

ULRR

Crystallization and structure determination of gramicidin D and the cytosolic domain of CzcB

Item Type	Thesis
Authors	Höfer, Nicole
Download date	2026-04-15 12:22:38
Item License	https://creativecommons.org/licenses/by-nc-sa/1.0/
Link to Item	https://hdl.handle.net/10344/3282

Crystallization and Structure Determination of Gramicidin D
and the Cytosolic Domain of CzcB

Nicole Höfer



UNIVERSITY of LIMERICK

OLLSCOIL LUIMNIGH

Title

Crystallization and Structure Determination of
Gramicidin D and the Cytosolic Domain of CzcB

Author

Nicole Höfer M.S.

Supervisors

Dr. Tewfik Soulimane, Prof. Martin Caffrey

Submitted to the University of Limerick
in partial fulfilment of the requirements for a
Doctor of Philosophy
2013

Abstract

Crystallization of membrane proteins and peptides represents a challenge in the field of structural biology. Lipidic cubic phase (LCP) has become an important medium for crystallogensis of membrane proteins of different molecular weight. Here, the small membrane peptide gramicidin is used as an example peptide to test if LCP can produce diffraction quality crystals for membrane proteins and peptides in the lower molecular weight range. This approach was initially tested with the standard LCP lipid monoolein and later on extended to a variety of different monoacylglycerols varying in their acyl chain length. Data sets for three different crystal forms were obtained. In each case gramicidin was found in the double stranded double helical (DSDH) conformation. One crystal form shows stabilizing hydrogen bonds between adjacent tryptophan residues indicating how DSDH can be stabilized as an aggregate in the membrane.

The cytoplasmic domain of the putative zinc transporter CzcB was solved in the apo and zinc-bound forms. NMR, X-ray scattering, and size-exclusion chromatography provide support for dimer formation. Full-length variants of CzcB in the apo and zinc-loaded states were generated by homology modeling with the $\text{Zn}^{2+}/\text{H}^{+}$ antiporter YjiP.

Acknowledgement:

First I would like to thank my supervisors, Dr. Tewfik Soulimane and Prof. Martin Caffrey for their help and support.

I would like to especially thank Dr Soulimane, Dr. Aragao and Dr. Pye for their support in the final stages of my Ph.D.

A warm thanks goes to Dr. David Aragao for introducing me to crystallography and helping me in the development of my work. Much of this thesis was facilitated by his guidance.

Furthermore, I would like to thank Dr. Cherezov, Dr. Boyle-Roden and Prof. Jaroniec for supervising me at the Ohio State University.

I would like to thank the members of the Membrane Structural and Functional Biology Group, that is now housed at Trinity College Dublin for their support throughout the years.

I would like to thank my mother-in-law, Patricia Porter, who travelled repeatedly to Ireland from Columbus, Ohio to babysit our baby daughter Amalia. A special thanks goes to my husband Dr. Christopher D. Porter for support and encouragement. Thanks also go to Amalia, Elena and my parents Monika and Wolfgang Höfer for their continued motivation.

Finally, I extend my heartfelt gratitude to my committee members for taking the time and energy they devoted to my defense and in reviewing my thesis itself.

List of Publications:

Höfer N, Aragão D, Lyons J, Caffrey M. Membrane Protein Crystallization in Lipidic Mesophases. Hosting lipid affects on the crystallization and structure of a transmembrane peptide J Crystal Growth and Design. 2011. 11(4): 1182-1192.

Höfer N, Aragão D, Caffrey M. Crystallizing transmembrane peptides in lipidic mesophases. Biophys J. 2010. 99(3): L23-25.

Höfer N, Kolaj O, Li H, Cherezov V, Gillilan R, Wall JG, Caffrey M. Crystallization and preliminary X-ray diffraction analysis of a soluble domain of the putative zinc transporter CzrB from *Thermus thermophilus*. Acta Crystallogr Sect F. 2007. 63: 673-677.

Cherezov V, Höfer N, Szebenyi DM, Kolaj O, Wall JG, Gillilan R, Srinivasan V, Jaroniec CP, Caffrey M. Insights into the mode of action of a putative zinc transporter CzrB in *Thermus thermophilus*. Structure. 2008. 16(9): 1378-1388.

Boyle-Roden E, Höfer N, Dey KK, Grandinetti PJ, Caffrey M. High resolution ¹H NMR of a lipid cubic phase using a solution NMR probe. J Magn Reson. 2007. 189(1): 13-19.

Mukherjee S, Pondaven SP, Höfer N, Jaroniec CP. Backbone and side-chain (¹H), (¹³C) and (¹⁵N) resonance assignments of LEN, a human immunoglobulin kappaV light-chain variable domain. Biomol NMR Assign. 2009. 3(2): 255-259.

Helmus JJ, Nadaud PS, Höfer N, Jaroniec CP. Determination of methyl ¹³C-¹⁵N dipolar couplings in peptides and proteins by three-dimensional and four-dimensional magic-angle spinning solid-state NMR spectroscopy. J Chem Phys. 2008. 128(5):052314.

Nadaud PS, Helmus JJ, Höfer N, Jaroniec CP. Long-range structural restraints in spin-labeled proteins probed by solid-state nuclear magnetic resonance spectroscopy. J Am Chem Soc. 2007. 129(24): 7502-7503.

PDB entries: 2XDC, 2ZQ8, 2Y5M, 2Y6N

Declaration

I hereby declare that this thesis is entirely my own work, and has not been submitted for any other awards at this or any other academic institution. Where use has been made to the work of other people it has been fully acknowledged and referenced.

Nicole Höfer

Date

Table of Contents

Title	I
Abstract	III
Acknowledgements	V
List of publications	VII
Declaration	IX
Table of contents	XI
List of figures	XV
List of tables	XVII
Abbreviations	XIX
Chapter 1 – Crystallization and Structure Determination of Gramicidin	1
1.1 Lipidic cubic phase	5
1.1.1 Introduction to lipidic cubic phases	5
1.1.2 Crystallization of membrane proteins	7
1.1.2.1 Lipidic cubic phases	7
1.1.2.1.1 Lipidic crystalline phases	7
1.1.2.1.2 Description of lipidic bicontinuous cubic phases	9
1.1.2.2 Lipidic cubic phase forming lipids	10
1.1.3 The <i>in meso</i> crystallization method	12
1.1.3.1 Development of this method	12
1.1.3.2 Advantages of the <i>in meso</i> method	14
1.1.3.3 Proteins crystallized by the <i>in meso</i> method	15
1.1.3.4 Proposed crystallization mechanism	20
1.1.3.5. Type I crystal packing	21
1.1.4 Crystallization of small proteins and peptides using lipidic cubic phases	23
1.1.4.1 Gramicidin D as test peptide	24
1.2 The small membrane peptide gramicidin	26
1.2.1 Introduction	26
1.2.2 Sequence of gramicidin D	27
1.2.3 Expression of gramicidin D in <i>Bacillus brevis</i>	29
1.2.4 Functions of gramicidin D	30
1.2.5 Conformations adopted by gramicidin D	31
1.2.6 Proposed mechanism of HSH (ion channel) conformation	36
1.2.7 Proposed mechanism of the DSDH conformation	37
1.2.8 Influence of tryptophan residues on conformation and ion	

channel activity in gramicidin D	38
1.2.9 Sensitivity of gramicidin D towards its surrounding environment ..	40
1.2.10 Gramicidin and the lipidic cubic phase	41
1.3 Materials and methods	44
1.3.1 Materials	44
1.3.2 Methods	44
1.3.2.1 Solubility of gramicidin in precipitant solutions containing PEG	44
1.3.2.2 Gramicidin/lipid mixture preparation	45
1.3.2.3 Cubic phase preparation and in meso crystallization ..	46
1.3.2.4 Crystal harvesting	47
1.3.2.5 Data collection	48
1.3.2.5.1 Data collection for CF1	48
1.3.2.5.2 Data collection for CF2	50
1.3.2.5.3 Data collection for CF3	50
1.3.2.6 Phase determination and refinement	51
1.3.2.6.1 Phase determination and refinement for CF1	51
1.3.2.6.2 Phase determination and refinement for CF2	53
1.3.2.6.3 Phase determination and refinement for CF3	53
1.3.2.7 Structure validation of CF 1-3	54
1.4 Results	56
1.4.1 Crystallization and structure determination of CF1	56
1.4.2 Crystallization and structure determination of CF2	60
1.4.2.1 Ligand	61
1.4.2.1 Tryptophan residues and their interactions	61
1.4.3 Crystallization and structure determination of CF3	63
1.4.3.1 Description of bound lipid molecules	64
1.4.3.2 Interaction between tryptophan residues	66
1.4.4 Comparison between <i>in meso</i> grown crystal structures of Gramicidin	69
1.4.4.1 DSDH conformation observed in all three crystal forms	69
1.4.4.2 Crystal packing	71
1.4.4.3 Crystallographic water molecules	72
1.4.4.4 Solvent content of <i>in meso</i> grown crystals	72
1.5 Discussion	75
1.5.1 Crystallization of gramicidin D from LCP of different lipids	75
1.5.2 Crystallization mechanism	75
1.5.2.1 Three hosting lipids, one crystal form in case of CF1...	75
1.5.2.2 A role for PEG	78
1.5.2.3 An alternative crystallization mechanism?	82
1.5.2.4 Crystallizing “small proteins” <i>in meso</i> is worthy of consideration	83
1.5.2.5 Functional significance of the DSDH conformation	85
1.5.3 Do tryptophan interactions stabilize aggregates?	86
1.5.4 Similar origin of all in meso gramicidin crystal forms	88

1.6 Conclusion and outlook	89
1.7 References	91
Chapter 2 – Crystallization and Structure Determination of the cytosolic domain of the putative zinc-transporter CzcB from <i>Thermus Thermophilus</i>	103
2.1 Introduction	107
2.2 Material and methods	114
2.2.1 Materials	114
2.2.2 Cloning and expression methods	115
2.2.3 Protein purification	116
2.2.4 Crystallization	118
2.2.4.1 6x His-tag sf-CzcB crystallization	118
2.2.4.2 sf-CzcB crystallization	119
2.2.5 Crystallization, X-ray diffraction, structure determination, and refinement	121
2.2.6 Circular dichroism	124
2.2.7 SAXS	124
2.2.7.1 Data collection	124
2.2.7.2 Data analysis	125
2.2.8 Size-exclusion chromatography	126
2.2.9 NMR spectroscopy	126
2.2.10 Multiple alignment and homology modeling	126
2.3 Results	128
2.3.1 Crystallization	128
2.3.2 Crystal structures	128
2.3.2.1 Apo-CzcB _{sf}	129
2.3.2.2 Zinc-CzcB _{sf}	131
2.3.2.3 Zinc coordination sites	132
2.3.3 Characteristics of the protein in solution	136
2.3.3.1 Size-exclusion chromatography	136
2.3.3.2 Small-angle X-ray scattering	138
2.3.3.3 ¹ H NMR spectroscopy	138
2.3.3.4 Circular dichroism	140
2.4 Discussion	141
2.4.1 Zinc binding	141
2.4.2 Homology modelling and mode of action	144
2.5 Summary, conclusions and outlook	149
2.6 References	150

List of Figures

Figure 1.1.1	Number of membrane protein entrees into the MPDB	6
Figure 1.1.2	Phase diagram of monoolein	9
Figure 1.1.3	Schematic of 9.7 MAG (monopalmitolein)	11
Figure 1.1.4	Number of published LCP structures in MPDB	13
Figure 1.1.5	Schematic of proposed crystallization mechanism	20
Figure 1.1.6	<i>In meso</i> crystal packing arrangement and molecular structures of membrane proteins	21
Figure 1.1.7	Molecular weight range of different MPs crystallized <i>in meso</i>	23
Figure 1.2.1	Schematic representation of major gramicidin conformations	27
Figure 1.2.2	Hydrogen bonding in the HSHH conformation	32
Figure 1.2.3	Conformations adopted by linear gramicidin	34
Figure 1.2.4	Space-filling model in top and side view of the HSHH conformation	39
Figure 1.2.5	Sensitivity of the cubic-Pn3m phase of hydrated 9.9 MAG to gramicidin D	43
Figure 1.3.1	Crystals of gramicidin growing <i>in meso</i> using different hosting lipids.	47
Figure 1.4.1	Schematic representation of hydrogen bonding (dashed lines) between oxygen atoms (red, the subscript identifies the ethylene glycol unit number) of PEG-A and tryptophan N ϵ atoms (blue) of gramicidin.	58
Figure 1.4.2	Schematic representation of hydrogen bonding between oxygen atoms of PEG and tryptophan N ϵ atoms of gramicidin	62
Figure 1.4.3	Representation of the interaction of tryptophan residues in N-H $\cdots\pi$ configuration	65
Figure 1.4.4	Interactions between two tryptophan residues observed in crystal structures of proteins.	67

Figure 1.4.5	Layered (Type I) packing observed in crystals of gramicidin in different crystal forms grown in the lipidic mesophases	71
Figure 1.4.6	Solvent content (%) of gramicidin crystals grown from organic solvents	73
Figure 2.1.1	Sequence Alignment of CzcB from <i>Thermus thermophilus</i> with CDF Proteins from Other Organisms that Are Involved in Zinc Transport.	108
Figure 2.1.2	Ribbon representation of YiiP	110
Figure 2.1.3	Alternating access mechanism for Zn^{2+}/H^{+} antiport by YiiP	112
Figure 2.2.1	SDS-PAGE of sf-CzcB at different stages during preparation. Lane 1, molecular-weight standard markers	118
Figure 2.2.2	Crystals of sf-CzcB with and without the 6x His-tag grown with and without added zinc.	120
Figure 2.3.1	Structural Rearrangements in CzcB _{sf} upon Zinc Binding C α traces of apo- (red) and zinc-CzcB _{sf} (blue).	129
Figure 2.3.2	Cartoon Representation of Secondary Structure Elements in the Zinc-CzcB _{sf} Model Colour coding is as follows	130
Figure 2.3.3	Surface Potential and Hydrophobic Core of Apo-CzcB _{sf}	132
Figure 2.3.4	Zinc Coordination at the Dimer Interface in CzcB _{sf}	134
Figure 2.3.5	Characterization of the Solution Properties of CzcB _{sf}	137
Figure 2.3.6	Small-Angle X-Ray Scattering from CzcB _{sf} in Solution with and without Added Zinc	139
Figure 2.4.1	Overlay of Zinc-CzcB _{sf} and YiiP at Homologous Zinc Binding Sites	142
Figure 2.4.2	An Overlay of the C α Trace of Zinc-CzcB _{sf} and YiiP in the Region That Includes Strands β 1 and β 2 and Loops L2 and L3 and for the Full-length Zinc-CzcB _{sf}	143
Figure 2.4.3	A Model for CzcB Action as a Zinc Transporter	147
Figure 2.4.4	A Model to Explain How CzcB _{sf} Can Render CzcB Inactive as a Zinc Transporter	148

List of Tables

Table 1.1.1	Monoacylglycerols	11
Table 1.1.2	Structures of membrane proteins produced by the <i>in meso</i> method	17
Table 1.2.1	Gramicidin structures available in the PDB	35
Table 1.3.1	<i>In meso</i> crystallization conditions in three hosting MAGs and gramicidin solubility in the corresponding precipitant solutions	45
Table 1.3.2	Diffraction data, refinement, and model statistics for data sets 1-3	49
Table 1.3.3	Diffraction data, refinement and model statistics for CF2, CF3	52
Table 1.3.4	Structure validation	55
Table 1.4.1	Hydrogen bonds between lipid headgroups of MPG1, MPG2, MPG3 and gramicidin or water; distances between 2.8 and 3.6 Å were considered	65
Table 1.4.2	Hydrogen bonds formed by tryptophan residues in dimers AB and CD of CF3	68
Table 1.4.3	R.M.S.D. values between main chain backbone atoms of the closed DSDH conformation of gramicidin dimers from <i>in meso</i> and organic solvent structures	70
Table 1.4.4	Crystal form, solvent content, Matthews coefficient (V_m), space group and unit cell dimensions of gramicidin crystals grown from various organic solvents	74
Table 1.5.1	Possible scenarios for the way in which gramicidin crystallizes in mesophases formed with different MAGs	77
Table 2.2.1	Data Collection and Processing Statistics	122
Table 2.3.1	Geometry of Zinc-Binding Residues in C _{zr} B _{sf}	135
Table 2.3.2	Results of Analysis of CD Spectra from Apo-C _{zr} B _{sf}	140

Abbreviations

BLM	black lipid membrane
CD	circular dichroism
CF	crystal form
CF1	crystal form 1
CF2	crystal form 2
CF3	crystal form 3
D	dextrorotatory
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DSDH	double stranded double helical conformation
gD	gramicidin D, mixture of gA, gB, gC
EM	electron microscopy
gA	gramicidin A
gB	gramicidin B
gC	gramicidin C
GMCA CAT	General Medicine and Cancer Institutes, Collaborative Access Team
GPCR	G-protein coupled receptor
HSH	head to head single stranded helical conformation
IPTG	β -D-hiogalactopyranoside
KcsA	potassium crystallographically-sited activation channel
L	levorotatory
LB	Luria-Bertani broth
LH	left-handed
LCP	lipidic cubic phase also referred to as <i>in meso</i>
MAG	monoacylglycerol
MC	Mathews coefficient
MP	membrane protein
MPD	2-methylpentane-2,4-diol
MPDB	Membrane Protein Data Base
MX	macromolecular crystallography
NMR	nuclear magnetic resonance
NRPS	non-ribosomal peptide synthases
PDB	Protein Data Bank
PDB ID	Protein Data Bank Identification Code
PEG	polyethylene glycol
RH	right-handed
r.m.s.d.	Root mean square deviation
SAXS	small angle X-ray scattering
SEC	size exclusion chromatography
SDS	sodium dodecyl sulfate
TFE	trifluoroethanol
TM	transmembrane

Chapter 1

Crystallization of Gramicidin D from Lipidic Cubic Phases

The content of this chapter was published prior to this thesis in the following two publications:

Nicole Höfer, David Aragão, Joseph A. Lyons, Martin Caffrey, Membrane Protein Crystallization in Lipidic Mesophases. Hosting Lipid Effects on the Crystallization and Structure of a Transmembrane Peptide, 2011, Journal of Crystal Growth and Design, Volume 11, 1182 – 1192.

