (w/w). **b** Titration of initial intrinsic fluorescence changes as a function of LPR content of the solution. Saturation observed at 5 (w/w) gives a value of 100 mol of lipids per mol of TSPO

saturation lipid-to-protein ratio (Fig. 14.10). The addition of lipids to DPC-solubilized TSPO does not produce any change in intrinsic fluorescence, which suggests that TSPO might be well structured in this detergent. This could be due to the choline head group of the DPC detergent, which could mimic the polar head group of phospholipids.

14.2.7 Production of Labeled TSPO for Structural Studies

3D analysis of protein crystals requires the presence of heavy atoms for phase determination when no previous structure is available to perform molecular replacement. Single- or multiple-anomalous dispersion (SAD or MAD) near the absorption edge of selenium can be efficiently used for phase determination (Pebay-Peyroula 2007). Replacement of methionines (Met) by selenomethionines (SeMet) can be achieved by overexpressing TSPO in bacteria cultured in a minimum medium complemented with SeMet (Guerrero et al. 2001).

NMR studies of large proteins require isotopic enrichment (Montaville and Jamin 2010) for the acquisition of multiple dimension spectra needed for 3D structure determination. Bacteria are a good system for isotopic enrichment because the addition of nitrogen- and carbon-enriched sources ($^{15}\text{NH}_4\text{Cl}$ or $(^{15}\text{NH}_4)_2\text{SO}_4$; ^{13}C -glucose or ^{13}C -glycerol) leads to the production of isotopically enriched proteins. The growth of cells in media containing D_2O leads to the high level of deuteration required for aliphatic detection. Bacterial overexpression of uniformly ^{15}N - and ^{13}C -labeled TSPO, and partially deuterated TSPO, has been performed, and enriched TSPO has been purified in sufficient amounts for NMR analysis (Robert and Lacapere 2010).

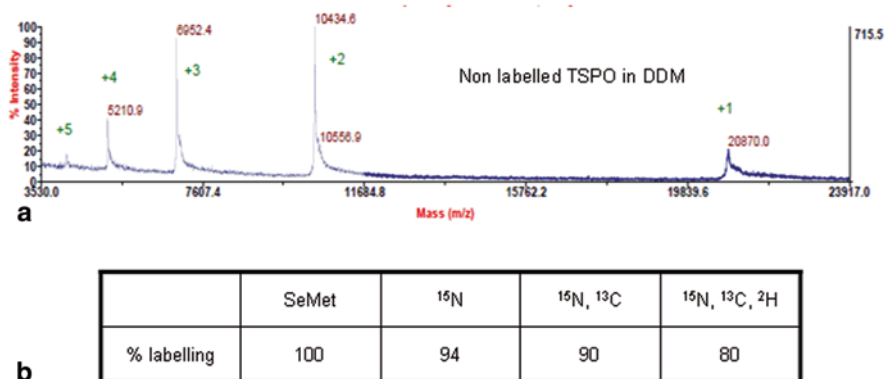


Fig. 14.11 Mass spectroscopy of TSPO. **a** MALDI-TOF spectrum of TSPO in DDM shows peaks corresponding to the different charges and enables calculation of the molecular mass of recombinant TSPO (20,870 Da). This corresponds to the full-length mouse TSPO containing the tag and omitting the first methionine. **b** Table presents the labeling percentage of recombinant mouse TSPO produced by bacterial overexpression in minimum medium supplemented with isotopically labeled sources

Mass spectroscopy is a perfect technique to measure protein labeling because the replacement of Met by SeMet and the replacement of ¹⁴N or ¹²C by ¹⁵N and ¹³C induces an increase in protein mass. However, it is difficult to perform mass spectroscopy with MPs due to the presence of detergent. The use of nonionic detergents such as DDM enabled MALDI-TOF analysis to be performed (Barrera et al. 2008; Sagan and Bolbach 2009). We exchanged the SDS surrounding TSPO with DDM (Fig. 14.3), and then obtained mass spectra of TSPO (Fig. 14.11a) that permitted to characterize the labeled TSPOs. We showed that all Met of TSPO were replaced by SeMet, and the incorporation of ¹⁵N, ¹³C, and ²H isotopes was very good (Fig. 14.11b).

14.3 Structural Studies of TSPO

14.3.1 Three-Dimensional Crystals

3D crystallization of MPs still remains challenging. The first bottleneck usually encountered is the quantity and quality of the available sample. In these studies of TSPO, the yield and concentration that were obtained were suitable for crystallization trials. The protein characterization (described in Sect. 14.2.5) was indicative of a well-folded and rather stable protein. The TSPO crystallization method utilized stand crystallization techniques called vapor diffusion.

14.3.1.1 Initial Crystallization Trials

The crystallization experiments were initially performed using the SDS-purified protein, but no protein crystals were obtained. Subsequently, crystallization trials were performed with DPC-purified TSPO, with protein and detergent concentrations estimated at 13 and 30 mg mL⁻¹, respectively. Conformational stability is an essential parameter for successful crystallization (see Sect. 14.2.6). Therefore, we tried from the beginning to add the specific ligand PK 11195 to the protein sample. Therefore, for comparison, only half of the sample contained 5 mM of the ligand. Screening of crystallization conditions was performed using the nanodrop HTX robot at EMBL (PSB Platform, Grenoble, France) and a sparse matrix approach, by testing approximately 400 commercial screen conditions from Qiagen. Several conditions that contained a PEG (polyethylene glycol)/salt mixture as precipitant led to phase separation or crystalline precipitates. No obvious difference was observed whether the ligand was present or not.

These conditions were further explored and refined using homemade screens combining different concentrations of PEG, salt, and different pH. A few crystals were obtained and investigated at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Some crystals showed a diffraction pattern that was typical for salt crystals; for others, no diffraction pattern was observed. Therefore, it was not possible to conclude whether the obtained crystals were protein or detergent crystals.

14.3.1.2 Exploring the Effects of Different Protein Concentrations and Ligands

Following the same strategy described above, different TSPO batches purified in DPC at protein concentrations ranging from 8 to 56 mg mL⁻¹ were subjected to crystallization trials. Crystallization in the presence of the PK 11195 ligand was explored. The addition of lipids (9:1 dimyristoylphosphatidylcholine, DMPC to phosphatidylethanolamine, PE) was tested for a stabilizing effect (see Sect. 14.2.6). We also performed the production and crystallization of the SeMet-labeled protein (see Sect. 14.2.7) concentrated to 18 mg mL⁻¹.

Phase separations, crystalline precipitates, and urchins or microcrystals were observed in some drops by days or weeks after the initial setup (Fig. 14.12a–f). Typically, the most favorable precipitants contained PEG (with different molecular masses) and different salts. Objects of sufficient size were tested at the ESRF. All crystals obtained with the protein–lipid complex were lipid crystals. For others, diffraction spots that could be characteristic of protein crystals were observed. The best crystals showed diffraction patterns up to a resolution of 20 Å (Fig. 14.12g).

Extensive optimization of setup parameters around the best conditions was conducted using commercial additives, detergents, screens, and homemade plates. Although some conditions could be reproduced, it was not possible to improve the diffraction quality.