Nicole Höfer, David Aragão, Martin Caffrey, Crystallizing transmembrane Peptides in lipidic cubic phases, 2010, Biophysical Journal 99, L23-L25.

Text, figures and tables have been copied.

1.1 Lipidic Cubic Phase

1.1.1 Introduction

The field of membrane protein (MP) structure determination has advanced considerably since the first structure of a membrane protein, the photosynthetic reaction center by J. Deisenhofer, H. Michel and M. Huber was determined (Deisenhofer et al., 1985). This structure was determined by X-ray crystallography to a resolution of 2.8 Å. Preceding the photosynthetic reaction center was a model of bacteriorhodopsin (bR) determined by electron microscopy (EM) (Henderson and Unwin, 1975) showing basic structural elements like transmembranal helices. These two structures proved that MPs can be crystallized and their structure can be determined. As of March 8th of 2013 there have been 386 unique membrane protein structures available in the protein data bank (PDB) (sourced from membrane proteins of known structure from Stephen White; <http://blanco.biomol.uci.edu/mpstruc/listAll/list>). The MPDB (Membrane Protein Data Base; www.mpdb.trinity.ie) (Raman et al., 2006) lists 1246 PDB entries at end of April 2012 (figure 1.1.1). The database ‘membrane proteins of known structure’ included only unique structures, whereas the MPDB lists multiple successful solutions of a given structure.

Over the years there were several landmark structures that represented milestones in the field of MP research. One of them was the structural characterization of the ion channel structure KcsA from *Streptomyces lividans* as the first ion channel structure, determined by the McKinnon group (Doyle et al., 1998). It should be noted that this structure was preceded by the ion channel structure of the small membrane peptide gramicidin A (Lomize et al., 1992). This initial structure of gramicidin A in ion channel conformation (see 1.2 for details) was published by Lomize et al., 1992 ((Lomize et al., 1992) using solution state nuclear magnetic resonance (NMR) in detergent micelles of sodium dodecyl sulfate (SDS). The validity of this structure was questioned mainly because SDS micelles are not membranes and the conformation of

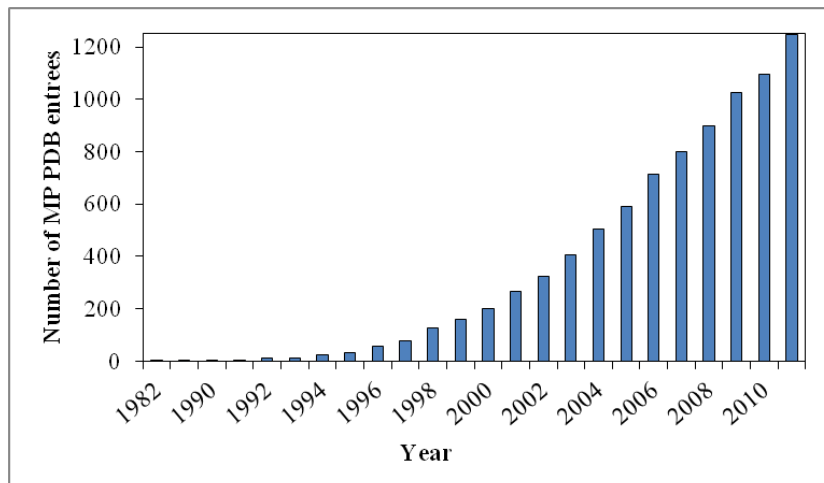


Figure 1.1.1: Number of membrane protein entries into the MPDB (Raman et al., 2006) per year. The entrees are by the end of April 2012.

gramicidin is known to be highly sensitive to its solvating medium. T. Cross and coworkers later determined the structure in dimyristoylphosphatidylcholine (DMPC) membranes by solid state NMR (Ketchum et al., 1993) proving the validity of the previously published structure. With the crystal structure from KcsA it became clear that the ion transporting pores in both structures are built by the peptide backbone and that gramicidin serves as an excellent study object to understand basic ion channel function (Yu and Pomes, 2003).

In recent years the structure determination of G protein coupled receptors (GPCR) has received wide attention. The first structure was published in 2007 by the B. Kobilka group in conjunction with the R. Stevens group (Cherezov et al., 2007). In 2011 the structure of an active GPCR was determined by the Kobilka group to an overall resolution of 3.2 Å (Rasmussen et al., 2011). With the knowledge of GPCR structures, their mechanism has become clearer and structural details can guide drug development.

About one third of the human genome encodes membrane proteins (Nugent and Jones, 2009). Yet by comparing the number of MPs compared to the number of soluble proteins structures in the PDB the vast majority of crystal structures originate from soluble proteins. In a sense this comparison reflects the difficulties involved in

obtaining a structure of a membrane protein. Over the years the ability to improve expression, purification, stabilization of MP has increased (Stroud, 2011). The focus in this chapter is on the crystallization of this type of protein, particularly on improvement of the *in meso* method.

1.1.2 Crystallization of Membrane Proteins

A variety of methods have been developed to crystallize MPs. According to the MPDB (www.mpdb.trinity.ie) (Raman et al., 2006) the most popular method is the crystallization of the MP from detergent micelle. In this mode crystals are produced by either vapor diffusion via hanging drop or sitting drop method or by batch crystallization. While these methods have been very successful and are the leading methods in producing crystals of MPs and subsequently protein structures of MPs, they do not mimic a natural membrane.

Over the past years crystallization methods involving lipidic bilayers have produced a number of successful crystal structures (Cherezov, 2011b). The most prominent one is the *in meso* method that uses lipidic cubic phases to embed membrane proteins (Caffrey, 2011).

1.1.2.1 Lipidic Cubic Phases

1.1.2.1.1 Lipidic Crystalline Phases

Lipidic crystalline phases are liquid crystals that have properties of a liquid and of a crystal, also referred to as mesophases (Caffrey, 2003). They are observed between a pure solid and a pure liquid phase (Caffrey, 2003). Lipidic crystalline phases are lyotropic meaning that a phase transition from one lipidic crystalline state into another is dependent on two functions; temperature and concentration (composition). Concerning the composition it means that at least two components (here lipid and water) need to be combined in order to form this phase. Lipids form liquid crystals because of their amphiphilic nature; they consist in their most simple representation of a hydrophilic headgroup and a hydrophobic tail. These two components (hydrophobic and hydrophilic) cannot be mixed and therefore tend to segregate. The

temperature – composition phase diagram describes different liquid crystalline phases best (figure 1.1.2). Depending on the nature of such a lipid a number of different liquid crystalline phases can be observed. The driving force to form different liquid crystalline phases is to minimize the exposure of hydrophobic lipid chains from the aqueous environment. Using the example of the lipid monoolein (9.9 MAG) different stages of liquid crystalline phases will be described (Qiu and Caffrey, 2000). In figure 1.1.2 different mesophases are described along the 20 °C isotherm with increasing concentration of water showing possible lipid polymorphism. The phase diagram for monoolein describes the composition of lipid to water in the X direction and the temperature in the Y direction. The different phases typically exist across a range of temperatures and compositions, as opposed to at only a single value. There are areas in which only one phase is present, but areas with multiple phases are also common (white regions). Starting at 0 %(w/w) water and 100 %(w/w) lipid the phase is referred to as liquid crystalline phase. The lipid is oriented in stacked bilayers. Upon heating, this sample will convert into the fluid isotropic phase, which is equivalent to a molten lipid. Hydrating the sample leads to water-induced mesophases starting with the lamellar phase at approximately 11 %(w/w) hydration. The lamellar phase is a set of stacked bilayers. The lipid is not as ordered as in the crystalline state. Water associates with the head group and forms sheets of ordered water. Upon increasing hydration, the phase changes into bicontinuous cubic phases of Ia3d and Pn3m. Even further hydration keeps the mesophase in the presence of excess water. If the temperature is increased and the composition kept constant an inverted hexagonal phase forms. This phase takes the shape of rods. Lipids are arranged around tubes of water channels.

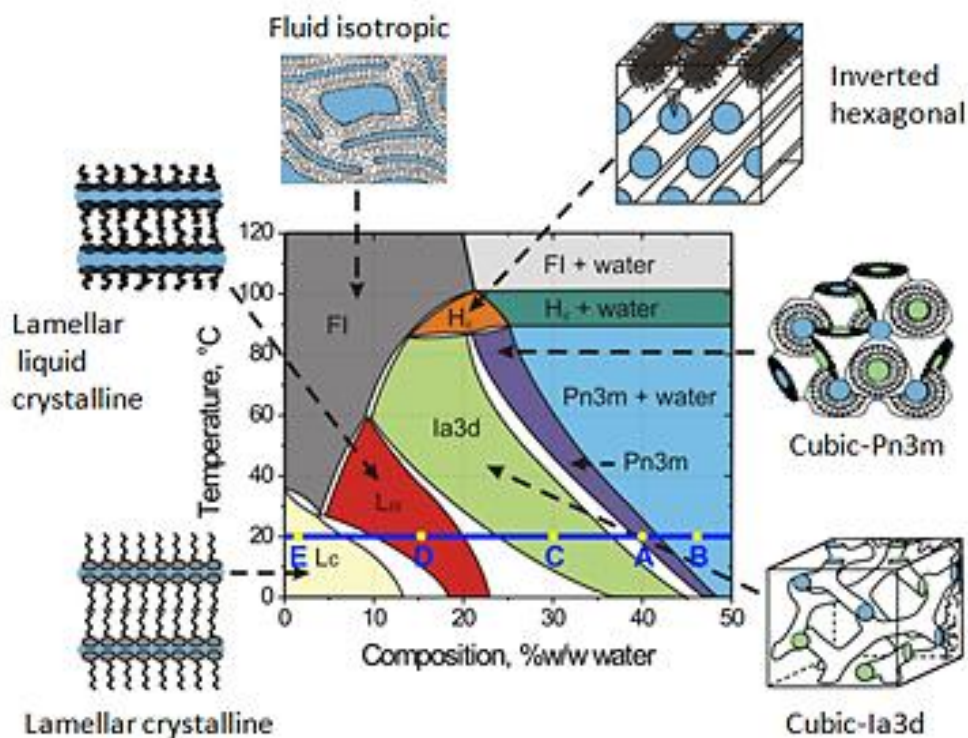


Figure 1.1.2: Phase diagram of monoolein (Qiu and Caffrey, 2000). The blue bar indicates the 20 °C isotherm. (Figure from Qiu and Caffrey, 2000).

1.1.2.1.2 Description of Lipidic Bicontinuous Cubic Phases

The observed lipidic cubic phases for monoacylglycerols water systems are Pn3m, Ia3d, and Im3m (Chung and Caffrey, 1995; Qiu and Caffrey, 2000). They are highly curved lipidic bilayers that permeate three dimensional space. Mathematically they can be described as triply periodic minimal surfaces that have zero curvature at any point (Anderson et al., 1988; Hyde, 1996). A minimal surface is a space filling geometry that has a locally minimal surface area. It is not a globally minimal surface area, but any deformation of the surface would result in an increase in total surface area. The zero mean curvature of the surface does not mean that there are no curves in the surface. On the contrary, these minimal surfaces are highly curved. The zero mean curvature means the following. For any point on the surface, a normal line could be drawn, and the curvature of the surface as it radiates from that normal line could be calculated along any direction. The average of the curvatures along all directions

radiating from the normal line will be zero. The minimal surface divides the bilayer into two even components that are continuous (Anderson et al., 1988). For this reason the phase is referred to as being bicontinuous. To position the lipid onto the minimal surface the terminal methyl group of the lipid chain is located closest to it (Chung and Caffrey, 1994). Moving from the minimal surface in both directions along the lipid, closest to the minimal surface is the terminal methyl group. Moving further away towards the double bond the neutral surface is crossed (Chung and Caffrey, 1994). Further moving outwards is the headgroup of the lipid in the interface of the bilayer. The curvature is negative towards the water channel (Chung and Caffrey, 1994). What then follows on both sides of the minimal surface are the water channels.

1.1.2.2 Lipidic Cubic Phase forming Lipids

Of interest for biological research are lipids that form the lipidic cubic phase at physiological pressure, temperature and pH range. To date a number of lipids, mainly monoacylglycerols (table 1.1.1), are known that form lipidic cubic phase under these conditions. They consist of a glycerol head group that is connected to an acyl chain via an ester bond. The chain is kinked at a cis double bond separating the chain into a neck (N) and a tail (T). Monoacylglycerols (MAG) can be described by their NT nomenclature (N.T MAG) (Misquitta et al., 2004b). This is explained in the case of monopalmitolein also referred to as 9.7 MAG. The lipid chain has 16 carbon atoms that are separated by a cis double bond into 9 carbon atoms observed in the neck portion and 7 atoms in the tail portion (figure 1.1.3).

So far a number of MAGs are known to form the lipidic cubic phase (table 1.1.1). These lipids have been confirmed by small angle x-ray scattering (SAXS) to produce lipidic cubic phase. Myverol, a mixture of different MAGs with the main lipid being monoolein has also been characterized by SAXS. It shows almost identical phase behavior to monoolein. In recent years there has been an effort in producing synthetic MAGs that can form the lipidic cubic phase and have altered characteristics like a thinner bilayer (Li et al., 2011).

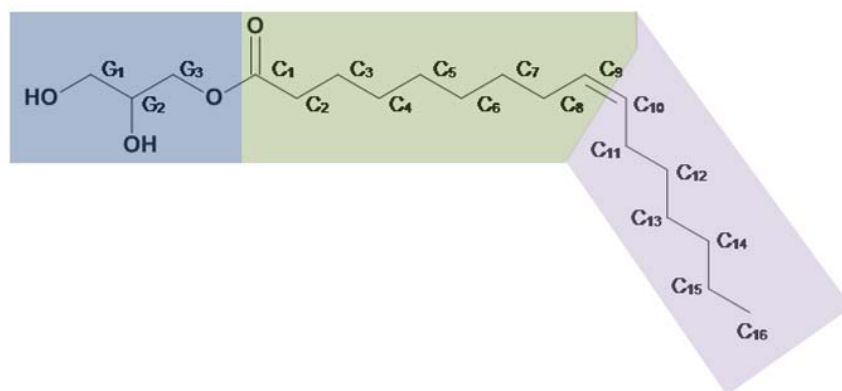


Figure 1.1.3: Schematic of 9.7 MAG (monopalmitolein). Blue indicates the glycerol head group; green describes the neck (N), and light purple indicates the tail (T). (Misquitta et al., 2004b)

Table 1.1.1: Monoacylglycerol

Common name / chemical name	N.T nomenclature (Lipid numbers)	MW	Reference
Monoeicosenoin 1-(11Z-eicosenoyl)-rac-glycerol	11.9 MAG (20:1c11)	384.6	(Misquitta and Caffrey, 2001)
Monovaccenin 1-(11Z-octadecenoyl)-rac-glycerol	11.7 MAG (18:1c11)	356.5	(Qiu and Caffrey, 1998)
Monoolein 1-(9Z-octadecenoyl)-rac-glycerol	9.9 MAG (18:1c9)	356.5	(Qiu and Caffrey, 2000)
Monopalmitolein 1-(9Z-hexadecenoyl)-rac-glycerol	9.7 MAG (16:1c9)	328.5	Thesis Jason Briggs
Monomyristolein	9.5 MAG (14:1c9)	300.5	(Briggs and Caffrey, 1994b)
Monopentadecenoin	10.5 MAG (15:1c10)	314.5	(Briggs and Caffrey, 1994a)
1-(7Z-tetradecenoyl)-rac-glycerol	7.7 MAG (14:1c7)	304.5	(Misquitta et al., 2004a)
1-(7Z-hexadecenoyl)-rac-glycerol	7.9 MAG (16:1c9)	328.5	(Misquitta et al., 2004b)
myverol	18-99K (mixture)	n.a	(Clogston et al., 2000)

1.1.3 The *in meso* Crystallization Method

The *in meso* method was introduced by Landau and Rosenbusch in 1996 with the structure determination of bacteriorhodopsin (bR) to high resolution (Landau and Rosenbusch, 1996; Pebay-Peyroula et al., 1997). Since then a number of membrane protein structures have been determined (figure 1.1.4; table 1.1.2). The *in meso* method uses a lipidic bilayer usually made of monoacylglycerols to host the membrane proteins (Caffrey, 2009). By homogenizing the proper ratio of lipid and protein solution the lipidic cubic phase forms (Caffrey, 2008). This phase is a lyotropic liquid crystal that forms a bilayer that is highly curved and bicontinuous in three dimensions. It has two interpenetrating but non-contacting aqueous channels (Caffrey, 2009). This lipidic cubic phase forms in specific regions of the temperature concentration phase diagram. A number of different MAGs are available in which the lipidic cubic phase can be tailored to specific needs of the crystallization experiment (Cherezov et al., 2002; Misquitta et al., 2004a; Misquitta et al., 2004b). However, a great many additional advancements were required before the *in meso* method could become a widely accessible method of MP crystallization.

1.1.3.1 Development of this Method

Since its introduction in 1996 this method has experienced tremendous improvements (Caffrey, 2003; Caffrey and Cherezov, 2009). It is now a high throughput method that uses very few materials and is fully automated (Caffrey and Cherezov, 2009; Cherezov, 2011a). The original protocol required that each individual crystallization experiment be set up in a glass tube. At first the dry lipid was placed into a small glass tube. On top the proper amount of protein solution was placed. The sample was centrifuged several times until a homogenous lipidic cubic phase had formed. This procedure needed to be repeated for every crystallization condition and was time consuming and difficult (Caffrey, 2003; Landau and Rosenbusch, 1996). The sample was incubated until crystals had reached the desirable size. This preparation

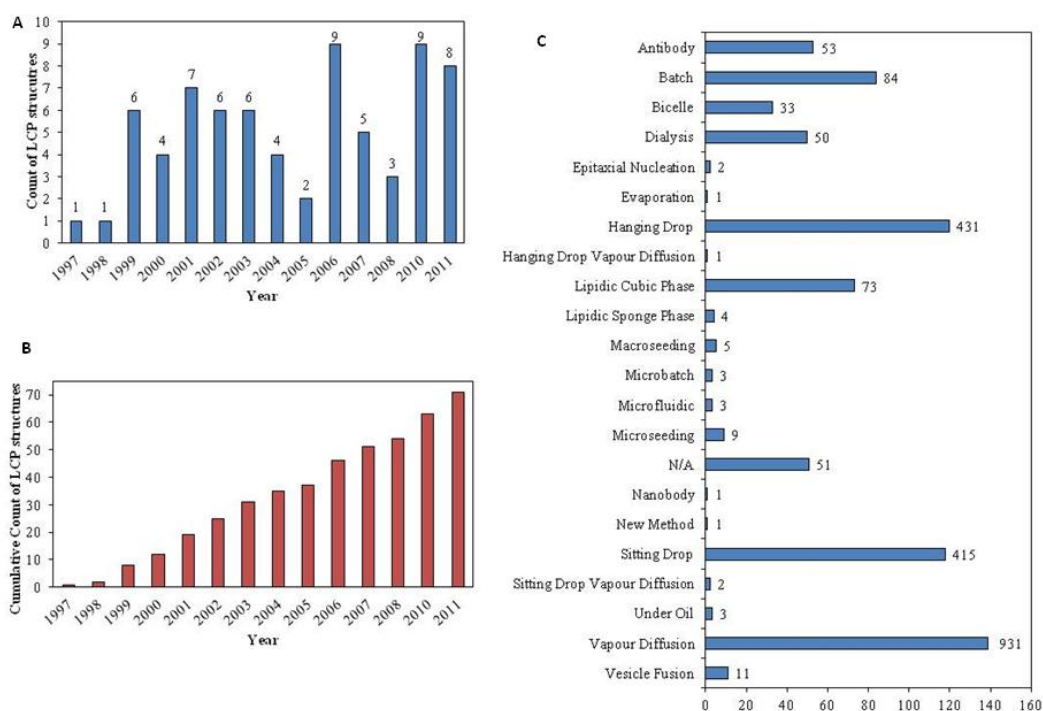


Figure 1.1.4: Number of published lipidic cubic phase (LCP) structures in MPDB; A: LCP structures by count per year; B: Cumulative count of LCP structures per year. C: Published structures of MPs in the MPDB by crystallization method (Raman *et al.*, 2006).

procedure had a number of disadvantages. Firstly, as already described each individual crystallization condition needed to be prepared by hand and individually centrifuged. The second major disadvantage was the large amount of lipid and membrane protein needed in order to prepare one crystallization experiment. Due to the setup in a tube and the large amount of lipidic cubic phase it was difficult to view crystals under the microscope. The difficulty translating into the harvesting the process. At that time, this method could only be used for colored proteins. With time there were improvements. The sample mixer first described by Cheng *et al.*, 1998 (Cheng *et al.*, 1998) made it possible to prepare cubic phase without having to use lengthy centrifugation steps. The process was further advanced by introducing a hand dispenser that could distribute small amounts of cubic phase on glass slides ranging

from 50 to 300 nl in volume (Caffrey, 2003). Individual crystallization spaces were created by spacers separating the crystallization drops and evaporation was avoided by covering the well with a glass slide. With these improvements the experimental time needed was reduced. Furthermore, by using smaller amounts of lipidic cubic phase, the amount of protein sample and lipid used was reduced significantly. An advantage of smaller amounts of lipidic cubic phase used made the viewing under the microscope easier and more precise. In addition the flat plates removed anomalous bending of light in the cubic phase sample.

With the advent of the *in meso* robot (Cherezov et al., 2004) and the imaging system (Caffrey and Cherezov, 2009) the *in meso* method became high throughput. Each individual crystallization experiment requires only 50 nl or less of lipidic cubic phase and 1 μ l of precipitant solution. The imaging system uses white light, polarized light and UV light to visualize crystals (Caffrey and Cherezov, 2009). The imaging system takes pictures of each well over a set imaging schedule that usually includes days 1, 3, 5, 8, 14, 21, 28 after crystallization setup (Caffrey and Cherezov, 2009). This automation and better understanding of the method turned this method from a highly specialized niche into a broadly available, high throughput method.

1.1.3.2 Advantages of the *in meso* Method

This method has a number of advantages over other techniques of crystallization. The first and major advantage is that the *in meso* method uses continuous bilayers compared to detergent micelle crystallization. The MP is reconstituted into the lipidic bilayer and can freely diffuse within the constraints of the bilayer. It is also assumed that the MP experiences a more biologically relevant environment compared to a detergent micelle. Additives like native lipids (such as cholesterol) (Cherezov et al., 2002), detergents (Misquitta and Caffrey, 2003), or spongefiers (Cherezov et al., 2006) can be added to the bilayer to keep the MP in the most stable conformation. Furthermore, effects from detergent micelles that may destabilize the MP are minimized due to diluting the detergent in the lipidic bilayer. This is important because detergents are used in the purification of MPs, meaning that MPs being put into a lipidic bilayer bring a certain amount of detergent with them. Additives like

spongefiers (Cherezov et al., 2006) can be added to the lipidic cubic phase drop that can modulate the size of the water channels and the lipidic phase to adjust to accommodate the MP better or improve crystallization. Due to the stiff nature of the lipidic cubic phase the cubic phase drop can be easily localized within the crystallization drop. When in the cubic or sponge phase, the lipidic drop appears clear under polarized light (in contrast to the lamellar phase, which is strongly birefringent). This makes the crystals much more visible, provided that they are birefringent. Harvesting crystals from lipidic cubic phase can pose a problem due to the stiff nature of the lipidic phase.

1.1.3.3 Proteins Crystallized by the *in meso* Method

The *in meso* method has produced a number of unique structures of MPs (table 1.1.2, figure 1.1.4) and is applicable to α -helical and β -barrel proteins. The most prominent examples for α -helical proteins are the GPCR structures determined over the past few years (Cherezov et al., 2007; Rasmussen et al., 2011). This method was also very productive for rhodopsins especially bacteriorhodopsin. The MPDB lists 33 entries for this protein (Raman et al., 2006). Both the GPCRs and rhodopsins span the membrane 7 times. α -helical structures like the light harvesting complex II span the membrane 26 times. Even more membrane spanning helices are found in the photosynthetic reaction center and in *baa₃* (table 1.1.2). So far the crystal structure of two β -barrel proteins (OpcA, BtuB) (table 1.1.2) have been published.

The Matthews coefficient (MC) and solvent content (V_m) of membrane proteins have been shown to be higher compared to soluble proteins (Schulz, 2011). The Matthews coefficient and solvent content describes the ratio of the volume of the asymmetric unit that is occupied by the protein versus solvent. Schulz, 2011 lists the average Matthews coefficient ($\text{\AA}^3/\text{Da}$) as 2.68, for monotopic MPs as 3.09 ± 0.08 , β -barrel proteins as 3.49 ± 0.09 , and α -helical proteins as 4.06 ± 0.08 . Matthews coefficient and solvent content observed for *in meso* structures (table 1.1.2) ranges from 2.3 to 3.6 and 43 to 66 % respectively. Table 1.1.2 shows that the peptides and rhodopsin structures have the lowest solvent content, followed by the two β -barrel structures. The photosynthetic proteins and GPCR structures have the highest solvent content.

The solvent content is lower for *in meso* crystallized proteins compared to the numbers published by Schulz, 2011. The difference is explained by the large volume the detergent micelle occupies.

Protein crystals that originated from lipidic cubic phase show the same tendency to crystallize in the most abundant space groups observed for proteins (C2, P2₁2₁2₁, P2₁, C222₁, P2₁2₁2, P3_{1,2}2₁ and P4_{1,3}2₁2) (Schulz, 2011), (table 1.1.2). The highest resolution for a MP grown *in meso* was obtained for bR to 1.48 Å. The small membrane peptide gramicidin diffracts to even higher resolution (1.12 Å). With the exception of GPCR structures most *in meso* crystal structure diffract to approximately 2 Å.

Table 1.1.2: Structures of membrane proteins produced by the *in meso* method (Information has been taken from MPDB, April 2012)

Protein name	MC ^a	MW ^b (kDa)	Space group	High. Reso. ^c (Å)	No. of PDB ID	MC ^d (Vs)	PDB ID
Rhodopsins							
Bacteriorhodopsin	7	27.1	P6 ₃	1.43	33	2.3 (47)	1AP9, 1BRX, 1C3W, 1C8R, 1C8S, 1CWQ, 1E0P, 1F4Z, 1F50, 1KG8, 1KG9, 1KGB, 1M0K, 1M0L, 1M0M, 1O0A, 1P8H, 1P8I, 1P8U, 1QHJ, 1QKO, 1QKP, 1VJM, 2I1X, 2I20, 2I21, 2NTU, 2NTW, 2WJK, 2WJL, 3MBV, 3NS0, 3NSB
			C222 ₁	2.3	5	2.4 (49)	1JV6, 1JV7, 1MGY, 1S8J, 1S8L
Halorhodopsin	7	28.9	P6 ₃ 22	1.8	1	2.3 (46)	1E12
			H3 ₂	1.93	2	2.3 (47)	2JAF, 2JAG
Sensory Rhodopsin	7	30.2	C222 ₁	2.0	1	3.13 (61)	1H68, 1JGJ, 1GU8, 1GUE, 3QAP, 3QDC
Sensory Rhodopsin II / Transducer Complex	7	43.7	P2 ₁ 2 ₁ 2	1.93	3	2.0 (39)	1H2S, 2F93, 2F95
Sensory Rhodopsin II	7	34.4	C222 ₁	2.1	4	2.8 (56)	1XIO
Heme-Copper Oxidases							
ba3-type cytochrome oxidase	15	84.0	C2	1.8	1	3.2 (62)	3S8F, 3S8G

Protein name (abbreviation)	MC ^a	MW ^b (kDa)	Space group	High. Reso. ^c (Å)	No. of PDB ID	MC ^d (Vs)	PDB ID
caa ₃ type cytochrome oxidase	23	140	I2	2.3	1	2.78 (56)	2YEV
G- protein coupled receptors (GPCR)							
β ₂ adrenergic receptor	7	55.9	P2 ₁ 2 ₁ 2 ₁	2.8	4	2.4 (49)	3D4S, 3NY8, 3NY9, 3NYA
		56.5	C2	2.4	1	3.1 (60)	2RH1
		52.4	P2 ₁ 2 ₁ 2	3.5	1	3.5 (65)	3PDS
β ₂ adrenergic receptor- Gs protein complex	7	44.3	P2 ₁	3.2	1	3.1 (60)	3SN6
A ₂ A adenosine receptor	7	54.7	P2 ₁	2.7	1	2.9 (57)	3EML, 3QAK
dopamine D ₃ receptor	7	53.6	P2 ₁ 2 ₁ 2 ₁	3.2	1	3.3 (64)	3PBL
Histamine H1 receptor	7	52.3	I422	3.1	1	3.1 (60)	3RZE
CXCR4 chemokine receptor	7	56.6	C2	2.9	1	3.6 (66)	3OE0
		56.6	P1	3.1	2	3.0 (58)	3OE8, 3OE9
		56.9	P2 ₁	2.5	1	2.8 (56)	3ODU
		57.5	I222	3.2	1	2.9 (58)	3OE6

Protein name (abbreviation)	MC ^a	MW ^b (kDa)	Space group	High. Reso. ^c (Å)	No. of PDB ID	MC ^d (Vs)	PDB ID
Photosynthetic Proteins							
Light Harvesting Complex II	18	129.6	P2 ₁ 2 ₁ 2 ₁	2.45	1	3.3 (66)	2FKW
Photosynthetic Reaction center	11	143	P4 ₂ 2 ₁ 2	2.0	1	2.8 (56)	1OGV, 2GNU, 2BNP, 2BNS
		143	P2 ₁ 2 ₁ 2	1.86	2	2.7 (54)	2WJM, 2WJN, 2X5U, 2X5V
Outer Membrane Proteins							
Outer membrane cobalamin transporter BtuB	22	66.3	P2 ₁ 2 ₁ 2 ₁	1.95	1	2.6 (53)	2GUF
Outer membrane adhesion protein OpcA	10	28.1	P2 ₁ 2 ₁ 2 ₁	1.95	1	2.2 (43)	2VDF
Peptides							
Gramicidin D	1	2.8	P2 ₁	1.08	5	2.6 (52)	2XDC, 2Y5M, 2Y6N
			P2 ₁	1.7	1	(40.6)	3ZQ8
			P2 ₁ 2 ₁ 2 ₁	1.2	2	2.3 (42)	Not in PDB

^a) membrane crossings; ^b) molecular weight, ^c) highest resolution, ^d) Matthews coefficient (solvent content) extracted from PDB file

1.1.3.4 Proposed Crystallization Mechanism

The proposed mechanism that describes the crystallization *in meso* is shown in figure 1.1.5. At first the MP that is dissolved in a detergent micelle is reconstituted into the lipidic cubic phase (figure 1.1.5.A). This is done by mixing dry or molten lipid with a solution of the MP. Once the LCP has formed, the MP is distributed evenly in the lipidic cubic phase and can diffuse freely (figure 1.1.5.A). The lipidic cubic phase system is in equilibrium. With the addition of precipitant solution the equilibrium is disturbed (figure 1.1.5.B). The precipitant solution interacts in many ways with the lipidic cubic phase. Some precipitants interact with the water channels, the MP, and the lipids in the curved bilayer. Each type of detergent has a different effect on the lipidic cubic phase system. It is believed that the MP diffuses from areas forming the LCP phase into areas containing the lamellar phase. The MPs start to aggregate in these lamellar regions and form a crystallization nucleus. If the precipitant solution has a crystal stabilizing effect on the nucleus and more protein molecules migrate into the nucleus, a protein crystal forms (figure 1.1.5.C).

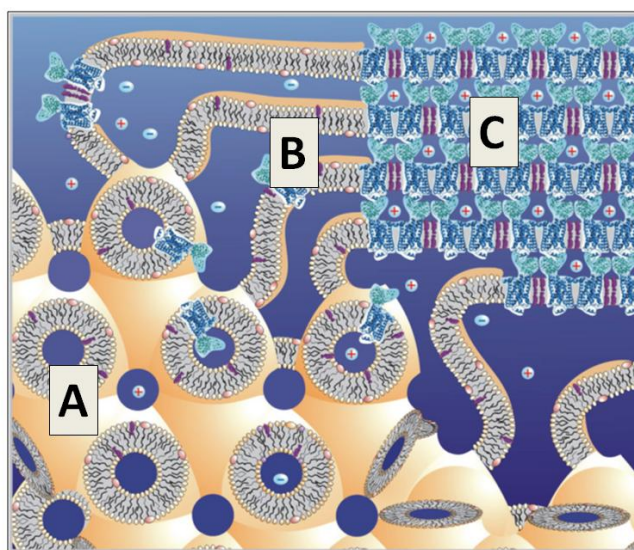


Figure 1.1.5: Schematic of proposed crystallization mechanism, see text for detailed description (Figure taken from Caffrey, 2009).

1.1.3.5. Type I Crystal Packing

Membrane protein crystal packing can be described by two different crystal types; type I and type II. In crystal type I the membrane protein is arranged in planar sheets, which are stacked on top of each other. Crystal contacts are formed between the crystals layers and also between MPs within one layer. Both the hydrophobic and the hydrophilic portion of the MP are involved in crystal contacts. In crystal type II crystal contacts can only form between the hydrophilic parts of the MP that is not covered by a detergent micelle or membrane particles. This type of packing tends to have a higher solvent content than type I.

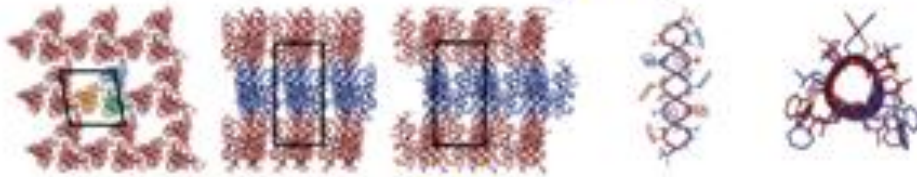
All *in meso* crystals so far have been found as type I crystals (figure 1.1.5). The crystal contact between proteins within one layer and can be extensive. An interesting aspect is that one unit vector is always perpendicular (within a reasonable tolerance) to the plane of the crystal layer.

Figure 1.1.6: Figure displayed on next page

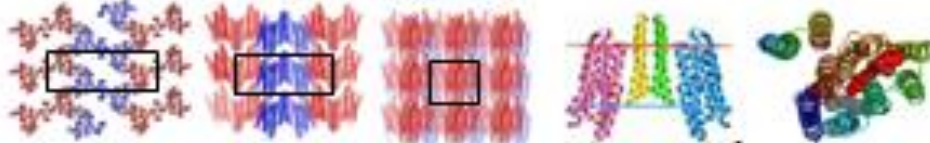
In meso crystal packing arrangements and molecular structures of membrane proteins. The packing arrangement is shown in columns 1–3 within each panel. Expanded views of individual proteins or oligomers are shown in columns 4 and 5. The views in columns 1 and 5 represent a projection along the stacking axis. The views in columns 2–4 are from within the plane of the stacked lamellae along the two other unit cell axes. In columns 1–3 black outlines are projections of the unit cell. Images in column 4 are from the Orientations of Proteins in Membranes (OPM) database (<http://opm.phar.umich.edu/>), in which the red and blue horizontal lines define the hydrophobic thickness of the protein. In column 4 (*d, e*) the horizontal red and blue lines correspond to the extracellular and periplasmic surfaces, respectively. In column 4 (*a–c, f–i*) the red and blue lines correspond to the extracellular and intracellular surfaces, respectively. Proteins are identified by name, source organism, resolution, and Protein Data Bank (PDB; <http://www.rcsb.org/pdb/home/home.do>) accession number. This figure and figure legend have been adapted from Caffrey, 2009.