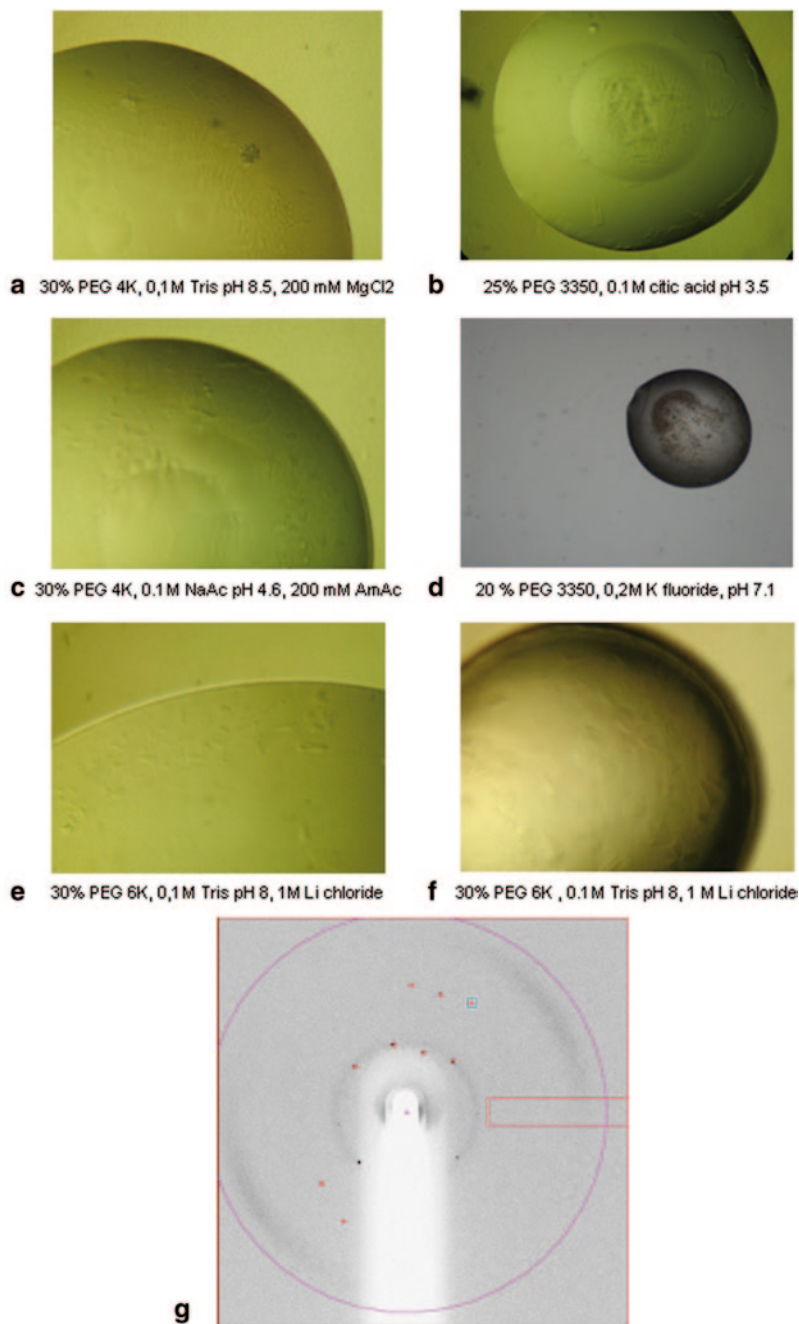


Fig. 14.12 Three-dimensional crystallization of TSPO. **a–e** Hits obtained for native TSPO purified in DPC, except **d**, where TSPO was purified in DDM, and **f** for TSPO in the presence of 5 mM PK 11195. Experimental conditions are described in each figure. **g** Typical diffraction pattern obtained for a crystal grown in the conditions shown on panel **e**. Data were collected on ID23-2 at the ESRF. The cell parameters predicted from the diffraction images are compatible with protein crystals and the presence of several TSPO molecules per asymmetric unit

14.3.1.3 Exploring the Effects of Other Detergents

Choosing the right detergent for protein crystallization is an essential key to success. Although some crystals can be obtained using various detergents, the diffraction quality is usually highly dependent on the detergent used. Therefore, in an effort to increase the success rate of the TSPO crystallization, alternative detergents were considered. Because the protein was purified and characterized in DDM, a detergent that has been successful for MP crystallization (more successful than DPC), new crystallization experiments were performed in this detergent. The protein and detergent concentrations were estimated at 12 and 30 mg mL⁻¹, respectively. Screening with the robot and the commercial kits yielded promising hits, but they could not be reproduced (e.g., Fig. 14.12d).

Crystallization is an empirical technique, especially for MPs, and it is not standardized. Important parameters include protein stability, homogeneity, and detergent properties. The crystallization trials performed with TSPO generated promising results and low-diffracting protein crystals. The optimization parameters, screening additives, protein concentration, and detergent were not sufficient to improve the diffraction quality of the protein crystals. Many other factors could be explored, including other types and concentrations of detergent. The control of protein purity, homogeneity, and stability is crucial for protein crystallization. *In meso* crystallization could offer an interesting alternative to approach TSPO crystallization.

In the meantime, another group working on TSPO of *Rhodobacter sphaeroides* reported that protein purified in DDM was able to crystallize in a vapor diffusion and bicelle setup (Li et al. 2012). This suggested that the use of bacterial TSPO for 3D crystallization could be a better choice than the use of mammalian TSPO. However, no diffraction data on the protein structure of *Rhodobacter* TSPO crystals have been published yet.

14.3.2 NMR Studies

Early NMR structural studies of TSPO were performed on protein fragments because of the high molecular weight of TSPO (18 kDa). TSPO is a mitochondrial MP that functions in the transport of cholesterol. A cholesterol-recognition amino acid consensus sequence (CRAC) has been identified (Li and Papadopoulos 1998) and observed in other proteins that interact with cholesterol. A solution ¹H-NMR study of a synthetic peptide corresponding to the C-terminus of TSPO containing the CRAC domain suggested that a binding site for cholesterol was in a groove made by aromatic residues (Jamin et al. 2005). This was confirmed by site-directed mutagenesis of the entire TSPO, and docking of cholesterol in the peptide structure was consistent with cholesterol bound in the groove and capped by an arginine (Jamin et al. 2005). Another solution ¹H-NMR study was performed on synthetic peptides corresponding to the five TM domains of TSPO (Murail et al. 2008). It clearly revealed that these peptides form helical structures. Recently, the production of a recombinant fragment containing a double-TM domain with its connecting loop

(TM4TM5) was successfully performed by overexpression in bacteria (Galvagnion et al. 2013). Optimization of the solubilization conditions for NMR enabled the 1D ^1H -NMR and well-resolved 2D ^1H - ^{15}N HSQC (heteronuclear single quantum correlation) of the recombinant double TM domain fragment.

For the entire TSPO solubilized in SDS, the solution 1D ^1H -NMR spectrum shows peaks in the backbone amide protons, tryptophan side chain indole NH groups, and upfield methyl region (Fig. 14.13a). These results suggest that the TSPO is at least partially structured. A solution 2D ^1H - ^{15}N HSQC spectrum of TSPO solubilized in SDS exhibits poorly resolved peaks especially in the central region at approximately 8 ppm (Fig. 14.13f). The 12 expected indole NH resonances ($\delta^1\text{H}$ and $\delta^{15}\text{N}$ at approximately 10 and 128 ppm, respectively) are not well separated, whereas the glycine region ($\delta^{15}\text{N}$ at approximately 106 ppm) shows more resolved peaks. The solution 2D ^1H - ^{13}C HSQC spectrum of TSPO solubilized in SDS shows poor chemical shift dispersion. In particular, the methyl groups of the five expected methionines ($\delta^1\text{H}$ at approximately 2 ppm) are not resolved (Fig. 14.13i).

To analyze the effect of detergents on the NMR spectra of TSPO, protein purified in SDS was added to solutions containing CMC of SDS, DDM, or DPC. The presence of DDM induces only minor modifications of the TSPO spectrum, such as a small low-field displacement of tryptophan indole NH protons (Fig. 14.13b). By contrast, the presence of DPC induces a notable low-field shift of tryptophan indole NH of TSPO (Fig. 14.13b). No modifications are observed in the upfield region of the 1D ^1H spectrum of TSPO irrespective of the detergent (Fig. 14.13c). SDS can be fully exchanged by DPC, and this exchange induces modifications of the 1D ^1H spectrum of TSPO both in the tryptophan indole NH (Fig. 14.13d) and in the upfield (Fig. 14.13e) regions. However, the 1D ^1H , 2D ^1H - ^{15}N HSQC, and 2D ^1H - ^{13}C HSQC spectra are not well resolved and show broadening of some peaks, which suggests conformational exchange.