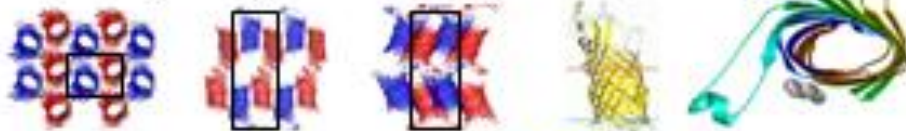
(a) Bacteriorhodopsin, *H. salinarum*, 1.43 Å, [1MOK](#)



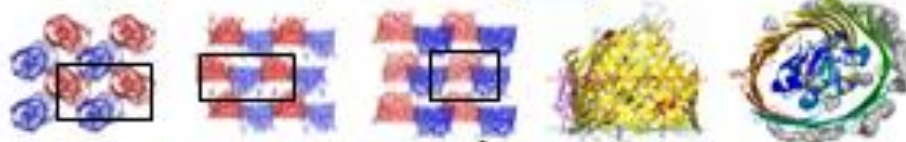
(b) Gramicidin D, *B. brevis*, 1.70 Å, [2XDC](#)



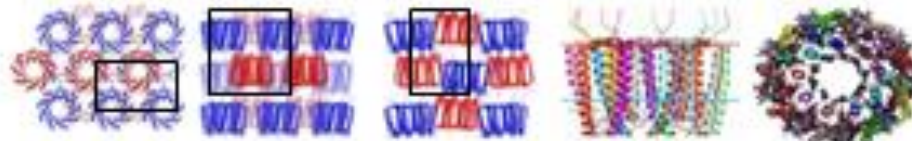
(c) Sensory rhodopsin II / transducer, *N. pharaonis*, 1.93 Å, [1H2S](#)



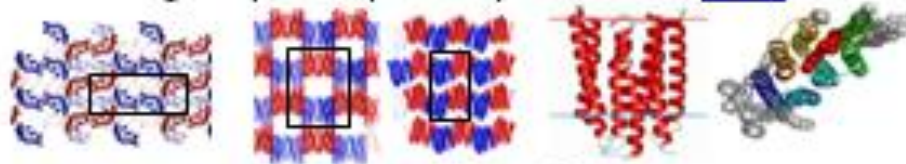
(d) Adhesin OpcA, *N. meningitidis*, 1.95 Å, [2VDF](#)



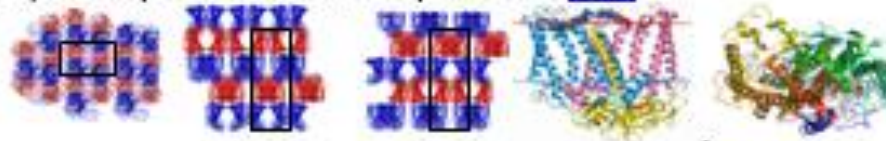
(e) Cobalamin transporter, *E. coli*, 1.95 Å, [2GUF](#)



(f) Light harvesting complex II, *Rps. acidophila*, 1.95 Å, [2FKW](#)



(g) Sensory rhodopsin II, *Anabaena* sp., 2.00 Å, [1XIO](#)



(h) Photosynthetic reaction center, *R. sphaeroides*, 2.20 Å, [2GNU](#)

1.1.4 Crystallization of Small Proteins and Peptides using Lipidic Cubic Phases

The *in meso* method has produced a variety of structures over the years ranging in molecular weight from 143 kDa with the photosynthetic reaction center to 25 kDa with numerous rhodopsin structures (figure 1.1.7). The weight range not explored so far is for membrane proteins and membrane peptides that are smaller than 25 kDa and have less than 6 transmembrane spanning helices. For this reason the small membrane peptide gramicidin (section 1.2) was chosen to explore the potentially smallest molecular weight range for membrane protein / peptides.

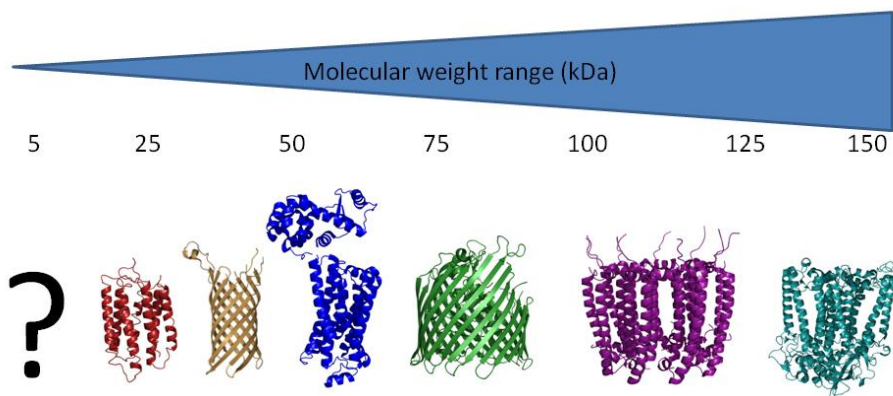


Figure 1.1.7: Molecular weight range of different MPs that has been explored for *in meso* crystallization. Beginning from the right: photosynthetic reaction center (143 kDa, PDB ID: 1OGY), light harvesting complex (129.6 kDa, PDB ID: 2FKW), outer membrane cobalamin transporter BtuB (66.3 kDa, PDB ID: 2GUF), β_2 adrenergic receptor β_{AR} (56.5 kDa, PDB ID: 2RH1), outer membrane adhesion protein OpcA (28.1 kDa, PDB ID: 2VDF), bacteriorhodopsin (27.1 kDa, PDB ID: 1AP9)

1.1.4.1 Gramicidin D as Test Peptide

To test the applicability of crystallizing membrane peptides from lipidic cubic phases the small membrane peptide gramicidin D was chosen. Each monomer has 15 amino acids that alternate in an L and D fashion. This allows the peptide to adopt a variety of conformations, the main conformations are the head to head single stranded helix (HSSH) and the double stranded double helix (DSDH) conformations. Furthermore, the peptide is protected from proteolytic cleavage through capping with ethanolamine on the C-terminus and through formylation in the N-terminus. The peptide is well studied by NMR, MX, circular dichroism (CD), etc.. The peptide functions as an ion channel and has been identified as a second messenger in *Bacillus brevis*, the bacterium that produces it in a non-ribosomal fashion. Gramicidin D is described in detail in section 1.2. There has been discussion in the literature of whether or not small membrane peptides can be crystallized using the *in meso* method (Grabe et al., 2003). This chapter shows that this is possible and that physiologically relevant structures can be obtained. Gramicidin was chosen because it is a small membrane peptide that is readily available without the need for further purification.

The *in meso* method was chosen as the crystallization method since it provides the peptide with a more natural environment. The bilayer can be modulated by choosing different monoacylglycerols. In the past, crystal structures from gramicidin all used organic solvents. This brought into question the physiological relevance of the structures. Organic solutions and bilayers of different thickness effect the conformation of gramicidin significantly.

In the work presented here, three different crystal forms (CF1, CF2, CF3) of gramicidin are described. Gramicidin crystallized as CF1 under similar conditions, in the same crystal form from a variety of lipids of different bilayer thickness (9.9, 8.8, 7.7 MAG). The conformation observed is again the DSDH conformation. Crystallization appeared to be independent of bilayer thickness. A second crystal form (CF2) is described. These crystals were grown from the lipidic cubic phase monovaccenin (11.7 MAG). The conformation observed is identical to the one in crystal form 1 (CF1). A third crystal form (CF3) was observed and is also described.

The conformation observed is the DSDH conformation. Of interest are the hydrogen bonds that are observed mainly between the inner ring tryptophan residues. Their interaction may possibly be an explanation of how aggregates of gramicidin stabilize in the membrane.

1.2 – The Small Membrane Peptide Gramicidin

1.2.1 Introduction

The membrane spanning pentadecapeptide gramicidin is expressed non-ribosomally in *Bacillus brevis* (Sarkar N, 1972). Gramicidin (gD) is a peptide antibiotic and has a molecular weight of 1886 Da. It adopts a variety of different conformations. The two main families of conformations (figure 1.2.1) are a head to head single stranded helix (HSH) (Ketchum et al., 1993) and a double stranded double helix (DSDH) (Burkhart et al., 1998b) (Lin, 2005). The HSH conformation has been identified as an ion channel (Ketchum et al., 1993). Over the years gramicidin has established itself as a model ion channel for the transports of monovalent cations (Jones et al., 2010). It was the first ion channel that was structurally well-studied and it yielded a significant amount of information which could be transferred to more complex and more difficult to study ion channels such as KcsA (Roux et al., 2004). The DSDH conformation can be left- or right-handed, parallel or antiparallel (Wallace, 1983) (figure 1.2.1). Some could potentially transport monovalent ions (Burkhart et al., 1998b; Doyle and Wallace, 1997). There has been controversy about the biological function of this conformation (Burkhart and Duax, 1999; Cross et al., 1999). The DSDH conformation has been observed in membranes, but is considered to be a non-ion-conducting conformation. The gD peptide has also been identified as a second messenger in *Bacillus brevis* in the transfer from vegetative growth to sporulation (Mandl and Paulus, 1985; Modest et al., 1984). This chapter discusses different aspects of gramicidin that allow it to adopt this wealth of conformations and let it function as an ion channel and as a second messenger. In the first part of this chapter, the characteristics of the sequence are introduced. This is followed by the expression of gramicidin in *Bacillus brevis*. The different conformations that it adopts are summarized. The mechanism of ion transport in HSH and DSDH is explained. The special importance of tryptophan residues to gD conformation is explained as is the

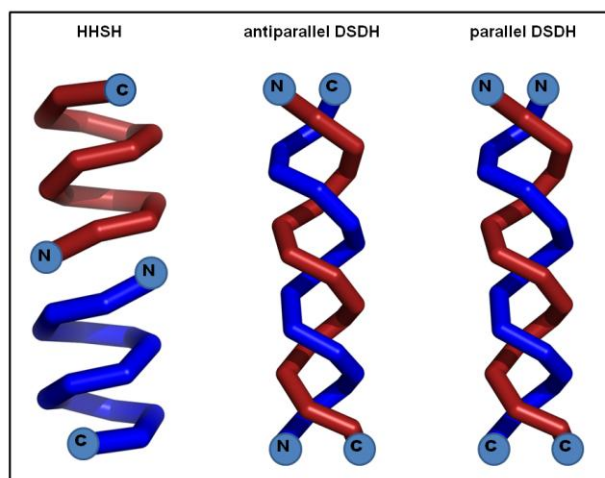


Figure 1.2.1: Schematic representation of major gramicidin conformations; HSSH (head to head single stranded helical and the parallel and antiparallel DSDH (double stranded double helical) conformation (Burkhart et al., 1998a; Burkhart et al., 1998b; Chen et al., 1996b; Ketchem et al., 1993).

influence that the medium in which gramicidin is dissolved. Spectroscopic results and the behavior of gramicidin in the lipidic cubic phase are also introduced.

1.2.2 Sequence of Gramicidin D

Gramicidin D (gD) is a small membrane-inserting peptide consisting of 15 amino acids (Hotchkiss and Dubos, 1940; Sarges and Witkop, 1965). The sequence alters between L and D amino acids and has the following composition:

N- terminus --- For₀-L-Val₁-Gly₂-L-Ala₃-D-Leu₄-L-Ala₅-D-Val₆-L-Val₇-D-Val₈-L-Trp₉-D-Leu₁₀-L-Trp₁₁-D-Leu₁₂-L-Trp₁₃-D-Leu₁₄-L-Trp₁₅-Eta₁₆ --- C- terminus.

The N- terminus is formylated (For) and the C- terminus is capped by ethanolamine (Eta). The N- and C- termini-protecting groups shield the peptide from proteolytic cleavage and keep the peptide uncharged at physiological pH (Kessler et al., 2004;

Schoenafinger et al., 2006; Schracke et al., 2005). The peptide would otherwise carry a charge at physiological pH since the N- terminus (α -amino group) has a pK_a value between 6.8 and 8.0 and the C- terminus (α -carbonyl group) has a pK_a value between 3.5 and 4.3. The lack of charge has two principle effects. Firstly it facilitates the insertion of the N- terminus into the membrane and secondly N- termini cannot repel one another. This is important because the N- termini face each other in the HSSH conformation in the membrane.

Examining the amino acid composition of gramicidin reveals a prevalence of hydrophobic amino acids (leucine (Leu), valine (Val), tryptophan (Trp)). This makes the peptide highly hydrophobic and almost insoluble in water, causing it to form colloidal solutions (MerckIndex, 1996). The modification at the N- and C- termini (formylation, capping with ethanolamine) decreases the solubility in water even further.

Gramicidin exists as a mixture of different isomers of gramicidin A (gA, 80 %), B (gB, 6 %), C (gC, 14 %) and is usually named Gramicidin D (Sarges and Witkop, 1965). gA has a tryptophan at position 11, gB a phenylalanine, and gC a tyrosine residue. There is also some variability in residue 1. It is valine in 95 % and isoleucine in the remaining 5 %. The variability is caused during the non-ribosomal synthesis in *Bacillus brevis* (Kessler et al., 2004).

Furthermore, D and L amino acids alternate in this peptide sequence with the result that all side chains are found on one side of the β -sheet (Urry, 1971) (see sequence above). All tryptophan and alanine residues are L amino acids. All leucine residues are D and most valine residues are D also. The alternating nature of the amino acid chirality has one exception: the second amino acid is glycine, which has no chirality. Gramicidin has a large number of tryptophan residues (8 tryptophan residues per dimer), which influence the conformation and character of the molecule significantly (Chattopadhyay et al., 2008; Sun et al., 2008). The functional group in tryptophan is the indole ring. Tryptophan has a dipole moment of 2 D along the C_5-N_1 axis (Gu et al., 2011). The N_e nitrogen can participate in hydrogen bonding (Gu et al., 2011). It is this combination of hydrophobic and hydrophilic interactions that makes tryptophan

prone to participating in the interface of the biological membrane (Liu and Caffrey, 2006). In membrane proteins tryptophan is therefore often found on surfaces of the membrane protein that border the interface. It is considered to anchor the membrane protein in the bilayer (Killian and von Heijne, 2000; Yau et al., 1998). Tryptophans can also stabilize protein structures through π - π or cation- π interactions (Burley and Petsko, 1986; Chakrabarti and Bhattacharyya, 2007).

1.2.3 Expression of Gramicidin D in *Bacillus brevis*

Gramicidin is naturally expressed by *Bacillus brevis* during early stages of sporulation (Sarkar N, 1972). Analysis of the cell content at the time of this event has also shown a high abundance of tyrocidin, a circular decapeptide. Both are considered to promote the shift from vegetative growth to sporulation (Mandl and Paulus, 1985; Marahiel et al., 1993; Modest et al., 1984). Research has shown that *Bacillus brevis* mutants lacking the ability to synthesize gramicidin, cannot switch into sporulation. Furthermore, research has shown that gramicidin interacts with RNA polymerase (Fisher and Blumenthal, 1982; Paulus et al., 1979; Ristow and Paulus, 1982; Sarkar et al., 1979). The exact nature of this interaction is not known.

Due to the unusual composition of the amino acids (presence of D amino acids), gramicidin cannot be synthesized by the ribosome. Some bacteria have developed a system of enzymes that can synthesize peptides on non-ribosomal peptide synthases (NRPSs) (Schracke et al., 2005). NRPSs are multifunctional enzymes that carry out several reactions in a specific and coordinated manner. For gramicidin these NRPSs are four multienzymes (LgrA, LgrB, LgrC, and LgrD) that together consist of 16 modules with 56 domains (Kessler et al., 2004). Crystal structures have been determined for proteins that have a similar activity, but not yet for the NRPSs of gramicidin (Strieker et al., 2010). Each module can perform activation, covalently bind an amino acid, and perform a condensation reaction through peptide bond formation. This approach is widely used in bacteria and allows for modifications that increase the biological stability through implementing D amino acids, modify the N-

and C- termini, or insert methylation or glycosylation to the peptide (Strieker et al., 2010).

1.2.4 Functions of Gramicidin D

Gramicidin is an antibiotic which is biologically active against gram positive bacteria with the exclusion of bacilli (Drug Bank, DB00027; www.drugbank.ca). Activity has also been found against some gram negative bacteria. It was observed that gramicidin disturbs asymmetric ion gradients across biological membranes, depleting them of their ability to produce ATP and in turn causing cell death. This function was later identified as an ion channel function and was studied in great detail. Gramicidin is able to transport monovalent ions across the biological membrane. It is insensitive to bivalent cations and also anions (Ketchum et al., 1993; Townsley et al., 2001). The function of gramicidin as second messenger has already been described and is not the main focus here. The pharmaceutical industry uses tyrothricin (the ethanolic extract consisting of gramicidin and tyrocidin) as an antibacterial substance in sore throat medication, skin creams, eye medication, etc. (Drug Bank, DB00027; www.drugbank.ca). Internal administration has been associated with hemolysis and is therefore discouraged (Drug Bank, DB00027; www.drugbank.ca). In the research community gramicidin's ability to form membrane crossing ion channels and its easy production make it a model peptide. By increasing the understanding of simple ion channels (Jones et al., 2010), and the behavior between membrane lipids and membrane peptide/protein (Lundbaek et al., 2010), model peptides help to introduce new experimental techniques (Misra et al., 2009) and promote better understanding of the influence of tryptophan residues in membrane proteins (Gu et al., 2011).

1.2.5 Conformations Adopted by Gramicidin D

Gramicidin is a highly flexible peptide that can adopt a large number of conformations. This is possible due to its relatively short length, the positioning of all side chains on one side of the peptide (caused by alternation of D and L), and the ability to form dimers (Kessler et al., 2004). Two main families of conformations are known; the **H**ead to **H**ead **S**ingle stranded **H**elical (HSHS) conformation (Urry, 1971) and the **D**ouble **S**tranded **D**ouble **H**elical conformation (DSDH) (Burkhart et al., 1998a; Burkhart et al., 1998b). Both conformations are dimers and show β -sheet-like hydrogen bonding.

The head to head conformation (table 1.2.1) forms a π -helix with 6.3 amino acids per turn (Urry et al., 1971). It shows parallel β -sheet-like hydrogen bonding with 12 intramolecular hydrogen bonds (figure 1.2.2) with the exception of the head to head junction. In this region the hydrogen bonding pattern changes to antiparallel. Both monomers are connected via 6 hydrogen bonds at their N- termini in a head to head fashion. The helices are both right handed (figure 1.2.3(A)). The inner diameter of this conformation is wide enough to transport ions across the membrane (Ketchum et al., 1993). Tryptophan residues are localized in the interface of the biological membrane. The structure has been determined by solid state and solution state NMR in a phospholipid membrane (Ketchum et al., 1993) and sodium dodecylsulfate (SDS) micelles. This conformation is considered to be the active ion channel conformation. It should be mentioned that a crystal structure has not been determined to date.