The addition of the high-affinity drug ligand PK 11195 leads to a spectacular improvement of NMR spectral quality (Murail et al. 2008). The 1D ^1H spectrum of TSPO shows numerous resolved peaks for the tryptophan indole NH (Fig. 14.13d) and in the upfield region of methyl resonances (Fig. 14.13e). Similarly, solution 2D ^1H - ^{15}N (Fig. 14.13h) and 2D ^1H - ^{13}C (Fig. 14.13k) HSQC of TSPO solubilized in DPC exhibit large chemical shift dispersion. In particular, the expected 12 side chain tryptophan NH groups are well differentiated, which suggests distinct local environments in the PK 11195-bound TSPO structure. A similar observation can be drawn for the five expected methionine methyl groups ($\delta^1\text{H}$ and $\delta^{13}\text{C}$ at approximately 2 and 17 ppm, respectively), the three isoleucine methyl resonances ($\delta^{13}\text{C}$ at approximately 14 ppm), and the alanine resonances ($\delta^1\text{H}$ and $\delta^{13}\text{C}$ at approximately 1.6 and 18 ppm, respectively). These data confirm that PK 11195 stabilizes the conformation of TSPO. However, the stabilization of a unique conformation of TSPO over a long period, which is required for 3D NMR experiments acquisition, a first step to determine the protein atomic structure, remains a difficult challenge.

Good stabilization has been recently obtained by adding a (R) enantiomer of PK 11195 leading to the first atomic structure of mouse TSPO (Jaremko et al. 2014).

Stabilization of TSPO can be obtained by placing the protein in a lipid environment (see Sect. 14.2.6). Recent work shows that MP structure can be determined

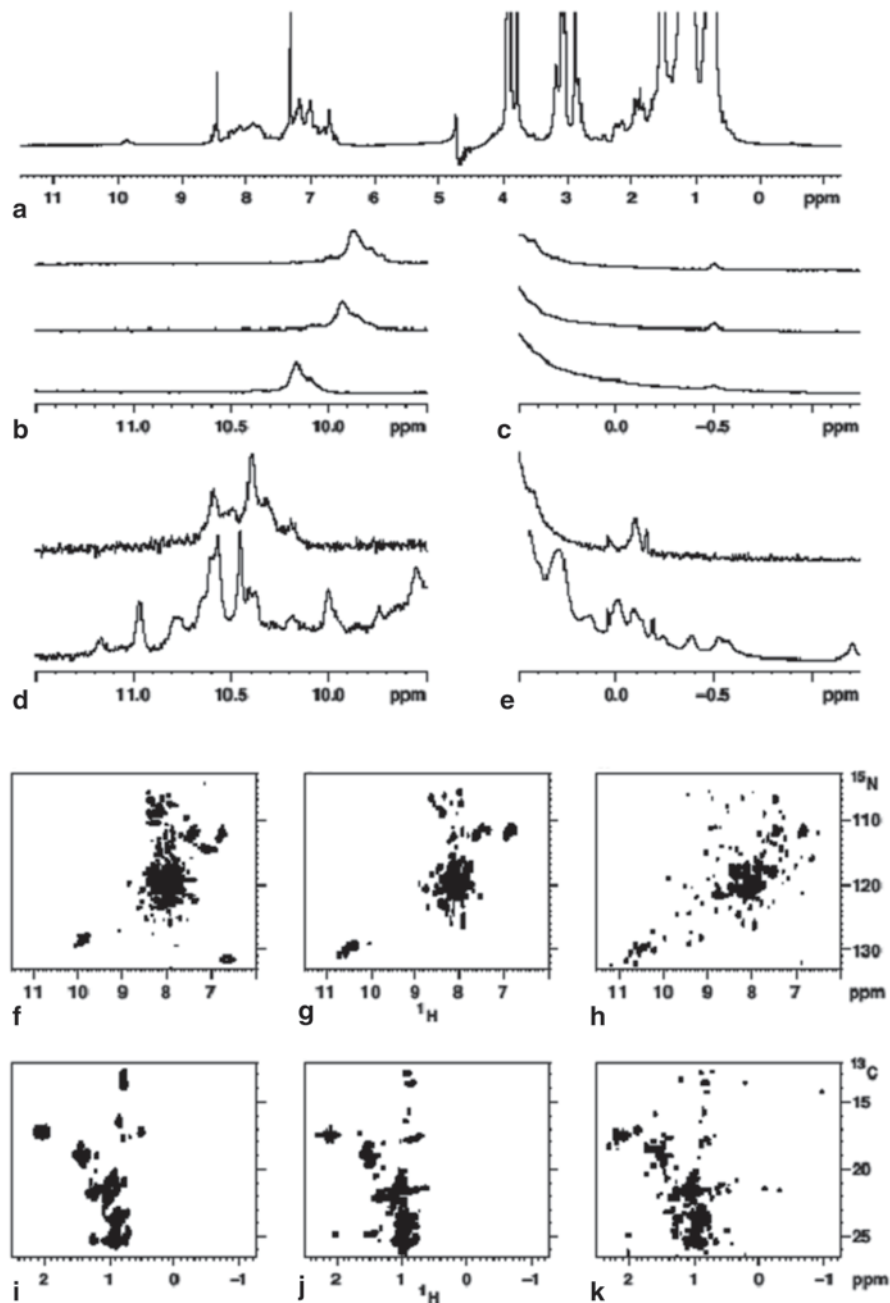


Fig. 14.13 Solution NMR spectra of TSPO. **a** 1D ^1H -NMR spectrum of TSPO solubilized in SDS. **b** and **c** 1D ^1H -NMR spectra of indole NH **b** and methyl proton **c** regions of TSPO solubilized in the presence of SDS (*top*), DDM (*middle*), and DPC (*bottom*). **d** and **e** 1D ^1H -NMR spectra of indole NH **d** and methyl proton **e** regions of TSPO solubilized in DPC in the absence (*top*) and in the presence (*bottom*) of PK 11195. **f**, **g**, and **h** 2D ^1H - ^{15}N HSQC spectra of TSPO solubilized in SDS **f** or DPC **g**, **h**, in the absence **g** and presence **h** of PK 11195. **i**, **j**, and **k** Methyl region of 2D ^1H - ^{13}C HSQC spectra of TSPO solubilized in SDS **i** or DPC **j**, **k**, in the absence **j** and presence **k** of PK 11195

using solid-state NMR (ssNMR) of proteoliposomes (Marassi et al. 2012; Das et al. 2012; Park et al. 2012; see also Chap. 12 in this volume). ^{15}N - and ^{13}C -labeled TSPO has been reconstituted at a lipid-to-protein ratio of 10 (w/w) and ssNMR spectra were recorded at magic angle spinning (MAS). The 1D MAS ^{13}C spectrum displays sharp, intense peaks (Fig. 14.14a) that correspond to the natural abundance of ^{13}C lipid signals that can be seen on the 1D spectrum of pure lipids (Fig. 14.14b). A lower lipid-to-protein ratio (1 w/w) was used to reduce the lipid contribution to the NMR spectra (Fig. 14.14c). The 2D MAS ^{13}C - ^{13}C correlation spectra in the absence and presence of PK 11195 were recorded (Fig. 14.14d, e). In the absence of PK 11195, the spectral quality of TSPO in its native lipid environment is promising and could be amenable to structural studies. A dramatic effect of PK 11195 is clearly observed on the spectrum in terms of chemical shift dispersion and number of observed correlations. This might be due to a conformational change of TSPO induced by PK 11195 that results in improved structural homogeneity, and could therefore open the way toward the structural determination of an alternative ligand-bound conformation of TSPO.

14.3.3 Two-Dimensional Crystals and Electron Microscopy

The 2D crystallization of MPs combined with EM observation and analysis has been an alternative to 3D crystallization. This approach requires low protein concentration and enables direct imaging of the objects (Lacapere 2010). Several types of 2D crystals can be grown such as sheets or tubes (Lacapere 2010).

The first 3D structure of TSPO was obtained by electron cryomicroscopy of helical crystals of bacterial TSPO from *Rhodobacter sphaeroides* (Korkhov et al. 2010). Although the data were of low resolution (10 Å), the 3D reconstruction revealed monomers with five TM domains and dimeric association. More recently, images of mouse TSPO interacting with a functionalized monolayer have been obtained (Teboul et al. 2012). The EM image (Fig. 14.15a) reveals particles (Fig. 14.15b) that are aligned and averaged (Fig. 14.15c). The diameter of a particle is approximately 6 nm, whereas the bacterial monomer is 2.5 nm (Fig. 14.15e, f). This raises a question about the degree of polymeric association of mammalian TSPO. The bacterial monomer can be enclosed either in a circle or in an ellipse (Fig. 14.15f) that can be used in an attempt to fit mouse TSPO using an isodensity contour level of the averaged image. Several densities surrounding a central hole can be clearly seen (Fig. 14.15d), and at least two different fittings can be obtained with three ellipses or four circles (Fig. 14.15g). The projection map of the bacterial dimer accounts for the entire protein because it is obtained from cryoimages, whereas the mammalian map is obtained from negative staining images that account for only the external domain of TSPO that is accessible to the negative stain. Thus, counting and comparing the number of densities (3 for bacterial TSPO, Fig. 14.15e, and 4–6 for mammalian TSPO, Fig. 14.15d) is not possible, whereas comparing the overall dimensions is more simple and accurate. At this stage, the following two questions can be raised: (1) Are there different TSPO polymers depending on the species (bacterial or mammalian)? (2) Are different TSPO polymers relevant for transporter function, or are they a crystallization artifact?

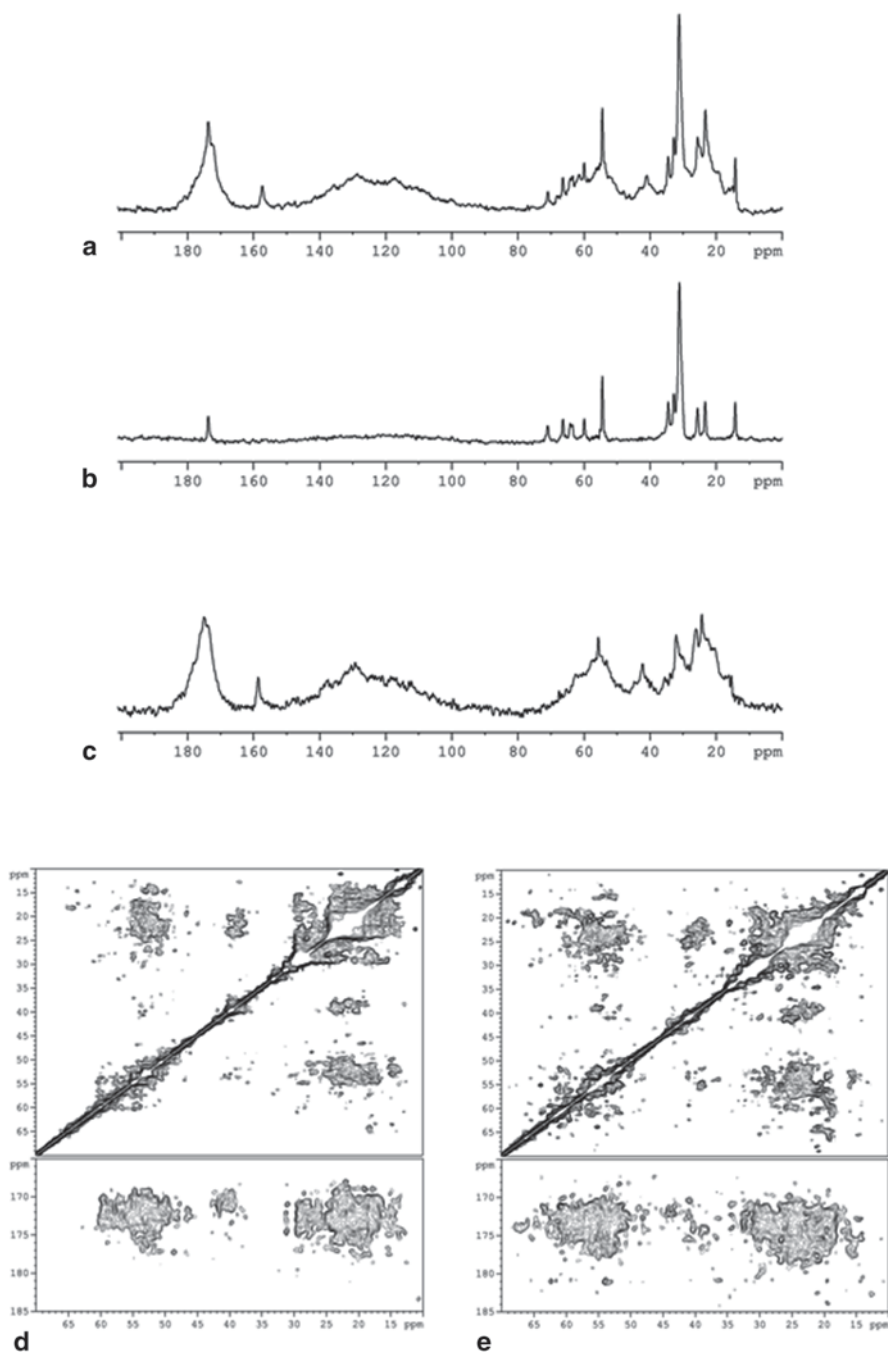


Fig. 14.14 Solid-state MAS NMR spectra of TSPO. **a** and **c** 1D MAS ^{13}C NMR spectra of TSPO reconstituted in proteoliposomes at a lipid-to-protein ratio of 10 (w/w) (**a**) or 1 w/w (**c**), and of pure lipid liposomes (DMPC:PE, 9:1) (**b**). **d** and **e** 2D MAS ^{13}C - ^{13}C NMR correlation spectra of TSPO reconstituted in proteoliposomes in the absence (**d**) or presence (**e**) of PK 11195, recorded using PARIS recoupling (Weingarth et al. 2009) and covariance treatment (Weingarth et al. 2010)