The DSDH conformations are intertwined and have been observed in parallel and antiparallel fashion, as right or left handed conformation, and in an open ion-bound - or closed conformation (Chen, 1997; Wallace, 1983) (figure 1.2.3(B)). Structure determination was performed by either NMR or X-ray crystallography (table 1.2.1). The intertwined helices can vary in their stagger, which results in a shorter and wider or longer and thinner conformation. PDB entries of DSDH include one parallel, left-handed, ion-bound conformation, several left-handed antiparallel, non-ion-bound

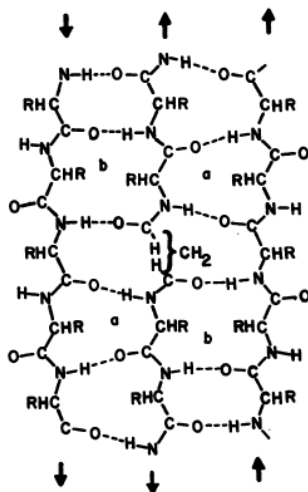


Figure 1.2.2: The HSSH (π -type helix) conformation is shown in an unrolled fashion showing the head to head junction. The HSSH conformation shows hydrogen bonding of an antiparallel β -pleated sheet (regions b). At the head to head junction the hydrogen bonding pattern changes to a β -pleated sheets. The junction site of two formyl moieties at the center of the diagram facilitates the transition between the two hydrogen-bonding patterns. This figure has been taken from Urry et al., 1971 (Urry et al., 1971).

conformations, and several right- and left-handed, ion-bound antiparallel conformations (see table 1.2.1 for PDB ID and reference). The left-handed antiparallel, closed DSDH conformation (figure 1.2.3(C)) forms a β -helix with a pitch of 5.6 amino acids per turn. The two strands are held together by hydrogen bonding between the oxygen and nitrogen atoms of the peptide backbone. This conformation was crystallized from a variety of organic solvents (methanol, ethanol, propanol). The structure was first described by Langs et al., 1991 (no PDB ID available) and later confirmed by (Burkhart et al., 1998a) (PDB ID 1AL4, 1ALX, 1ALZ). The dimer is so tightly wound that its core is considered to be too tight to transport ions. The molecule has a length of approximately 35 Å. The tryptophan residues are located along the entire dimer. It is this conformation that will be the main focus of the remaining chapter.

The left or right handed antiparallel ion-bound conformation has 6.4 amino acids per turn (Burkhart et al., 1998b; Doyle and Wallace, 1997; Wallace and Ravikumar, 1988). The different stagger makes the dimer shorter and wider (length 26 Å). This has the effect that the inner width is great enough to support ion transport. It should be emphasized that this is not the predominant conformation for ion transport. The structure is associated with a pore function. These structures have been crystallized from organic solvents in the presence of different ions. Two of the structures (PDB IDs: 1C4D, 1GMK) are associated with the left-handed antiparallel DSDH. In the literature, particularly Duax and coworkers, have repeatedly defended that the left-handed crystal structures as described in 1GMK and 1C4D are invalid (Burkhart et al., 1998b). Several PDB IDs 3L8L, 2IZQ, 1AV2, 1BWD, 1W5U describe the right-handed antiparallel DSDH conformation.

The parallel left-handed DSDH conformation (1MIC) (figure 1.2.3(D)) was described by Chen et al., 1996 (Chen et al., 1996a). Due to the parallel orientation of the two intertwined chains, all tryptophan residues are located on one side of the dimer. The biological function of the conformation is unknown.

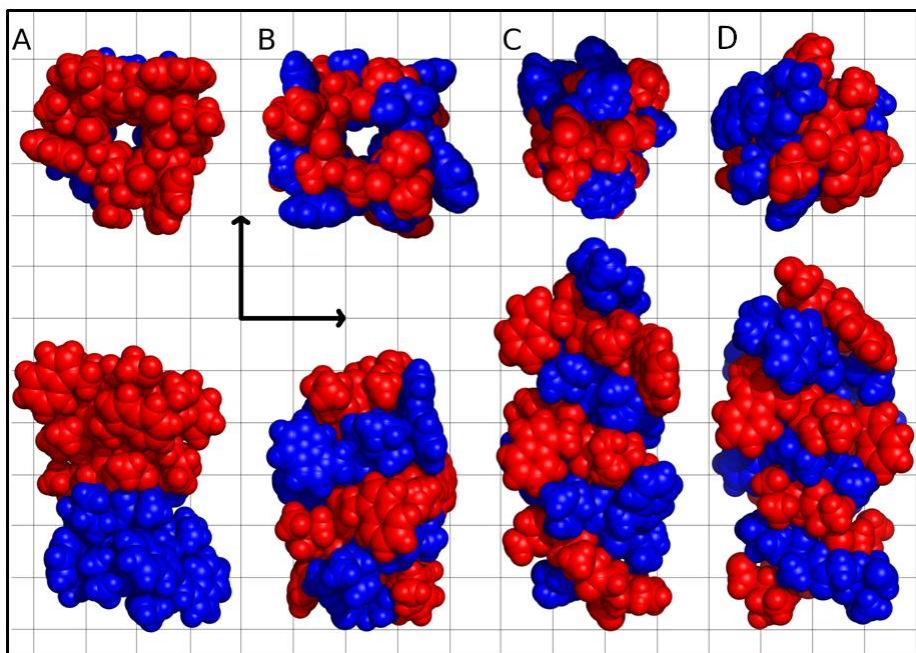


Figure 1.2.3: Conformations adopted by linear gramicidin. End on view along the normal to the membrane plane (upper panels). Side view from within the membrane (lower panels). The conformations shown include the HSSH (PDB ID, 1JNO; A), right-handed DSDH – ion bound form (PDB ID, 3L8L; B, ions not shown), left-handed DSDH – non-ion bound (closed) form (PDB ID, 1ALZ; C) and parallel, left-handed DSDH form (PDB ID, 1MIC; D). Individual monomers are colored red and blue for clarity. The background grid has vertical and horizontal lines every 5 Å for scale. This figure has been taken from Höfer *et al.*, 2011.

Table 1.2.1: Gramicidin structures available in the Protein Data Bank; this table has been taken from Höfer *et al.*, 2011.

Method	Resolution (Å)	Dispersing Medium	PDB ID ^a	Reference
DSDH (AP ^b , RH ^c , ion-bound)				
MX	1.25	Methanol (+ NaI)	3L8L	Olczak et al., 2010
MX	0.80	Methanol (+ KI)	2IZQ	Olczak et al., 2007
MX	1.40	Methanol (+ CsCl)	1AV2	Burkhart et al., 1998b
MX	1.70	Glacial acetic acid	1BDW	Burkhart et al., 1998b
MX	1.14	Ethanol (+ RbCl)	1W5U	Glowka et al., 2005
DSDH (AP ^b , LH ^d , ion-bound)				
MX	2.50	Methanol (+ KSCN)	1GMK	Doyle & Wallace, 1997
MX	2.00	Methanol (+ CsCl)	1C4D	Wallace & Ravikur, 1988
DSDH (AP ^b , LH ^d , ion-free)				
MX	1.13	n-Propanol	1AL4	Burkhart et al., 1998a
MX	0.86	Ethanol	1ALZ	Burkhart et al., 1998a
MX	1.20	Methanol	1ALX	Burkhart et al., 1998a
MX	1.70	LCP ^e , 9.9 MAG	2XDC	Höfer et al., 2010, Höfer et al., 2011, Chapter 1
MX	1.28	LCP ^e , 8.8 MAG	2Y6M	Höfer et al., 2011, Chapter 1
MX	1.08	LCP ^e , 7.7 MAG	2Y5N	Höfer et al., 2011, Chapter 1
MX	1.8	LCP ^e , 11.7 MAG	-	Chapter 1
MX	1.7	LCP ^e , 11.7 MAG	3ZQ8	Chapter 1
DSDH (P ^f , LH ^d)				
NMR	n/a	Methanol (+ CaCl ₂)	1MIC	Chen et al., 1996b

Method	Resolution (Å)	Dispersing Medium	PDB ID ^a	Reference
HSHH (AP ^b , RH ^c)				
NMR	n/a	SDS micelle	1GRM	Lomize et al., 1992
NMR	n/a	SDS micelle	1JNO	Townsley et al., 2001
NMR	n/a	SDS micelle	1JO3	Townsley et al., 2001
NMR	n/a	SDS micelle	1JO4	Townsley et al., 2001
NMR	n/a	DMPC bilayer	1MAG	Ketchum et al., 1996
NMR	n/a	DDPC micelle	1NRU	To be published
NMR	n/a	DDPC micelle	1NRM	To be published

^a PDB ID – protein data bank identification code. ^b AP – anti-parallel. ^c RH - right-handed. ^d LH - left-handed. ^e LCP – lipidic cubic phase. ^f P – parallel.

1.2.6 Proposed Mechanism of HSHH (Ion Channel) Conformation

The HSHH conformation has been identified as the ‘ion channel conformation’ (Ketchum et al., 1993; Lundbaek et al., 2010). The eight tryptophan residues are located at the interface of the membrane, anchoring the structure in the membrane (Gu et al., 2011). Looking at the structure from the top there is an opening that is capable of transporting monovalent cations. This conformation has typical features that are found for ion channels like gating and removal of hydration shells from ions (Li et al., 2010).

The channel is gated, meaning that the channel has an open or closed conformation (Jones et al., 2010). The mechanism by which gramicidin does this is still debated. There are two proposals. One proposal suggests that the opening and closing of the channel happens by the channel going from a dimeric conformation towards a monomeric, less folded conformation (Mo et al., 2004). This requires the breakages of six hydrogen bonds that hold the dimer together in the mid-portion of the membrane raising the question if this is energetically possible. A more recent

proposal (Jones et al., 2010) suggests that the channel remains dimeric and that small conformational changes in the area of the N- terminus (For₀ and Val₁) may trigger the open or closed conformation. It is proposed that the carbonyl orientations of For₀ and Val₁ can alternate between an inward position pointing towards the inside of the channel and an outwards position. In the HSH conformation these residues are located in the middle of the dimer. Pointing outward would inhibit conduction and stop the flow of ions and result in a closed ion channel.

The channel has two cation binding sites (Li et al., 2010), which are placed near the entrance of the channel. These binding sites are the Leu₁₀ - Trp₁₁ peptide linkage. The ion is stabilized here through electrostatic interaction. The large energetic loss of ion dehydration is compensated by coordination with main-chain carbonyl oxygens, which is not involved in hydrogen bonding with the carbonyl backbone. This feature of using the carbonyl backbone of a peptide is also observed in larger ion channels like KcsA (Doyle et al., 1998). The pore of gramicidin D is filled with water molecules that are replaced as ions are passed through the channel.

1.2.7 Proposed Mechanism of the DSDH Conformation

This right-handed, antiparallel DSDH has been introduced as having a pore diameter wide enough to transport ions. Several crystal structures describe this conformation (Burkhart et al., 1998b; Glowka et al., 2005; Olczak et al., 2007; Olczak et al., 2010). All crystals are grown from organic solvents. Literature does not consider this conformation the major conformation responsible for ion transport and so not nearly as much research effort has been put into this conformation. Furthermore, there has been controversy in the literature about the biological function of this conformation (Burkhart and Duax, 1999; Cross et al., 1999).

1.2.8 Influence of Tryptophan Residues on Conformation and Ion Channel Activity in Gramicidin D

As mentioned above, gramicidin has a total of eight tryptophan residues per dimer (four per monomer). In the sequence they are located towards the C-terminus in position 9, 11, 13, and 15. Tryptophan residue 11 can also be substituted by phenylalanine or tyrosine (gA – Trp₁₁, gB – Tyr₁₁, gC – Phe₁₁). Tryptophan residues 13 and 15 (outer ring tryptophan residues) are always found in the interface of the membrane independent of the conformation that gramicidin adopts (Chattopadhyay et al., 2008). The second set of tryptophan residues 9 and 11 (inner ring tryptophan residues) can either be found in the interface of the membrane for the HSHH conformation or inside of the membrane for the DSDH conformations (figure 1.2.4) (Chattopadhyay et al., 2008). Black lipid membrane (BLM), circular dichroism (CD), size exclusion chromatography (SEC) and nuclear magnetic resonance (NMR) experiments on the natural gramicidin derivatives (gA, gB, gC) show a different tendency to conduct ions (Townsend et al., 2001). These derivatives also tend to have different ratios between the conformations (HSHH and DSDH). With an increase of hydrophobic character as seen in phenylalanine the equilibrium is shifted towards the DSDH conformation. An effect is also observed for the conductance; it reduces by a factor of 2 when Trp₁₁ is replaced by Phe₁₁ (Chattopadhyay et al., 2008).

In order to better understand the importance of these tryptophan residues, single or multiple substitutions have been performed on all tryptophan residues. The tryptophan residue can be substituted with phenylalanine. If all tryptophan residues in gramicidin are replaced by phenylalanine (gramicidin M), gramicidin predominantly adopts the DSDH conformation and has a six-fold reduced conductance (Markham et al., 2001). By substituting with phenylalanine the influence on the outer ring tryptophan residues is less severe than in the inner ring tryptophan residues (Sun et al., 2008). This indicates that tryptophan 9 and 11 have a significant effect on conductance and conformation. Further experiments were carried out in which either the dipole moment in tryptophan or tryptophan's ability to form hydrogen bonds was

altered. These experiments show that conductance is influenced by the dipole moment (Chattopadhyay et al., 2008) of tryptophan and the ability of gramicidin to adopt the HSHH conformation is influenced by forming hydrogen bonds (Chattopadhyay et al., 2008).

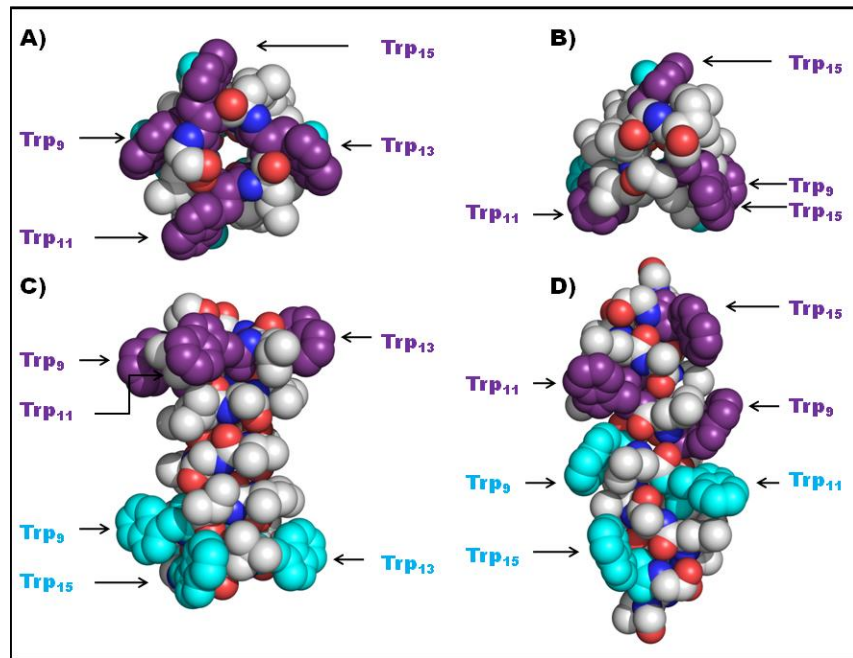


Figure 1.2.4: Space-filling model in top and side view of the HSHH conformation in A, C based on PDB ID 1GRM and the DSDH conformation in B, D based on PDB ID 2XDC. Tryptophan residues are indicated by name and colored in purple and cyan. Non-tryptophan carbon, nitrogen and oxygen atoms are shown in grey, blue, and red, respectively (figure has been adapted from Liu and Caffrey, 2005).

1.2.9 Sensitivity of Gramicidin D towards its Surrounding Environment

The conformation of gramicidin is highly dependent on the environment in which it is dissolved. Gramicidin dissolves without problems in organic solvents like ethanol, methanol, propanol, benzene, trifluoroethanol (TFE), chloroform, octanol to concentrations 20 mg/ml and higher (MerckIndex, 1996). In these solvents gramicidin only adopts a number of different DSDH conformations (Chen, 1997). These conformations are in equilibrium. To date no solvent has been found in which the HSHH conformation can be stabilized. The preferred conformation in membranes is HSHH (Killian et al., 1988). In situations of hydrophobic mismatch between gramicidin and the bilayer, the peptide is in equilibrium between HSHH and DSDH. Increasing membrane thickness causes an increase in DSDH conformation. Kelkar and Chattopadhyay, 2007 showed that in a phospholipid membrane of 1,2-diarachidoyl-sn-glycero-3-phosphocholine (DAPC; membrane thickness 50 Å) the DSDH conformation is the predominant conformation. In very thin membranes with less than 10 carbon atoms per acyl chain an increase in DSDH conformation is also observed (Greathouse et al., 1994).

Gramicidin D adopts the HSHH conformation in lipid bilayers of different composition (de Godoy and Cukierman, 2011). The character of the lipid (phospholipid, monoacylglycerol, etc.) does not appear to influence the conformation. The conformation is only dependent on bilayer thickness. The choice of lipid has more of an effect on channel lifetime and ion conductance. As an example, the initial ion conductance work on gramicidin in black lipid membranes was performed in different MAGs including monoolein (Pope et al., 1982). Gramicidin adopted the HSHH conformation. Although not a lipid, the first structure of gramicidin in HSHH conformation was determined by NMR in micelles of sodium dodecyl sulfate (SDS) (Lomize et al., 1992).

Noteworthy is also the tendency of gramicidin to be affected by its solvent history (Bouchard and Auger, 1993; LoGrasso et al., 1988). The general procedure for

preparing gramicidin in membranes is to co-dissolve the peptide with the lipid in an organic solvent. As mentioned above gramicidin adopts the DSDH conformation in organic solvents (Chen, 1997). This conformation is one of many possible low energy conformations in which gramicidin can be observed. Experiments have shown that this conformation can be carried over into the bilayer (LoGrasso et al., 1988). The expected equilibrium can be reinstated by heating the sample for 48 hours above 50 °C (LoGrasso et al., 1988). This lock in conformation is particularly pronounced for chloroform and alcohols. Trifluoroethanol is a good choice as a co-solubilization solvent, because gramicidin exists in a monomeric less folded conformation not locking it *a priori* in a particular conformation (LoGrasso et al., 1988).