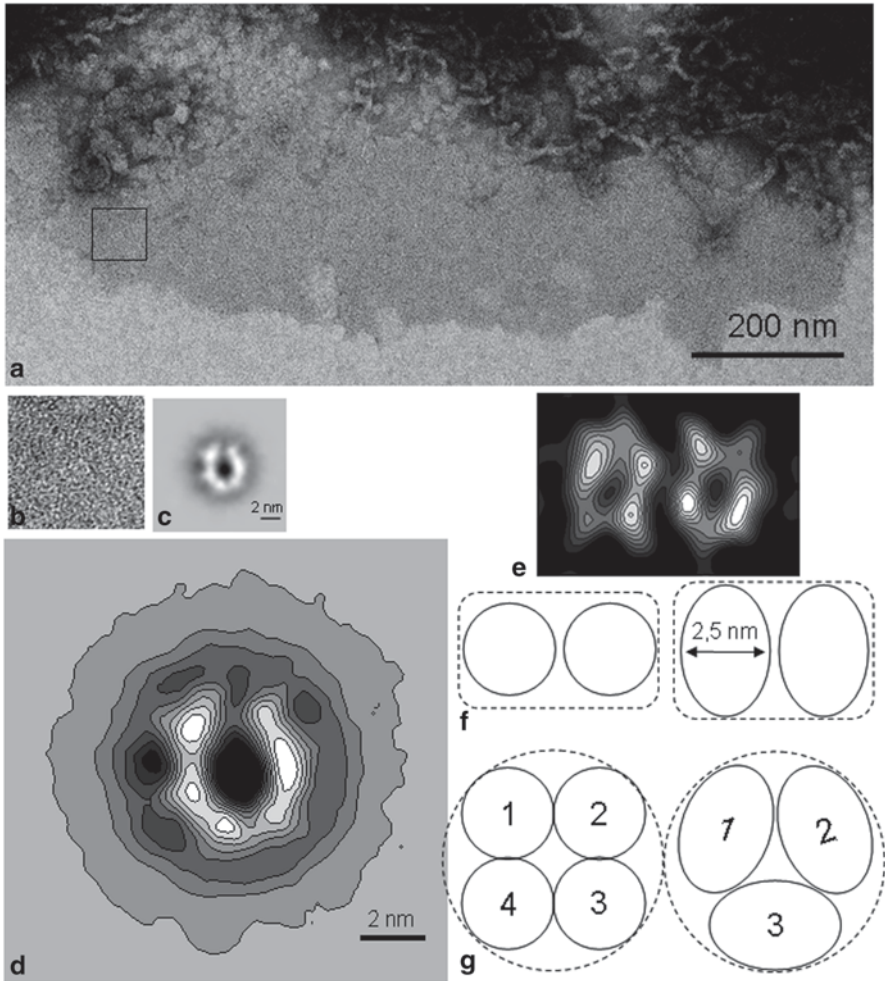


Fig. 14.15 Two-dimensional crystallization of TSPO. **a** Transmission electron microscopy image of recombinant TSPO adsorbed under a functionalized monolayer. **b** Enlargement of particles contained in the *black box* shown in panel **a**. **c** Average of particles obtained after correspondence analysis (Teboul et al. 2012). **d** Superimposed contour level of average particle. **e** Computed projection of a dimer from the cryo-TEM density map of bacterial TSPO (Khorkov et al. 2010). **f** Schemes depicting the dimer of bacterial TSPO represented either by two *circles* or two *ellipses* enclosed in a *rectangle* with rounded corners. **g** Schemes depicting overall size of average image of recombinant TSPO enclosing either three or four putative monomers

In bacterial TSPO crystals, lipid-inserted TSPO dimers are associated in a ring and proteolipid tubes are formed by ring stacking. The 3D structure shows close interactions in a ring between dimers at the level of TM helices, but some space between dimers, suggesting the interactions of loops (Korkhov et al. 2010). The rings are stacked with a rotation from one ring to the neighboring ring. Interactions

involved in the stabilization of adjacent rings within the tube are not well understood because no atomic structure is available. It may involve the C-terminal region that contains the poly-histidine tag. The crystal structure may suggest a dimeric association as a functional unit, but could not exclude higher degrees of polymerization. The functional interpretation of bacterial TSPO arrangement is rather complex and depends on the functional unit considered. Substrate translocation can occur within the core of a monomer; the presence of a dimer remains unexplained. Substrate translocation may occur at the interface of a dimer and thereby provide functional data for the existence of the dimer.

The structural determination of the mammalian TSPO is more complicated. The association observed in the particle can be due to close packing of proteins adsorbed under a functionalized monolayer. In this case, it often occurs through hexameric associations that minimize energy interactions, but the presence of a protein in the middle of the putative hexamer would be expected. The overall size would be greater than that measured. Both tetrameric and trimeric associations, which fit into the density, give a central hole that could accommodate substrate translocation. This would indicate a stoichiometry of one transported molecule per trimer or tetramer, and a possible cooperative mechanism for transport. No data are available to distinguish between these possibilities. It has been reported that mouse TSPO can form a covalent polymer in response to steroid and ROS production (Delavoie et al. 2003), but no unique polymeric stoichiometry has been described. The most common mechanism suggested for cholesterol transport activated by pharmacological ligands proposes the lowest stoichiometry of one cholesterol and one PK 11195-binding site per TSPO.

We cannot exclude that bacterial and mammalian TSPO have different unit associations because they have different functions. *Rhodobacter* TSPO regulates photosynthetic gene expression involving tetrapyrrole homeostasis (Yeliseev and Kaplan 2000). Mammalian TSPO regulates biosynthesis of steroids and bile salts involving cholesterol transport (Lacapere and Papadopoulos 2003). The cytosolic loop involved in PK 11195 binding is different in bacterial and mammalian TSPOs; bacterial TSPO is shorter and lacks residues important for the ligand binding.

14.3.4 Molecular Modeling

Bioinformatic- and molecular-modeling approaches are powerful tools to construct the atomic structure of proteins from X-ray crystallography, NMR spectroscopy, or amino acid sequence. For the first two approaches, experimental data are needed that are not available for TSPO. For the third approach, several structures have been modeled by sequence alignment and homology mapping with proteins of known atomic structure.

The first 3D model of TSPO was generated using molecular dynamics simulation and the bacteriorhodopsin atomic structure (Bernassau et al. 1993). In this model, the five helices are arranged clockwise from one to five (Fig. 14.16a). This model for the

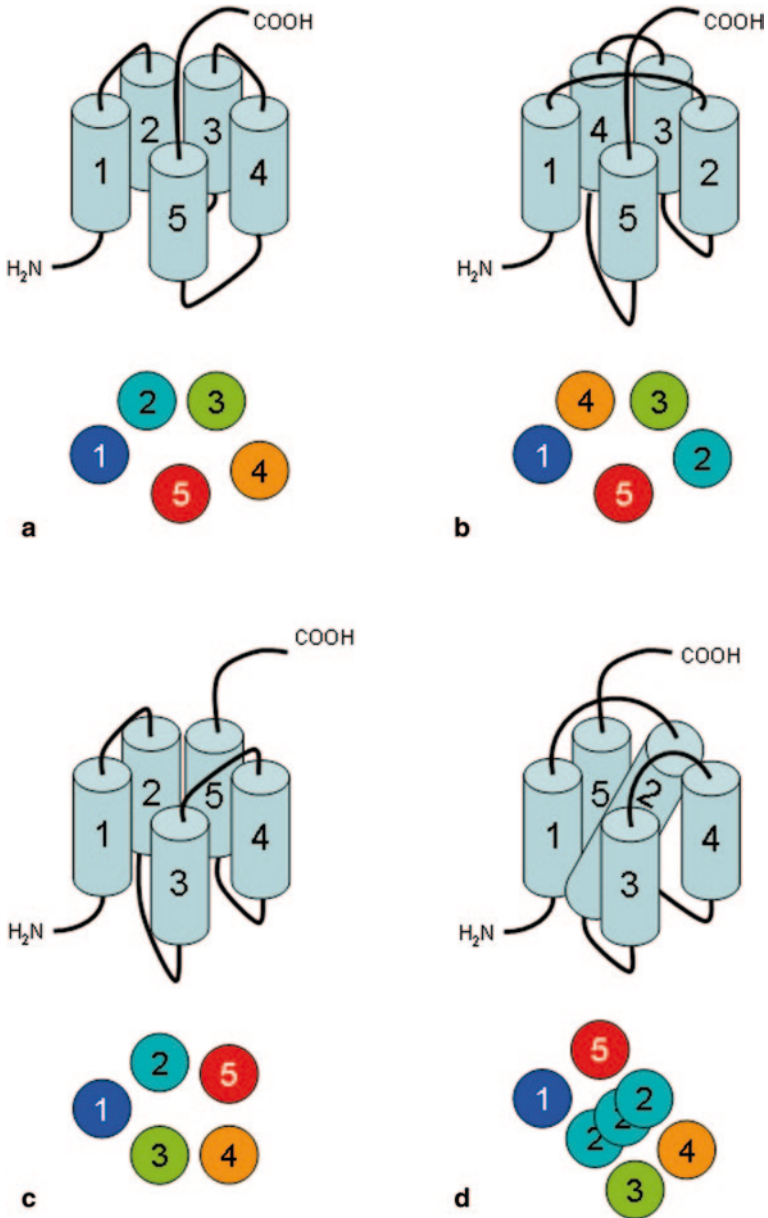


Fig. 14.16 Molecular modeling of TSPO. Schematic representation of different models that have been computed from existing 3D structures. **a** Bacteriorhodopsin (Bernassau et al. 1993), **b** apolipoprotein III (Anzini et al. 2001), or *ab initio* from evolutionary sequence variations (**c** and **d**; Hopf et al. 2012). For each model, side view and top view are shown. The top view is from the C-terminus

human TSPO was used later to build another mammalian TSPO model after mutation of the human sequence to the mouse sequence (Culty et al. 1999). This approach positioned the TSPO ligand-binding site using a series of amino acid deletions and mutations that reduced the binding affinity for PK 11195 and cholesterol (Li and Papadopoulos 1998). The binding of PK 11195 involved external loops and the C-terminal region of TSPO. The binding of cholesterol was proposed to occur at the CRAC domain (Li and Papadopoulos 1998), and it was passed through a channel located in the middle of the bundle made by the five TM helices (Ruppecht et al. 2010).