Gramicidin tends to form aggregates in membranes. Both HSHH and DSDH conformation have been observed to show this behavior (Diociaiuti et al., 2002; Kelkar and Chattopadhyay, 2007). For the HSHH conformation it has been seen by atomic force microscopy that the aggregate consists of hexameric units (Diociaiuti et al., 2002).

1.2.10 Gramicidin and the Lipidic Cubic Phase

Gramicidin has been used as a model peptide for the introduction of many new techniques and methodologies (Misra et al., 2009). The characterization of gramicidin incorporated into the lipidic cubic phase was first described in Liu and Caffrey, 2005. In this paper several protocols were reported for the incorporation of gramicidin into lipidic cubic phases of different lipids. Liu and Caffrey, 2005 describe a direct reconstitution of the peptide into molten monoolein and two indirect ways of reconstitution. In the first co-solubilization protocol, gramicidin and lipid are dissolved in methanol as an organic solvent and dried under vacuum to remove the solvent. The second protocol uses SDS micelles that are mixed with a solution of gramicidin pre-dissolved in TFE. The peptide incorporates into the SDS micelle and is then mixed with cubic phase hosting lipid to form the lipidic cubic phase. In the following chapters the co-solubilization protocol is used in which gramicidin, lipid,

and methanol are combined and then dried under vacuum. The gramicidin lipid mixture was then mixed with buffer or water to form the lipidic cubic phase. Crystallization experiments were also performed using the second co-solubilization protocol with SDS. Unfortunately, crystals were never observed with this technique. Liu and Caffrey show that the lipidic cubic phase of 9.9 MAG and 9.7 MAG can incorporate large quantities of gramicidin up to 10 mole%. At amounts greater than 10 mole% the lipidic phase changes from cubic to a hexagonal phase (H_{II}) in the case of 9.9 MAG. The phase identity remains stable for 9.7 MAG in the measured range of up to 10 mole%. The lattice parameter of monoolein free of gramicidin at 20 °C and a water lipid ratio of 2/3 is ~ 100 Å and decreases to approximately 70 Å at 10 mole% gramicidin. The lattice parameter for 9.7 MAG is approximately 10 Å larger (figure 1.2.5). For the work presented in this thesis a mid-range of 4.8 mole% was chosen. This is equivalent to a ratio of lipid to gramicidin of 20:1 mol/mol, surrounding each gD dimer with 40 molecules of lipid.

It was also investigated which conformation gramicidin adopts in a bilayer of the lipidic cubic phase. Tests with different modes of reconstitution as described above and different bilayer thicknesses of 9.9 MAG, 11.9 MAG, 9.7 MAG were performed. The conformation adopted by gramicidin was determined using circular dichroism. It was determined that gramicidin adopts a mixture of HSHH and DSDH conformations in the lipidic cubic phase with the equilibrium shifted towards the DSDH conformation. Liu and Caffrey found that the ratio HSHH to DSDH is shifted towards HSHH with decreasing bilayer thickness. It should be mentioned that taking spectra in the lipidic cubic phase has its challenges due to strong absorption of the lipid in the region of wavelengths shorter than 210 nm. It is almost impossible to obtain useful information in this range.

In summary, gramicidin can be incorporated into the lipidic cubic phase of a range of lipids. Powder diffraction analysis shows that the lipidic phase remains cubic. Gramicidin adopts an equilibrium between both conformations. The majority of gramicidin dimers are in the DSDH conformation. Gramicidin lipid samples were

prepared. Gramicidin was crystallized from several lipidic cubic phases and three crystal forms (CF1-3) were observed and are described in detail.

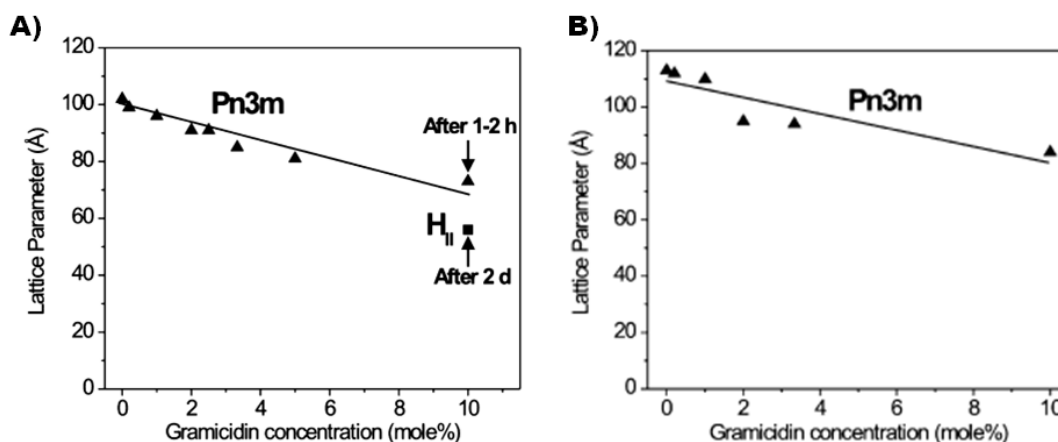


Figure 1.2.5: Sensitivity of the cubic-Pn3m phase of hydrated 9.9 MAG (panel A) to gramicidin D and 9.7 MAG (panel B) at 20 °C. Phase identity and phase microstructure (lattice parameter) were determined using small-angle X-ray diffraction. Gramicidin was incorporated into the lipidic mesophase by the co-dissolution mixing method. The concentration of gramicidin is expressed as mole% corresponding to $100 \text{ (mole gramicidin)}/[(\text{mole gramicidin}) + (\text{mole 9.9 MAG})]$. At the highest concentration of gramicidin used the initial phase observed was of the cubic-Pn3m phase. This transformed to the H_{II} phase over a period of time at room temperature and at an elapsed time of 2 days is completely H_{II}. This figure and figure caption was copied from Liu and Caffrey, 2005.

1.3 Materials and Methods

1.3.1 Materials

Monoolein (9.9 MAG, ref 53, Lots: M-239-A21-Q, M-239-516-5, 356 g/mol; 11.7 MAG, Lots: M-249-J25-S, 356 g/mol) was obtained from NuChek Prep Inc. (Elysian, MN, USA). Linear gramicidin D (Lot: 1345609 13307134, 1,880 g/mol) and 2,2,2-trifluoroethanol (TFE, Lot: S25895-495) were purchased from Sigma-Aldrich (Dublin, Ireland). The lipids 7.7 MAG (Lots: JPL-4-92, JL-3-68, 300 g/mol) and 8.8 MAG (Lot: JPL-2-42, 328 g/mol) were synthesized and purified in-house following established procedures. PEG/Ion (Lot: 212629), Index (Lot: 214407), Crystal Screen I (Lot: 211094), Crystal Screen II (Lot: 211251) and MembFac (Lot: 211423) crystallization screens were obtained from Hampton Research (Aliso Viejo, CA, USA). Wizard I, Wizard II, and Wizard III (Lot: EBS000322200499) were sourced from Emerald BioSystems, Inc. (Bainbridge Island, WA, USA). The crystallization screen MemStart (Lot: 011-121) was obtained from Molecular Dimensions (Suffolk, UK). PEG 400 (Lot: 0001418256) was purchased from Sigma- Aldrich (Dublin, Ireland). Water, with a resistivity of $>18 \text{ M}\Omega \cdot \text{cm}$, was purified using a Milli-Q Water System (Millipore, Bedford, MA, USA) consisting of an Elix 5 UV compartment (Lot: F4HN34349) with a Prograd^{®2} cartridge (Lot: F9HNO1157) to pre-purify water and a Synergy compartment (Lot: F4EN79695B) with a Simpapak^{®1} cartridge (Lot: F9HN06031) to produce highly purified water followed by sterile filtration through a 0.22 μm MilliPAK40 filter (Lot: F5PN18060).

1.3.2 Methods

1.3.2.1 Solubility of Gramicidin in Precipitant Solutions Containing PEG

To determine the solubility of gramicidin in the precipitant solutions that yielded crystals, a 10 mg/mL stock solution of gramicidin in TFE was prepared. This solution was mixed with the relevant precipitant solutions using a volume ratio of 1 part gramicidin solution to 9 parts precipitant solution and shaken by hand for 5 min at

room temperature (RT, 18-23 °C). The precipitate, which formed in all cases, was removed by centrifugation at 16300 x g for 5 min at RT. Absorbance of the supernatant was measured at 280 nm (A_{280}) and converted to peptide concentration using a measured extinction coefficient of $18590 \text{ M}^{-1} \text{ cm}^{-1}$. The latter was determined from the slope of an A_{280} versus gramicidin concentration curve established with the peptide dissolved in a solution composed of TFE, polyethyleneglycol (PEG) 400, and water in the volume ratio 2:9:9 (table 1.3.1). The solubility of gramicidin in precipitant solution for crystal form 2 and 3 was not tested.

Table 1.3.1: *In meso* crystallization conditions in three hosting MAGs and gramicidin solubility in the corresponding precipitant solutions. This table has been taken from Höfer *et al.*, 2011.

	7.7 MAG	8.8 MAG	9.9 MAG
Precipitant ^a	25 % PEG 3350	30 % PEG 8000	20 % PEG 6000
Salt	0.2 M LiSO ₄	0.2 M NH ₄ SO ₄	-
Buffer	0.1 M Bis-Tris pH 5.5	-	0.1 M Bicine pH 9.0
gD solubility ^b	0.04 g/L, 21 μM	0.10 g/L, 53 μM	0.02 g/L, 11 μM

^a Percentage is reported as weight by volume. ^b Gramicidin solubility in precipitant solutions was determined in the absence of lipid.

1.3.1.2 Gramicidin/Lipid Mixture Preparation

Lipids and gramicidin were combined in a molar ratio of 20:1 followed by co-solubilization in 4-5 mL TFE in a 15 mL glass vial. Typically, between 50 and 100 mg of MAG were prepared with the appropriate amount of gramicidin. The solution was shaken by hand for approximately 3 min at room temperature (RT) until optically clear. TFE was removed by evaporation under a stream of nitrogen gas followed by high vacuum drying (10 mbar, Büchi Vac V500, Büchi, Flawil, Switzerland) at RT

for 24 h (Liu and Caffrey, 2005). Samples were stored at RT in the dark under nitrogen for a maximum of one week.

1.3.2.3 Cubic Phase Preparation and *in meso* Crystallization

The cubic mesophase was prepared by combining the dry gramicidin / MAG mixture with Milli-Q water in coupled 0.1 mL gastight RNtype syringes (Hamilton Company, Reno, Nevada, USA) at different weight ratios depending on the lipid, as previously described (Cheng et al., 1998). Specifically, lipid to water weight ratios of 1:1, 1:1, 1:1 and 3:2 were used for 7.7, 8.8, 11.7 and 9.9 MAG, respectively.

Crystallization trials were set up in 96-well glass sandwich plates using 50 nL mesophase and 800 nL precipitant solution with an *in meso* robot, as previously described (Caffrey and Cherezov, 2009; Caffrey and Porter, 2010). Plates were stored in an incubator/imager (RockImager RI1500, Formulatrix, Inc., Waltham, MA, USA) at 20 °C. Crystals for crystal form 1 (CF1) measuring 30 x 30 x 30 μm^3 to 60 x 60 x 60 μm^3 appeared after 3-5 days in a variety of precipitants, all of which contained PEG (figure 1.3.1). Full details regarding the composition of the precipitant solutions that produced crystals for structure determination are reported in table 1.3.1.

Crystals for CF2 were not observed in the initial screening period of 30 days. After approximately 9 months, block shaped crystals were observed in 11.7 MAG and 9.7 MAG. The crystals measured 30 x 30 x 30 μm^3 and grew in identical precipitant solution of precipitant 10 %(w/v) PEG 6000, 5 %(v/v) MPD, 0.1 M HEPES, pH 7.5.

Crystals for CF3 were also not observed in the initial screening period and first observed after 9 months. They had an egg shaped appearance and measured 50 x 30 x 30 μm^3 in precipitant 1 M lithium sulfate, 10 mM nickel chloride, 0.1 M Tris HCL, pH 8.5.

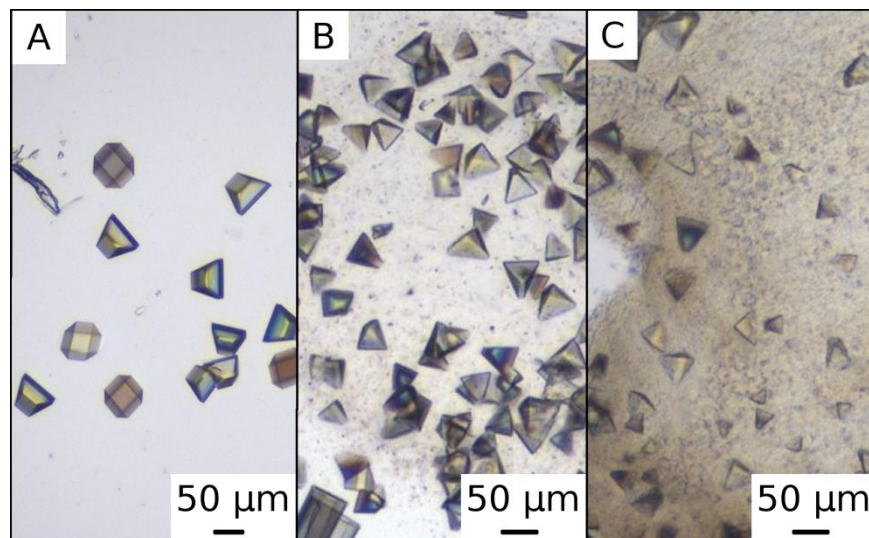


Figure 1.3.1: Crystals of gramicidin growing *in meso* using different hosting lipids. A) 7.7 MAG [28 %(w/w) PEG 2000 MME, 0.1 M Bis-Tris, pH 6.5], (B) 8.8 MAG [30 %(w/v) PEG 8,000, 0.2 M NH_4SO_4], (C) 9.9 MAG [20 %(w/v) PEG 6,000, 0.1 M Bicine, pH 9.0]. This figure has been taken from Höfer et al., 2011.

1.3.2.4 Harvesting Crystals

Plates with crystals were removed from the RI1500 and transferred to a 20 °C room. Cover glass was scored manually using a tungsten carbide glass cutter (Model: 633657, TCT Scriber & Glass Cutter, Silverline, Yeovil, UK) and gently removed from the well. To minimize changes in composition of well contents, the lipid cubic phase drop was immediately overlaid with approximately 1 μl precipitant solution. Cryo-loops (Micro Mounts, MiTeGen, Ithaca, NY, USA) ranging from 30 to 100 μm were used to harvest crystals from the *in meso* bolus. Crystals were directly cryo-cooled in liquid nitrogen and stored in liquid nitrogen until further use (Caffrey and Cherezov, 2009).

1.3.2.5 Data Collection

1.3.2.5.1 Data Collection for CF1

Diffraction data were collected at the General Medicine and Cancer Institute's Collaborative Access Team (GM/CA CAT) beamline 23ID-B at the Advanced Photon Source (APS) with a MAR 300 CCD detector using X-rays ranging from 0.827 to 1.033 Å (table 1.3.2). Because of the small size of the crystals, a minibeam of 10 x 10 μm² was used (Fischetti et al., 2009). The crystals were kept cryo-cooled in a stream of nitrogen at 100 K during data collection with exposures of 1 - 2.3 s for each 1° or 2° oscillation. To ensure that all high-resolution reflections were collected, the detector was moved to the minimum crystal-to-detector distance of 100 mm and the wavelength was set to 0.83 Å (data set 1, 7.7 MAG) and to 0.98 Å (data set 2, 8.8 MAG).

Data reduction was performed in HKL2000 (Otwinowski, 1997) and then converted to structure factors using the CCP4 program TRUNCATE (CCP4, 1994). High resolution data were collected for data set 1 (7.7 MAG) with a highest resolution shell of 1.08 Å. Crystals for data set 2 (8.8 MAG) diffracted to 1.26 Å. For this lipid, a low resolution data set was collected and merged with two high resolution data sets. Crystals for data set 3 (9.9 MAG) diffracted to 1.70 Å. The same test set flag for Rfree calculation assigned to data set 1 was carried over to data sets 2 and 3. All crystals were found to be in space group P2₁ with similar cell dimensions. The asymmetric unit contained three dimers of gramicidin (total molecular weight, ~11.3 kDa) corresponding to a Matthews coefficient (Matthews, 1968) of 2.6 Å³/Da and a solvent content of 52 %. Diffraction and refinement statistics are given in table 1.3.2.

Table 1.3.2: Diffraction Data, Refinement, and Model Statistics for Data Sets 1-3.

This table has been taken from Höfer *et al.*, 2011

Dataset			
	1 ^a	2 ^b	3 ^c
Hosting lipid	7.7 MAG	8.8 MAG	9.9 MAG
Diffraction data			
Space group	P2 ₁	P2 ₁	P2 ₁
Unit cell	a = 30.6 Å b = 62.8 Å c = 30.7 Å β = 100.0°	a = 30.6 Å b = 62.8 Å c = 30.6 Å β = 100.0°	a = 30.5 Å b = 62.6 Å c = 30.5 Å β = 100.0°
X-ray source	23ID-B, APS	23ID-B, APS	23ID-B, APS
Wavelength [Å]	0.83	0.98	1.03
Resolution limits [Å]			
Overall	1.08-30.20	1.26-31.40	1.70-31.30
Highest shell	1.08-1.09	1.26-1.29	1.70-1.74
Number of unique reflections	46435	29856	11751
Completeness [%]	95.0 (65.0)	96.4 (76.9)	94.2 (79.4)
Redundancy [%]	6.3 (3.6)	7.1 (3.2)	3.0 (2.2)
Rmerge [%]	9.5 (41.8)	6.4 (27.9)	6.7 (26.5)
I/σ(I)	15.8 (2.4)	22.6 (3.5)	13.1 (3.3)
Refinement model statistics ^d			
Number of reflections			
Working set	44055 (2238)	28320 (1670)	11147 (706)
Test set	2380 (132)	1536 (94)	604 (29)
Rfactor [%]	13.0	14.8	17.9
Rfree [%]	15.5	16.9	21.4
Protein atoms	816	816	816
Solvent atoms	94	100	107
Average B factor [Å ²]			
All atoms	14.0	14.7	20.3
Peptide	13.0	14.0	18.7
Chain A	11.5	12.4	16.9
Chain B	11.4	12.3	17.2
Chain C	12.8	13.6	18.4

Chain D	14.7	16.0	20.0
Dataset			
	1 ^a	2 ^b	3 ^c
Hosting lipid	7.7 MAG	8.8 MAG	9.9 MAG
Chain E	12.8	13.6	18.7
Chain F	14.7	16.0	18.7
Solvent	22.9	24.9	33.0
PEG	22.6	24.9	32.8
Water	31.5	25.9	40.0
Sodium ion			48.7
Rms deviations from ideal			
Bond lengths [Å]	0.02	0.02	0.02
Angles [°]	1.70	1.70	1.62

^d Values in parentheses are for the highest resolution shells.