Another TSPO model was built by considering the structural restraints drawn by the apolipoprotein III for the TM region and the myohemerythrin for the first loop, whereas the third loop was modeled *ab initio* (Anzini et al. 2001). In this full-length model, the five helices are not in a clockwise sequence, because the positions of TM2 and TM4 permuted in agreement with the apolipoprotein structure (Fig. 14.16b). The main consequence is that the exit of the channel in the middle of the bundle made by the five TM helices might be closed more tightly on the N-terminal side by the short loops connecting TM2–TM3 and TM4–TM5. On the opposite side, the loops connecting TM1–TM2 and TM3–TM4 were used to analyze mutations that affect ligand binding and the interaction of different chemical classes of ligands that are not structurally related (Anzini et al. 2001). The flexibility of these loops is sufficient to open and close the entrance of the putative channel, especially if cholesterol binds on the CRAC domain located in a groove at the C-terminus of TSPO (Jamin et al. 2005).

Several publications have described 3D protein structure computed from evolutionary sequence variation (Marks et al. 2011). This was applied recently to human TSPO; two different 3D models were proposed (Hopf et al. 2012), and two distinct TM topologies have been proposed (Fig. 14.16c, d). The former model does not show a sequential clockwise position of the five TM, because the positions of TM3 and TM5 have been permuted (Fig. 14.16c). The exit of the putative channel for cholesterol might be closed on the N-terminal side of TSPO. The later model is more distinct because it suggests a highly tilted TM2 that completely occupies the putative channel for cholesterol. Thus, it implies a completely different mechanism for cholesterol transport.

All these models are very useful to produce a 3D structure of TSPO, but no model is based on experimental data from the TSPO family. They may help to resolve amino acid mutations, but they cannot be used to simulate transport by TSPO. There is still a need for the first atomic structure of any TSPO.

It has to be mentioned that during the edition process of this book, a first atomic structure has been solved showing that topology model shown in figure 14.16c is the correct one (Jaremko et al 2014).

Acknowledgments We would like to thank S. Beaufils, G. Bolbach, F. Delavoie, J. Fan, C. Galvagnion, N. Jamin, P. Montaville, H. Li, S. Murail, J. M. Neumann, M. A. Ostuni, J. C. Robert, S. Sagan, J. C. Taveau, D. Teboul, V. Vie, C. Venien, and Z. X. Yao, who participated in the course of this investigation of TSPO overexpression and structure determination. This work was supported with funds from the Institute National de la Santé et de la Recherche Medical, the Centre National de la Recherche Scientifique, the French National Agency for Research (ANR BLAN-0190-01), and the Canadian Institutes of Health Research.

References

- Anzini M, Cappelli A, Vomero S, Seeber M, Menziani MC, Langer T, Hagen B, Manzoni C, Bourguignon JJ (2001) Mapping and fitting the peripheral benzodiazepine receptor binding site by carboxamide derivatives. Comparison of different approaches to quantitative ligand-receptor interaction modeling. *J Med Chem* 44:1134–1150
- Baneres JL, Popot JL, Mouillac B (2011) New advances in production of G-protein-coupled receptors. *Trends Biotechnol* 29:314–322
- Barrera NP, Di Bartolo N, Booth PJ, Robinson CV (2008) Micelles protect membrane complexes from solution to vacuum. *Science* 321:243–246
- Bernassau JM, Reversat JL, Ferrara P, Caput D, LeFur G (1993) A 3D model of the peripheral benzodiazepine receptor and its implication in intra mitochondrial cholesterol transport. *J Mol Graph* 11:236–244
- Culty M, Li H, Boujrad N, Amri H, Vidic B, Bernassau JM, Reversat JL, Papadopoulos V (1999) In vitro studies on the role of the peripheral-type benzodiazepine receptor in steroidogenesis. *J Steroid Biochem Mol Biol* 69:123–130
- Das BB, Nothnagel HJ, Lu GJ, Son WS, Tian Y, Marassi FM, Opella SJ (2012) Structure determination of a membrane protein in proteoliposomes. *J Am Chem Soc* 134(4):2047–2056
- Delavoie F, Li H, Hardwick M, Robert JC, Giatzakis C, Peranzi G, Yao ZX, Maccario J, Lacapere JJ, Papadopoulos V (2003) In vivo and in vitro peripheral-type benzodiazepine receptor polymerization: functional significance in drug ligand and cholesterol binding. *Biochemistry* 42(15):4506–4519
- Fan J, Lindemann P, Feuilloley MG, Papadopoulos V (2012) Structural and functional evolution of the translocator protein (18 kDa). *Curr Mol Med* 12(4):369–386
- Galvagnion C, Montaville P, Coïc YM, Jamin N (2013) Production and initial structural characterization of the TM4TM5 helix-loop-helix domain of the translocation protein. *J Pept Sci* 19(2):102–109
- Garnier M, Dimchev AB, Boujrad N, Price JM, Musto NA, Papadopoulos V (1994) In vitro reconstitution of a functional peripheral-type benzodiazepine receptor from mouse Leydig tumor cells. *Mol Pharmacol* 45:201–211
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) *The proteomic protocols handbook*. Humana, Totowa, pp 571–607
- Ginter C, Kiburu I, Boudker O (2013) Chemical catalysis by the translocator protein (18 Kda). *Biochemistry* 52(21):3609–3611
- Goormaghtigh E, Gasper R, Bénard A, Goldsztein A, Raussens V (2009) Protein secondary structure content in solution, films and tissues: redundancy and complementarity of the information content in circular dichroism, transmission and ATR FTIR spectra. *Biochim Biophys Acta* 1794(9):1332–1343
- Guerrero SA, Hecht HJ, Hofmann B, Biebl H, Singh M (2001) production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. *Appl Microbiol Biotechnol* 56:718–723
- Guillaumot D, Guillon S, Morsomme P, Batako H (2009) The Arabidopsis TSPO-related protein is a stress and abscisic acid-regulated, endoplasmic reticulum–Golgi-localized membrane protein. *Plant J* 60:242–256
- Hielscher R, Friedrich T, Hellwig P (2011) Far- and mid-infrared spectroscopy analysis of the substrate-induced structural dynamics of respiratory complex I. *Chemphyschem* 12(1):217–224
- Hopf TA, Colwell LJ, Sheridan R, Rost B, Sander C, Marks DS (2012) Three-dimensional structures of membrane proteins from genomic sequencing. *Cell* 149:1607–1621
- Issop L, Rone MB, Papadopoulos V (2013) Organelle plasticity and interactions in cholesterol transport and steroid biosynthesis. *Mol Cell Endocrinol* 371(1-2):34–46

- Jamin N, Lacapère JJ (2007) Circular dichroism as a tool for controlling membrane protein folding or structural modifications. In: Pebey-Peyroula E (ed) *Biophysical analysis of membrane proteins. Investigating structure and function*. Wiley-VCH, Weinheim, pp 243–258
- Jamin N, Neumann JM, Ostuni MA, Vu TK, Yao ZX, Murail S, Robert JC, Giatzakis C, Papadopoulos V, Lacapere JJ (2005) Characterization of the cholesterol recognition amino acid consensus sequence of the peripheral-type benzodiazepine receptor. *Mol Endocrinol* 19(3):588–594
- Jaremko L, Jaremko M, Giller K, Becker S, Zweckstetter M (2014) Structure of the mitochondrial translocator protein in complex with a diagnostic ligand. *Science* 343(6177):1363–6
- Joseph-Liauzin E, Delmas P, Shire D, Ferrara P (1998) Topological analysis of the peripheral benzodiazepine receptor in yeast mitochondrial membranes supports a five-transmembrane structure. *J Biol Chem* 273(4):2146–2152
- Kragh-Hansen U, Le Maire M, Moller J (1998) The mechanism of detergent solubilization of liposomes and protein-containing membranes. *Biophys J* 75:2932–2946
- Korkhov VM, Sachse C, Short JM, Tate CG (2010) Three-dimensional structure of TspO by electron cryomicroscopy of helical crystals. *Structure* 18:677–687
- Lacapere JJ (2010) Membrane protein structure determination: methods and protocols. *Methods in molecular biology*, vol 654. Humana, New York
- Lacapere JJ, Papadopoulos V (2003) Peripheral-type benzodiazepine receptor: structure and function of a cholesterol binding protein. *Steroids* 67:569–585
- Lacapere JJ, Delavoie F, Li H, Peranzi G, Maccario J, Papadopoulos V, Vidic B (2001) Structural and functional study of reconstituted peripheral benzodiazepine receptor. *Biochem Biophys Res Comm* 284:536–541
- Lacapère JJ, Peybay-Peyroula E, Neumann JM, Etchebest C (2007) Determining membrane proteins structures: still a challenge! *Trends Biochem Sci* 32(6):259–270
- Li H, Papadopoulos V (1998) Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern. *Endocrinology* 139:4991–4997
- Li F, Hiser C, Ferguson-Miller S (2012) Expression, purification and characterization of bacterial and human translocator protein 18 kDa (TSPO). *Biophys J* 102(3):247a–248a
- Marassi FM, Bibhuti BB, Lu GJ, Nothnagel HJ, Park SH, Son WS, Tian Y, Opella SJ (2012) Structure determination of membrane proteins in five easy pieces. *Methods* 55:363–369
- Marks DS, Colwell LJ, Sheridan R, Hopf TA, Pagnani A, Zecchina R, Sander C (2011) Protein 3D structure computed from evolutionary sequence variation. *PLoS One* 6(12):e28766
- Montaville P, Jamin N (2010) Determination of membrane protein structures using solution and solid-state NMR. *Methods Mol Biol* 654:261–282
- Mouillac B, Banères JL (2010) Mammalian membrane receptors expression as inclusion bodies in *Escherichia coli*. *Methods Mol Biol* 601:39–48
- Murail S, Robert JC, Coïc YM, Neumann JM, Ostuni MA, Yao ZX, Papadopoulos V, Jamin N, Lacapère JJ (2008) Investigating the structural properties of the translocator protein TSPO by CD and NMR. Ligand binding induces conformational changes. *Biochim Biophys Acta* 1778(6):1375–1381
- Neehaul Y, Juárez O, Barquera B, Hellwig P (2013) Infrared spectroscopic evidence of a redox-dependent conformational change involving ion binding residue NqrB-D397 in the Na(+)-pumping NADH:quinone oxidoreductase from *Vibrio cholerae*. *Biochemistry* 52(18):3085–93
- Ostuni MA, Péranzi G, Ducroc RA, Fasseu M, Vidic B, Dumont J, Papadopoulos V, Lacapere JJ (2009) Distribution, pharmacological characterization and function of the 18 kDa translocator in rat small intestine. *Biol Cell* 101(10):573–586
- Ostuni MA, Harbi-Iatmamen S, Teboul D, Robert JC, Lacapere JJ (2010) Characterization of membrane protein preparations: measurement of detergent content and ligand binding after proteoliposomes reconstitution. *Method Mol Biol* 654:3–18

- Park SH, Bibhuti BB, Casgrande F, Tian Y, Nothnagel HJ, Chu M, Kiefe H, Maier K, De Angelis AA, Marassi FM, Opella SJ (2012) Structure of the chemokine receptor CXCR1 in phospholipids bilayers. *Nature* 491:779–784
- Parola AL, Stump DG, Pepper DJ, Krueger KE, Regan JW, Laird HE (1991) Cloning and expression of a pharmacologically unique bovine peripheral-type benzodiazepine receptor isoquinoline binding protein. *J Biol Chem* 268(21):14082–14087
- Papadopoulos V, Miller WL (2012) Role of mitochondria in steroidogenesis. *Best Pract Res Clin Endocrinol Metab* 26(6):771–790
- Papadopoulos V, Baraldi M, Guilarte TR, Knudsen TB, Lacapère JJ, Linemann P, Noenberg MD, Nutt D, Weizman A, Ahang MR, Gavish M (2006) Translocator protein (18 kDa): new nomenclature for the peripheral-type benzodiazepine receptor based on its structure and molecular function. *Trends Pharmacol Sci* 27(8):402–409
- Pebay-Peyroula E (2007) Biophysical analysis of membrane protein. Wiley-VCH, Hoboken
- Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD (2012) The Pfam protein families database. *Nucleic Acids Res* 40(D1):D290–D301
- Rigaud JL, Pitard B, Levy D (1995) Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *Biochim Biophys Acta* 1231:223–246
- Riond J, Laplatois P, Laurent P, le Fur G, Caput D, Loison G, Ferrara P (1991) Expression and pharmacological characterization of the human peripheral-type benzodiazepine receptor in Yeast. *Eur J Pharmacol* 208(4):307312
- Robert JC, Lacapere JJ (2010) Bacterial overexpressed membrane proteins: an example, the TSPO. *Method Mol Biol* 654:29–45
- Rone MB, Midzak AS, Issop L, Rammouz G, Jagannathan S, Fan J, Ye X, Blonder J, Veenstra T, Papadopoulos V (2012) Identification of a dynamic mitochondrial protein complex driving cholesterol import, trafficking, and metabolism to steroid hormones. *Mol Endocrinol* 26(11):1868–1882
- Ruppecht R, Papadopoulos V, Rammes G, Baghai TC, Fan J, Akula N, Groyer G, Adams D, Schumacher M (2010) Translocator protein (18 kDa) (TSPO) as a therapeutic target for neurological and psychiatric disorders. *Nat Rev Drug Discov* 9:971–988
- Sagan S, Bolbach G (2009) Analyse des protéines membranaires par spectrométrie de masse: toujours un amour impossible? *Spectra Anal* 270:19–21
- Sprengel R, Werner P, Seeburg PH, Mukhin A, Santi RS, Grayson DR, Guidotti A, Krueger KE (1989) Molecular cloning and expression of cDNA encoding a peripheral-type benzodiazepine receptor. *J Biol Chem* 264(43):20415–20421
- Sreerama N, Woody RW (2004) On the analysis of membrane protein circular dichroism spectra. *Protein Sci* 13:100–112
- Tate CG, Schertler GFX (2009) Engineering G protein-coupled receptors facilitate their structure determination. *Curr Opin Struct Biol* 19:386–395
- Teboul D, Beaufile S, Taveau JC, Iatmanen-Harbi S, Renault A, Venien-Bryan C, Vié V, Lacapere JJ (2012) Mouse TSPO in a lipid environment interacting with a functionalized monolayer. *Biochim Biophys Acta* 1818:2791–2800
- Verma A, Nye JS, Snyder SH (1987) Porphyrins are endogenous ligands for the mitochondrial (peripheral-type) benzodiazepine receptor. *Proc Natl Acad Sci U S A* 84:2256–2260
- Vigano C, Smeyers M, Raussens V, Scheirlinckx F, Ruyschaert JM, Goormaghtigh E (2004) Hydrogen-deuterium exchange in membrane proteins monitored by IR spectroscopy: a new tool to resolve protein structure and dynamics. *Biopolymers* 74(1-2):19-26
- Weingarth M, Demco DE, Bodenhausen G, Tekely P (2009) Improved magnetization transfer in solid-state NMR with fast magic angle spinning. *Chem Phys Lett* 469:342–348
- Weingarth M, Tekely P, Brüschweiler R, Bodenhausen G (2010) Improving the quality of 2D solid-state NMR spectra of microcrystalline proteins by covariance analysis. *Chem Commun* 46:952–954
- Wendler G, Lindemann P, Lacapère JJ, Papadopoulos V (2003) Protoporphyrin IX binding and transport by recombinant mouse PBR. *Biochem Biophys Res Commun* 311(4):847–852

- Yeliseev AA, Kaplan S (1995) A sensory transducer homologous to the mammalian peripheral-type benzodiazepine receptor regulates photosynthetic membrane complex formation in *Rhodobacter sphaeroides* 2.4.1. *J Biol Chem* 270(36):21167–21175
- Yeliseev AA, Kaplan S (1999) A novel mechanism for the regulation of photosynthesis gene expression by the TspO outer membrane protein of *Rhodobacter sphaeroides* 2.4.1. *J Biol Chem* 274(30):21234–21243
- Yeliseev AA, Kaplan S (2000) TspO of rhodobacter sphaeroides. A structural and functional model for the mammalian peripheral benzodiazepine receptor. *J Biol Chem* 275(8):5657-67

Erratum

Foundations of Biomolecular Simulations: A Critical Introduction to Homology Modeling, Molecular Dynamics Simulations, and Free Energy Calculations of Membrane Proteins

Jérôme Hénin, Marc Baaden and Antoine Taly

Labratorire de Biochimie Théorique,
IBPC, CNRS, UPR9080, Univ Paris Diderot,
Sorbonne Paris Cité, 13 rue Pierre et Marie Curie, 75005, Paris France

I. Mus-Veteau (ed.), *Membrane Proteins Production for Structural Analysis*,
DOI 10.1007/978-1-4939-0662-8_13, © Springer Science+Business Media New York 2014

DOI 10.1007/978-1-4614-9642-7_15

The Publisher regrets that in chapter 13 the author sequence is incorrect i.e. Marc Baaden, Jérôme Hénin and Antoine Taly. The correct sequence is Jérôme Hénin, Marc Baaden and Antoine Taly.