1.3.2.5.2 Data Collection for CF2

Diffraction data for crystal form 2 (CF2) grown from 11.7 MAG were collected at ID14-4, Macromolecular Crystallography MAD Beamline at the ESRF with a Q315r ADSC CCD detector using X-rays at 0.9395 Å (table 1.3.3). A complete dataset was obtained by collecting 60 images with 2° oscillation at a sample detector distance of 191.5 mm. The dataset diffracted to a maximum resolution of 1.8 Å.

Data reduction was performed in Xia2. Mosfilm (Battye et al., 2011) was used for indexing and integration as a subroutine in Xia2. This was followed by Pointless and Scala. Data collection statistics are given in table 1.3.3. The crystals were found in space group P2₁2₁2 with unit cell dimensions a = 45.7 Å, b = 63.4 Å, c = 20.9 Å, $\alpha = \gamma = \beta = 90^\circ$.

1.3.2.5.3 Data Collection for CF3

Diffraction data were collected at ID14-4 at the ESRF with a Q315r ADSC CCD detector and using X-rays at 0.9395 Å (table 1.3.3). A complete 1.70 Å dataset was obtained by collecting 115 images with 2° oscillation at a sample detector distance of 267 mm.

Data reduction was performed in HKL2000 (Otwinowski, 1997) and then converted to structure factors using the CCP4 program TRUNCATE (CCP4, 1994). Data collection statistics are given in table 1.3.3. The crystal was found in space group $P2_1$ with unit cell dimensions $a = 24.0 \text{ \AA}$, $b = 42.0 \text{ \AA}$, $c = 32.4 \text{ \AA}$, $\alpha = \gamma = 90^\circ$, $\beta = 107.3^\circ$.

1.3.2.6 Phase Determination and Refinement

1.3.2.6.1 Phase Determination and Refinement of Datasets from CF1

The phase problem for the highest resolution data set (data set 1, 7.7 MAG) was solved by molecular replacement with the program Phaser (McCoy et al., 2007) using a poly-alanine pruned DSDH dimer (PDB entry 1AL4) and data truncated to 1.7 \AA . Refinement was performed using restrained maximum likelihood as implemented in Refmac (Murshudov et al., 1997). Side chains were modeled in using Coot (Emsley and Cowtan, 2004). Positional and isotropic B factor parameters were refined for each atom. Alternative conformations for side chains were added where suggested by $|F_o| - |F_c|$ maps with an initial occupancy of 0.5. In cases where B factors and maps justified it, occupancy was set to 0.3 and to 0.7 so that maps were clear and neighboring atoms had similar B factors. After this stage of refinement, PEG, lipid, water molecules, and ions were added to the model based on electron density, standard geometrical, and chemical constraints, and presence in the crystallization conditions. Molecules were subsequently deleted if they were not visible at the 1.0 sigma level in $2|F_o| - |F_c|$ electron density maps. Molecules assigned as PEG or lipid were also deleted if no clear hydrogen bonds could be assigned. Anisotropic B factor refinement was then turned on followed by addition of hydrogens in riding positions and refined until convergence.

Table 1.3.3: Diffraction data, refinement and model statistics for CF2 and CF3

Crystal form	CF2	CF3
Lipid matrix	11.7 MAG	11.7 MAG
Space group	P 2 ₁ 2 ₁ 2	P2 ₁
Unit cell	a = 45.7 b = 63.2 c = 20.8 $\alpha=\beta=\gamma=90^\circ$	a = 24.0 Å b = 42.0 Å c = 32.4 Å $\beta = 107.3^\circ$
X-ray source	ID14-4, ESRF	ID-14-4, ESRF
Wavelength [Å]	0.94	0.94
Resolution limits [Å]		
Overall	37.0 – 1.8	22.9 – 1.7
Highest shell	1.83 – 1.88	1.70 – 1.75
No. of unique reflections	5673	6612
Completeness [%] ^c	98.9 (100)	96.9 (78.9)
Redundancy [%] ^c	4.2 (3.5)	4.7 (3.9)
R _{merge} [%] ^c	5.2 (23.9)	5.2 (24.6)
I / $\sigma(I)$ ^c	21.8 (4.6)	23.0 (5.0)
Refinement model statistics ^c		
No. of reflections		
Working set	5386 (287)	6306 (352)
Test set	408 (22)	306 (26)
R _{factor} [%]	16.9	16.6
R _{free} [%]	22.3	21.1
Protein atoms	542	544
Solvent atoms	47	46
Average B factor [Å ²]		
All atoms	17.3	18.1
Peptide	16.1	16.4
Chain A	17.8	17.4
Chain B	16.3	18.5
Chain C	15.4	14.2
Chain D	14.8	15.7
Solvent		
Lipid	-	40.1
PEG	31.0	-
Water	31.5	30.1
Rms deviations from ideal		
Bond lengths [Å]	0.019	0.019
Angles [°]	1.82	1.83

A similar structure determination and refinement protocol was followed for data set 2 (8.8 MAG) using as molecular replacement search model a poly-Ala version of the final, previously refined gramicidin dimer (2Y5M, 7.7 MAG). Structure determination and refinement for data set 3 (9.9 MAG) are described in Höfer *et al.*, 2010.

1.3.2.6.2 Phase Determination and Refinement for CF2

The phase problem for the data set of CF2 was determined by molecular replacement with program Phaser using a poly-Ala pruned DSDH dimer (PDB entry 2XDC) and data truncated to 1.9 Å. Refinement was performed using restrained maximum likelihood as implemented in Refmac. Side chains were modeled in using Coot. Alternative conformations, PEG, MPD, and water molecules were added back into electron density where it was suggested. The data were refined. Weights were adjusted such that the root mean squared deviation (r.m.s.d.) from ideal was less than 0.02 Å for bond length and 2° for angles. Anisotropic B factors were not turned on since data to parameter ratio did not suggest it.

There are 2 dimers of gramicidin in the asymmetric unit giving a total molecular weight of approximately 7.5 kDa. The corresponding Matthews coefficient is 2.0 Å³ / Da, which corresponds to a solvent content of 38.8 %.

1.3.2.6.3 Phase Determination and Refinement for CF3

The phase problem for the data set was determined by molecular replacement with program Phaser (McCoy *et al.*, 2007) using a poly-Ala pruned DSDH dimer (PDB entry 2XDC, (Höfer *et al.*, 2010)) and data truncated to 1.9 Å. Refinement was performed using restrained maximum likelihood as implemented in Refmac (Murshudov *et al.*, 1997). Side chains were modeled using Coot (Emsley and Cowtan, 2004). Positional and isotropic B factor parameters were refined for each atom. Alternative conformations for side chains were added where suggested by |Fo|-|Fc| maps with an initial occupancy of 0.5. To obtain clear maps and similar B factors in neighboring atoms the occupancy was further adjusted up or down in steps of 0.05.

After this stage of refinement, lipid and water molecules were added to the model based on electron density, standard geometrical and chemical constraints, and presence in the crystallization conditions. Molecules were subsequently deleted if they were not visible at the 1.0 sigma level in $2|F_o| - |F_c|$ electron density maps.

There are 2 dimers of gramicidin in the asymmetric unit giving a total molecular weight of approximately 7.5 kDa. The corresponding Matthews coefficient (Matthews, 1968) is $2.07 \text{ \AA}^3 / \text{Da}$, which corresponds to a solvent content of 40.7 %.

1.3.2.7 Structure Validation of CF1-3

The quality of the structural models for CF1 was evaluated during the course of refinement using MolProbity (Chen et al., 2010), RAMPAGE (Lovell et al., 2003), and Sfccheck (Vaguine et al., 1999). All programs show bond length and angle values as expected for structures at these resolutions. Ramachandran analysis (Ramachandran and Sasisekharan, 1968) for each structure shows over 87 % of the residues in the most favored regions, over 11 % of the residues in the allowed regions and no outliers (table 1.3.4). These results are consistent with what is found for other gramicidin structures in the same conformation (see PDB entries 1AL4, 1ALZ, and 1ALX) (Burkhart et al., 1998a). The coordinates and structure factors for data set 1 (7.7 MAG), data set 2 (8.8 MAG), and data set 3 (9.9 MAG) have been deposited in the PDB under code identifiers: 2Y5M, 2Y6N, and 2XDC, respectively.

The quality of the structural models for CF2 and CF3 was evaluated as described for CF1. All programs show bond length and angle values as expected for structures at these resolutions. These results are consistent with what is found for other gramicidin structures in the same conformation (table 1.3.4).

The coordinates and structure factors for CF3 have been deposited in the PDB under code identifier: 3ZQ8.

Table 1.3.4: Structure validation (This table has been adapted from Höfer et al., 2011)

Crystal form	Crystallization matrix / solvent	PDB	Ramachandran			Molprobitry clash score (percentile)
			Residues in favored regions, Number (%)	Residues in allowed regions, Number (%)	Residues in outlier regions, Number (%)	
CF1	<i>in meso</i> , 9.9 MAG	2XDC	68 (87.2)	10 (12.8)	0 (0)	5.31 (94)
	<i>in meso</i> , 8.8 MAG	2Y6N	69 (88.5)	9 (11.5)	0 (0)	1.1 (99)
	<i>in meso</i> , 7.7 MAG	2Y5M	68 (87.2)	10 (12.8)	0 (0)	2.84 (93)
CF2	<i>in meso</i> , 11.7 MAG	Appendix 5	41 (78.8)	11 (21.2)	0 (0)	5.18 (96)
CF3	<i>in meso</i> , 11.7 MAG	3ZQ8	42 (80.8)	10 (19.2)	0 (0)	4.2 (97)
Propanol solvate		1AL4	21 (80.7)	5 (19.3)	0 (0)	22.73 (2)*
Methanol solvate		1ALX	22 (84.6)	4 (15.4)	0 (0)	33.79 (0)*
Ethanol solvate		1ALZ	20 (77.0)	6 (13.0)	0 (0)	21.28 (2)*

^a The ϕ and ψ torsion angles for all amino acids were calculated using RAMPAGE (Lovell et al., 2003). The torsion angle sign for the D- amino acids was reversed, as described by Glowka et al., 2005 (Glowka et al., 2005) moving them into the most favored region of the Ramachandran plot. All glycines are in glycine allowed regions.

*Structures 1AL4, 1ALX, and 1ALZ were submitted to the PDB before Molprobitry was used as a program to check the quality of a structure. Molprobitry assigned these structures with a high clashscore and low clash score percentile. The majority of clashes reported are clashes between peptide and solvent, which could be easily be prevented and fixed with recent crystallographic tools.

1.4. Results

The small membrane peptide gramicidin was crystallized from the lipidic cubic phase of different lipids. Three different crystal forms were found (CF1, CF2, CF3). CF1 was crystallized from 7.7, 8.8, and 9.9 MAG. These hosting lipids differ in their bilayer thickness. The purpose for using different hosting lipids was to see the effect they have on the crystal form and gramicidin D. CF2 and CF3 crystallized from 11.7 MAG. In the following section each crystal form is described in detail.

1.4.1 Crystallization and Structure Determination of CF1

The purpose of this study was to determine if the character of the hosting mesophase, dictated by the lipid used to form it, would impact the crystallization and structure of the transmembrane polypeptide gramicidin. What we found was unexpected: the three lipids used to create the hosting mesophases for *in meso* crystallogenesis yielded identical crystals and identical structures for gramicidin. In all cases, the crystals grew to their full size within a week at room temperature with maximum dimensions in the 30-60 μm range. The space group in every instance was $P2_1$ with unit cell metrics of $a = c = 30.6 \text{ \AA}$, $b = 62.8 \text{ \AA}$, $\beta = 100.0^\circ$ to within error and the structures refined to reasonable R factors with good geometry (table 1.3.4). PEG was a common precipitant ingredient and was identified in the final crystal. For each of the three hosting lipids, the phase from which crystals were harvested was of the sponge type (Cherezov et al., 2006).

The structure of gramicidin crystallized *in meso* in the three hosting lipids was solved initially by molecular replacement with a gramicidin model obtained using crystals grown from n-propanol (Burkhart et al., 1998a). The *in meso* and n-propanol structures are remarkably similar. The peptide exists as a left-handed, intertwined helical homodimer in an antiparallel arrangement. The asymmetric unit consists of three dimers and each unit cell has a total of six dimers. The two monomers within a dimer can be viewed as β -strands that associate to form a two-stranded lath or ribbon stabilized by either 24 or 26 interstrand hydrogen bonds, depending on the dimer,

typical of β -sheets. Because the sequence consists of alternating D- and L- residues, side chains are on one side of the ribbon. Their bulk contributes to a curving of the ribbon and to the formation of a β -helix with ~ 5.6 residues per turn (a $\beta^{5.6}$ -helix). Individual dimers are approximately 35 Å long and have an outer diameter of 18 Å. The core of the dimers is sealed and thus incapable of accommodating ions. Dimers are arranged in layers with their long axis oriented approximately normal to the layer plane. This so-called Type I or layered packing (Michel, 1982) is consistent with the proposed mechanism for crystallization *in meso* (figure 1.1.6); it has been observed in all crystal structures obtained to date by the *in meso* method (Caffrey, 2009).

Crystals formed from the three different lipids have the same crystal packing and composition. For simplicity therefore the following discussion will refer primarily to the highest resolution structure (data set 1, 7.7 MAG). As noted, there are three dimers within an asymmetric unit of the gramicidin crystal. For purposes of discussion, the three will be referred to as dimers AB, CD, and EF with monomers identified by individual letters. Within the asymmetric unit, the AB and CD dimers associate to form a V-shaped object. The pointed end of the V is created by the N-terminal ends of chain A and chain D that are in van der Waals contact. The angle of the V is 25° and its splayed end has the main chain of chains A and D separated by 14 Å. In contrast, the long axes of dimers CD and EF are almost parallel with a dimer center to center separation of around 20 Å. Dimers CD and EF are almost identical with an r.m.s.d. in $C\alpha$ atom positions of 0.03 Å. Small structural differences exist between dimer AB and dimers CD and EF as reflected in slightly higher r.m.s.d. values.

In addition to the peptide, each asymmetric unit within the *in meso* gramicidin crystal contains four molecules of PEG and two molecules of water. The four PEG molecules have different lengths and are located in distinct parts of the asymmetric unit. The longest, with 16 ethylene glycol units (PEG-A), has its bulk located toward the middle of the plane containing the gramicidin dimers (figure 1.4.1). The midsection of PEG-A runs approximately perpendicular to the long axis of all dimers

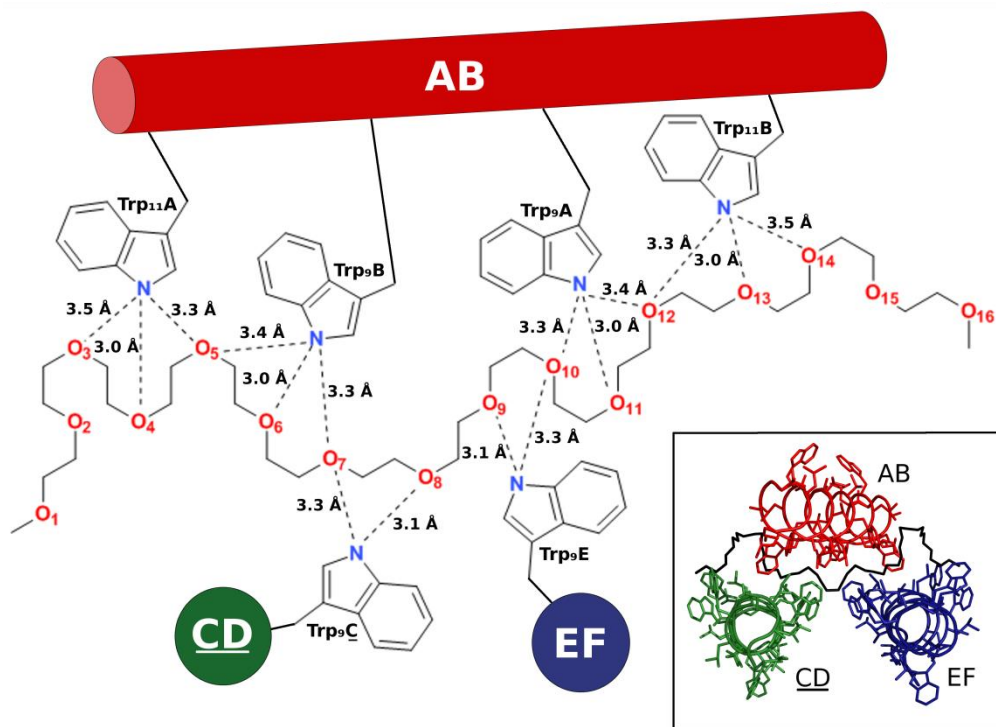


Figure 1.4.1: Schematic representation of hydrogen bonding (dashed lines) between oxygen atoms (red, the subscript identifies the ethylene glycol unit number) of PEG-A and tryptophan N ϵ atoms (blue) of gramicidin. Dimers AB (red cylinder) and EF (purple circle) derive from one asymmetric unit while dimer CD (green circle) is from an adjacent asymmetric unit. The inset shows dimers AB (red), EF (purple) and CD (green) in cartoon form with PEG-A in black stick representation. Crystals used for the structure shown were grown in the lipidic mesophase prepared with 7.7 MAG. This figure has been taken from Höfer *et al.*, 2011.

and is sandwiched between dimer AB and dimers EF and CD (the underscore indicates an adjacent asymmetric unit, see figure 1.4.1). The polymer chain therefore forms important crystal contacts within the layers of the Type I crystal lattice. The oxygen of its ethylene glycol units hydrogen bonds extensively with the N ϵ of the indole side chains of tryptophan 9 and tryptophan 11 that extend away from the long axis of the intertwined double helices (figure 1.4.1). Another two PEG molecules,

with seven ethylene glycol units each (PEG-B, PEG-B₀), are located close to the C-terminus of gramicidin chain A. They create crystal contacts between layers in the crystal lattice by hydrogen bonding with tryptophan 15 of one layer and the ethanolamine cap of gramicidin from another layer. It is possible that PEG-B and PEG-B₀ derive from a single PEG chain but with poor electron density between the two they have been refined as separate entities. The maps for all three gramicidin structures reported in this work also show a sphere of electron density between PEG-B and PEG-B₀. Modeling and refining this volume as water, ammonium, or sodium resulted in B factors that differ significantly from those of the surrounding atoms or resulted in improbable hydrogen bonding for a positively charged species. Because this density could not be reasonably fit it was left un-modeled. The last PEG with two ethylene glycol units (PEG-C) runs parallel to the long axis of dimer CD and is anchored by two hydrogen bonds to tryptophan 11 from chain D and a symmetry related chain F.