.....
The online version of the original book can be found at
http://dx.doi.org/10.1007/978-1-4614-9642-7_13

I. Mus-Veteau (ed.), *Membrane Proteins Production for Structural Analysis*,
DOI 10.1007/978-1-4939-0662-8_15, © Springer Science+Business Media New York 2014

Index

A

- Amphipol (APol) 81, 315
 - A8–35 56
 - advantages 325
 - applications 179
 - bound, measuring the amount of 185, 186
 - drawbacks 325
 - folding in 79
 - generalities 323
 - illustrations 323, 324
 - library of 176, 177
 - membrane protein
 - cell-free expression of 194
 - folding of 190, 191
 - immobilization of 196
 - MP/APol complexes, properties of 177, 178, 179
 - properties in aqueous solution 174, 175, 176
 - protein trapping in 181, 182
 - best MP/APol ratio identification 185
 - detergent removal 184, 185
 - determination of protein concentration 183
 - MP/APol mass ratio determination 183
 - preparation of stock solution 182, 183
- Analytical Ultracentrifugation (AUC) 177, 189, 268
 - and SEC/MALS 270
 - fluorescence in 273
 - interference in 272
- Artificial hydrophobic environments 57
- Asymmetry 256
 - lipid 173
 - creation of 258
 - loss of 151
 - transbilayer phospholipid 150, 151

B

- Bacterial expression 72, 81
- Bibliographic analysis 100
- Bound detergent 174, 272, 274, 276, 277, 283, 285
 - estimate of
 - and interference in AUC 272
 - and refractive index in SEC/MALS 272, 273

C

- Ca²⁺ transport 141, 149
- Cell-free expression (CFE) 181, 194
 - large-scale reaction 195
 - small-scale reaction 194, 195
- Circular dichroism (CD) 80, 193, 398, 400, 402
- Comparative modeling 347, 349, 350
- Co-translational solubilization
 - of CF expressed IMPs 53
- Crystallization
 - LCP 295, 296, 297
 - of membrane proteins 290, 292, 293
 - robot, LCP 304
 - trials, initial 407
- Cys-loop family 363

D

- Detergents
 - advantages 319
 - drawbacks 319, 320
 - effects of 409
 - folding of 77, 78
 - generalities 317
 - illustrations 317, 319
 - solubilization and stability in 160, 161, 163
- Dynamic light scattering (DLS) 268, 272, 273, 274, 276, 283

E

- E. coli 48, 74, 81, 93, 323
 - functional expression of GPCRs in 72, 73
- Electron microscopy (EM) 102
 - and 2D crystals 412, 415
 - Torpedo structures 359
- Escherichia coli *See* E. Coli 72

F

- FhuA
 - in F6-DigluM 280
 - in LDAO 283, 284
 - monomeric 281
 - sedimentation of 281
- Fluorinated surfactants (FS) 53
- Free energy calculations 349, 366, 376

G

- G Protein Coupled Receptor (GPCR)
 - functional expression of, in E. Coli 72, 73
 - in IBs, accumulation of 74
 - in vitro folding of 75, 76
 - purified, in vitro stabilization of 80, 81
 - solubilizing, from IBs 74, 75

H

- Heterogeneity 186, 256, 257, 258, 261, 263
 - of MP/APol complexes 179
- Heterologous expression 148, 163, 290, 395

I

- In-cell NMR 315, 337
- Interactions
 - electrostatic 179
- Intrinsic fluorescence 398, 400, 403, 404
- Ion channels 253, 289, 293
- Ion channels *See also* Ligand-gated ion channels (LGICs) 353
- Isothermal titration calorimetry (ITC) 400

L

- Ligand-gated channels 348
- Ligand-gated channels *See also* Ligand-gated ion channels (LGICs) 375
- Ligand-gated ion channels (LGICs) 348, 350, 353, 354
- Lipids 137
 - folding in 78
 - LCP host 297, 298
- Lipid transport 151, 153, 163
- Liposome 254, 255, 257
 - fusion 260, 261

M

- Mass spectroscopy 406
- Membrane mimetics 315, 334, 336
- Membrane proteins (MPs)
 - amphipol-mediated immobilization of 196
 - crystallization of 290, 292, 293
 - quality control of 61
- Molar mass 270, 272, 273
 - buoyant 275, 276
 - measure of, in SEC.MALS 273, 274
- Molecular dynamics (MD) 175, 259, 358
 - simulations 362–366, 368, 369

N

- Nanodisc 61, 79, 80, 81, 320, 326, 334
- NMDA receptor
 - building models of 356–358
- Normal mode analysis (NMA) 358–360
- Nuclear magnetic resonance (NMR)
 - in-cell 315, 337
 - solid state 328, 329
 - solution-state 317, 319, 320

P

- P2X receptor
 - structure of 351
- PK 11195 400, 402, 404, 407, 415
 - high-affinity drug ligand 398, 410
 - isoquinoline 394
- Production
 - of labeled TSPO 405, 406
 - of proteomicelles, D-CF expression 53, 56, 57
- Protein structure 335, 400, 404, 409, 417
- Proteoliposomes 254, 403, 404, 412
- Proteolysis 62, 74, 134, 395, 401
- Protocol
 - for applications of APols 181–198
 - for improving CF expression efficiency 48, 49
- P-type ATPases 141
 - crystallization of 160
 - crystallogenesis of 148
 - inhibitor of 157
 - transport cycle of 141

R

- Recombinant expression 146
- Reconstitution 253, 295, 327, 332
 - methods 254
- Refolding 73, 77, 78, 79, 80
 - in vitro, expression of GPCRs in 74–76

S

- Saccharomyces cerevisiae 88, 89, 133
 - as a host for heterologous expression of membrane proteins 134, 137
 - expression in 141, 142
 - serca1a expression in, optimization of 143, 145
- SEC/MALS 272
 - FhuA in LDAO 283, 284
 - light scattering in 273
 - measure of molecular masses in 273, 274
 - refractive index in 272, 273
- Sedimentation velocity (SV) 268, 270
 - theoretical background 274
- Serial femtosecond crystallography (SFX)
 - LCP 307, 308
- Solubilization
 - buffer, volume of 50
 - of MPs in denaturing conditions 191, 192
 - temperature of 50
- Stabilization
 - in vitro, of purified GPCRs 80, 81
 - protein 403, 404, 405
- Structural biology 90
- Structural characterization 148
- Surfactants 174, 177, 181

T

- T7 RNA polymerase 91, 92, 94
- Translocator protein (TSPO) 393, 394
 - labeled
 - production 405
 - recombinant
 - characterization 395, 398
 - production 395
 - purification 395, 398
 - structural studies
 - molecular modeling 415, 417
 - NMR studies 409, 410, 412
 - three-dimensional crystals 406, 407, 409
 - two-dimensional crystals and electron microscopy 412, 414, 415

X

- X-ray crystallography 62, 253, 289, 332, 337, 354, 415
- X-ray diffraction
 - by three-dimensional crystals 253
- X-ray free electron lasers (XFEL) 307