Noteworthy is the observation that the resolution of the gramicidin structures solved in this study scales with the chain length of the MAG used to create the hosting mesophase. Thus, resolutions of 1.08, 1.26, and 1.70 Å were obtained in 7.7 MAG, 8.8 MAG, and 9.9 MAG, respectively. In our hands, the shortest chain lipid, 7.7 MAG, generally produces data of the highest quality and resolution in separate *in meso* crystallization trials. While the number of instances is not great, this short chain MAG is generating structure-grade crystals where the default lipid, 9.9 MAG, is either not producing crystals or the crystals it does generate are of poor quality. It is not clear why 7.7 MAG is out-performing the other lipids. In addition to producing a thinner, less curved bilayer, it is more prone to forming the sponge phase. A looser, less strained bilayer characteristic of the sponge phase may favor the growth of bigger and better ordered crystals as has been speculated on previously (Caffrey, 2008; Cherezov et al., 2006).

1.4.2 Crystallization and Structure Determination of CF2

The small membrane peptide gramicidin was crystallized from the lipidic cubic phase of monovaccenin (11.7 MAG). The peptide crystallized in both cases in space group $P2_12_12$ with $a = 45.7 \text{ \AA}$, $b = 63.3 \text{ \AA}$, $c = 20.9 \text{ \AA}$ and $\alpha = \beta = \gamma = 90^\circ$. The major precipitant was PEG. Crystals had a maximum size of approximately 30 \mu m in all directions.

The structure of gramicidin was solved by molecular replacement using gramicidin model (2XDC) (Höfer et al., 2010; Höfer et al., 2011), the dataset from 11.7 MAG (Höfer et al., 2010; Höfer et al., 2011). The peptide was solved in the left-handed antiparallel DSDH conformation. The asymmetric unit consists of two dimers referred to here as dimer AB with chains A and B and dimer CD with chains C and D. Each unit cell has a total of 8 dimers. Two monomers within a dimer are held together by interstrand hydrogen bonds. Dimer AB is stabilized by 28 and dimer CD by 26 hydrogen bonds. The hydrogen bonding pattern formed is typical of β -sheets. Because the sequence consists of alternating D- and L- residues side chains are on one side of the ribbon. Their bulk contributes to a curving of the ribbon and to the formation of a β -helix with ~ 5.6 residues per turn (a $\beta^{5.6}$ -helix). Individual dimers are approximately 35 \AA long and have an outer diameter of 18 \AA . Due to the pitch the inner core of the dimers is too tight to accommodate ions.

Crystals are packed in the layered or type I crystal arrangement (typical for *in meso* crystals). Eight hydrogen bonds form the Type I crystal contacts between layers of gramicidin dimers. Unit cell axis b runs parallel to the layers, whereas unit cell axis a and c run perpendicular. Within a layer dimers are arranged to form V shaped objects. The pointed end of the V is created by the N-terminal ends of chain A and chain D' that are in van der Waals contact. The angle of the V is 25° and its splayed end has the main chain of chains A and D separated by 14 \AA . Dimers AB and CD are similar with an r.m.s.d in peptide backbone positions of 0.03 \AA .

1.4.2.1 Ligand

Aside from peptide and two molecules of water, one chain of polyethylene glycol (PEG) was modeled into electron density. The PEG chain has a length of 13 polyethylene glycol units $(-C-C-O-)_n$ and is located between dimers AB, AB', CD', CD'' from adjacent asymmetric units. The bulk is located in the mid section of the gramicidin dimers (figure 1.4.1). This PEG molecule forms crystal contacts within the layer of the type I lattice. Very similar crystal contacts have been seen in previous crystals structures of gramicidin (Höfer et al., 2010; Höfer et al., 2011) (PDB ID 2XDC, 2Y5M, 2Y6N). The oxygen atom of its ethylene glycol units hydrogen bonds extensively with the N_ϵ of the indole side chains of Trp₉ and Trp₁₁. It is interesting to note that only the inner ring tryptophan residues (Trp₉ and Trp₁₁) are involved in this kind of crystal contact. In addition, one molecule of 2-methylpentan-2,4-diol (MPD) was modeled into electron density. The alcohol groups of MPD form one hydrogen bond with oxygen 3 of PEG and one with the N_ϵ of the indole side chains of Trp₁₁ C'.

1.4.2.2 Tryptophan Residues and their Interactions

Each dimer of gramicidin has 8 tryptophan residues (see chapter 2.3). Residues 9 and 11 belong to the inner ring tryptophan residues, whereas residues 13 and 15 are considered the outer ring residues (Sun et al., 2008). All inner ring tryptophan residues in the asymmetric unit in this crystal form are involved in hydrogen bonding with PEG with the exception of tryptophan residue 9C. The hydrogen of the N_ϵ nitrogen of this side chain forms a hydrogen bond with an adjacent tryptophan residue (Trp₁₁D). The distance between N_ϵ of Trp₉C and the center position of the benzene ring is 3.1 Å. The angle between N_ϵ of Trp₉C the center position of the benzene ring and CD1 of Trp₁₁D was measured to be 87°. This type of hydrogen bond is not very common, but has been observed for tryptophan residues that are in close proximity (Chakrabarti and Bhattacharyya, 2007). The outer ring tryptophan residues do not form hydrogen bonds with molecules that were modeled into electron density.

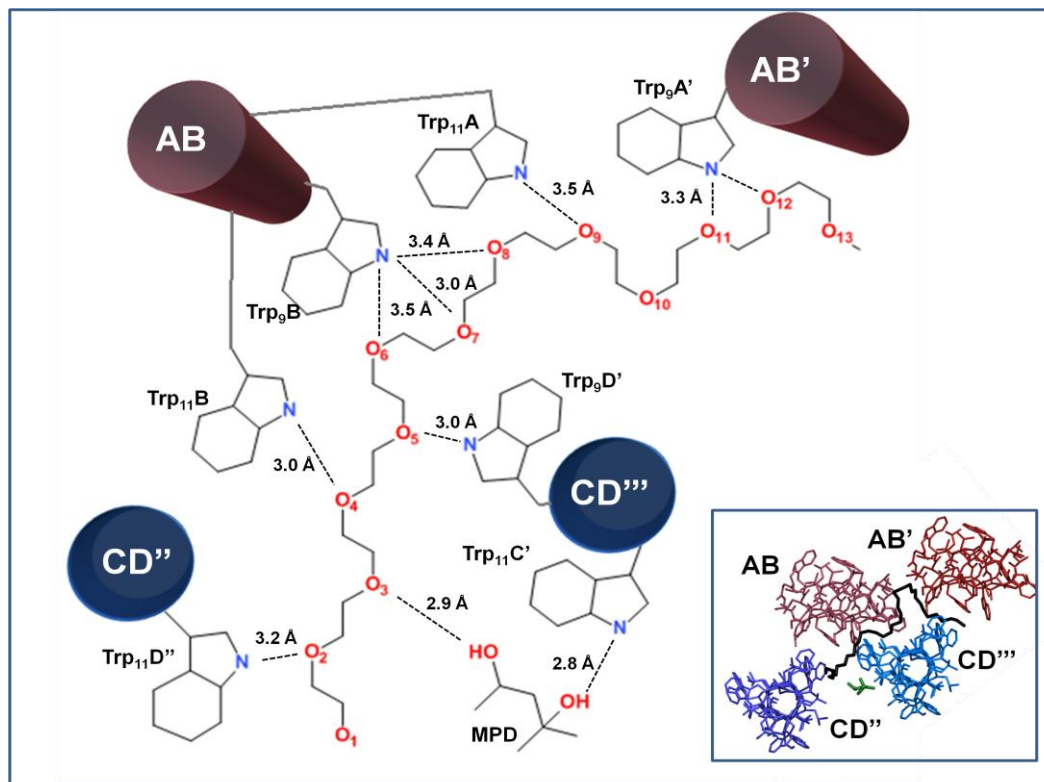


Figure 1.4.2: Schematic representation of hydrogen bonding (dashed lines) between oxygen atoms (red, the subscript identifies the ethylene glycol unit number) of PEG and tryptophan N ϵ atoms (blue) of gramicidin. PEG is coordinated by dimers AB (raspberry cylinder), AB' (firebrick cylinder), CD'' (blue cylinder), and CD''' (marina blue cylinder). The dimers originate from four adjacent asymmetric units indicated by ' . For completeness the hydrogen bonding with MPD is also shown. The inset shows dimers AB (raspberry), AB' (firebrick), CD'' (blue), and CD''' (marina) in cartoon form with PEG and MPD in black and green stick representation. Crystals used for the structure shown were grown in the lipidic mesophase prepared with 11.7 MAG. This figure has been adapted from a figure published as Höfer *et al.*, 2011.

1.4.3 Crystallization and Structure Determination of CF3

In this crystal form gramicidin was crystallized from the lipidic cubic phase using 11.7 MAG (monovaccenin) as a crystallization matrix. Space group $P2_1$ with unit cell metrics of $a = 24.0 \text{ \AA}$, $b = 42.0 \text{ \AA}$, $c = 32.4 \text{ \AA}$, $\alpha = \gamma = 90^\circ$, $\beta = 107.3^\circ$ was determined for this crystal form (CF3). The structure refined to R factors of 16.6 % and an R free of 21.2 % with good geometry (table 1.3.3). The structure of gramicidin was solved initially by molecular replacement with a gramicidin model (2XDC, (Höfer et al., 2010)), which was obtained from crystals grown *in meso*. The peptide exists as a left-handed DSDH conformation as seen in CF1 and CF2. The asymmetric unit consists of two dimers, which are referred to as dimer AB and dimer CD. The two dimers are similar with a main chain r.m.s.d. of 0.93 \AA (table 1.4.3). Each unit cell contains a total of four dimers. The two monomers within a dimer form a β -helix with ~ 5.6 residues per turn (a $\beta^{5.6}$ -helix) and are stabilized by either 26 or 30 inter-strand hydrogen bonds, depending on the dimer. Side chains point away from the peptide backbone due to the alternating nature between D and L amino acids in the peptide sequence. Individual dimers are approximately 35 \AA long and have an outer diameter of 18 \AA . The core of the dimers is sealed and thus incapable of accommodating ions. This structure is remarkably similar to previously described *in meso* (Höfer et al., 2010; Höfer et al., 2011) and alcohol based crystallization structures of gramicidin (Burkhart et al., 1998a). Type I crystal packing typically found for *in meso* structures and for gramicidin structures of CF1 and CF2 is observed. The dimers are arranged in layers with their long axis oriented approximately normal to the layer plane. The layers are held together by 8 hydrogen bonds in one asymmetric unit / per two connecting asymmetric units.

Within the asymmetric unit, dimers AB and CD associate to form a V-shaped object when viewed along unit cell axis a . The pointed end of the V is created by the N-terminal ends of chain A and chain C that are in van der Waals contact. The angle of the V is 25° and its splayed end has the main chain of chains A and D separated by 14 \AA .

1.4.3.1 Description of Bound Lipid Molecules

In addition to the peptide and two water molecules, each asymmetric unit within the *in meso* gramicidin crystal contains a total of three molecules of 11.7 MAG. One nearly complete molecule of 11.7 MAG (MPG1) is present (missing only three atoms at the end of the aliphatic chain), as well as two lipid headgroups, the tails of which are not firmly enough bound to be seen in electron density. The lipids originate from an 11.7 MAG crystallization matrix. The most complete 11.7 MAG molecule (figure 1.4.3) runs parallel to the long axis of dimer CD with its headgroup being close to the N-terminus of monomer D and close to the interface between crystal layers. The overall orientation of the lipid molecule is perpendicular to the crystal layer. The lipid tail is located in a more hydrophobic area of the gramicidin dimer. The lipid headgroup is held in place by 3 hydrogen bonds (table 1.4.1).

The remaining two lipid molecules are two lipid head groups (MPG2, MPG3), which are also located towards the ends of the dimer and are held in place by three or two hydrogen bonds, respectively (table 1.4.1). The two headgroups are located at the N- and C- termini of the gramicidin dimers and are closer to the outside of a crystal layer oriented in the lipid interface.

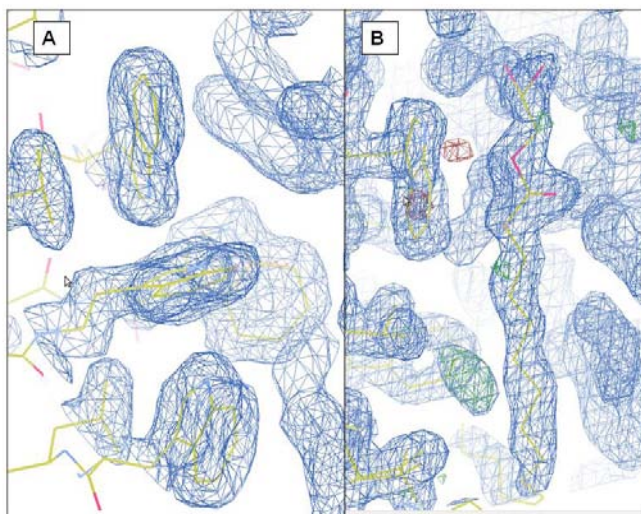


Figure 1.4.3: Representation of the interaction of tryptophan residues in N-H $\cdots\pi$ configuration (A); 11.7 MAG lipid (MPG1) in electron density (B). Taken from PDB ID 3ZQ8.

Table 1.4.1: Hydrogen bonds between lipid headgroups of MPG1, MPG2, MPG3 and gramicidin or water; distances between 2.8 and 3.6 Å were considered

Lipid atom	Hydrogen bond length (Å)	Bonding partner
Lipid MPG 1		
Carbonyloxygen	2.97	N _ε Trp ₉ D alternative conformation A
O1	3.39	N _ε Trp ₁₁ C
O2	3.03	N _ε Trp ₁₁ C
O3	-	-
Lipid MPG 2		
Carbonyloxygen	2.82	N _ε Trp ₁₁ A
O1	-	-
O2	3.26	O For ₁ C
O3	2.39	Water ₂
Lipid MPG 3		
Carbonyloxygen	2.87	N _ε Trp ₁₅ C
O1	-	-
O2	-	-
O3	3.54	O Val ₁ C

1.4.3.2 Interaction between Tryptophan Residues

As described previously, gramicidin is rich in tryptophan residues with eight of them per dimer (Sarges and Witkop, 1965). Tryptophan residues 13 and 15 (outer ring tryptophan residues) are localized in the interface of the bilayer. The inner ring tryptophan residues 9 and 11 are found in the interface of the bilayer for the HSHH conformation or buried in the hydrophobic portion of the membrane bilayer as observed for the closed DSDH conformation (Kelkar and Chattopadhyay, 2007). The latter is energetically less favorable, but the closed DSDH conformation has been observed in bilayers (Dzikovski et al., 2011; Kelkar and Chattopadhyay, 2007).

In CF2 and CF3 (PDB ID: 3ZQ8) weak hydrogen bonds between tryptophan residues are observed in form of an N-H $\cdots\pi$ interaction (figure 1.4.3). The side chains are oriented such that the aromatic moiety (indole ring) is perpendicular to the N $_{\epsilon}$. The distance between the N $_{\epsilon}$ nitrogen and the indole ring is between 3.3 and 3.6 Å. The angle between the two indole rings ranges between 80 and 100 ° (table 1.4.2). These weak hydrogen bonds have been observed between tryptophan residues within a dimer and between dimers. Four such hydrogen bonds have been observed in CF3, in particular for the inner ring tryptophan residues 9 and 11 (N $_{\epsilon}$ Trp₉A - π Trp₁₁C - N $_{\epsilon}$ Trp₉B, N $_{\epsilon}$ Trp₉B - π Trp₁₁A - N $_{\epsilon}$ Trp₉C, N $_{\epsilon}$ Trp₁₁D - π Trp₁₃B, N $_{\epsilon}$ Trp₉D - π Trp₁₁C). The multiple hydrogen bonds involving tryptophans help stabilize this structure within the crystal layer.

This type of hydrogen bond is an extension of the ‘conventional’ hydrogen bond, in which the distance between two electronegative atoms is shorter than their van der Waals radii (Chakrabarti and Bhattacharyya, 2007). N-H $\cdots\pi$ interaction between tryptophan residues has been observed in proteins and is considered to have a stabilizing effect on protein structure (Chakrabarti and Bhattacharyya, 2007). This kind of hydrogen bond is weaker than the conventional type hydrogen bond. Two tryptophan side chains can interact in different ways as can be seen schematically in figure 1.4.4. The most frequently observed interaction consists of one tryptophan residue aligning normal to the aromatic ring of a second tryptophan residue, such that the N $_{\epsilon}$ hydrogen is directly over the indole ring of the second residue (figure 1.4.4)

(Chakrabarti and Bhattacharyya, 2007; Samanta et al., 1999, 2000). The outer ring tryptophan residues (Trp13, Trp15) form a variety of hydrogen bonds including the ‘conventional’ hydrogen bond between lipid headgroup and Ne hydrogen and N-H $\cdots\pi$ interaction.

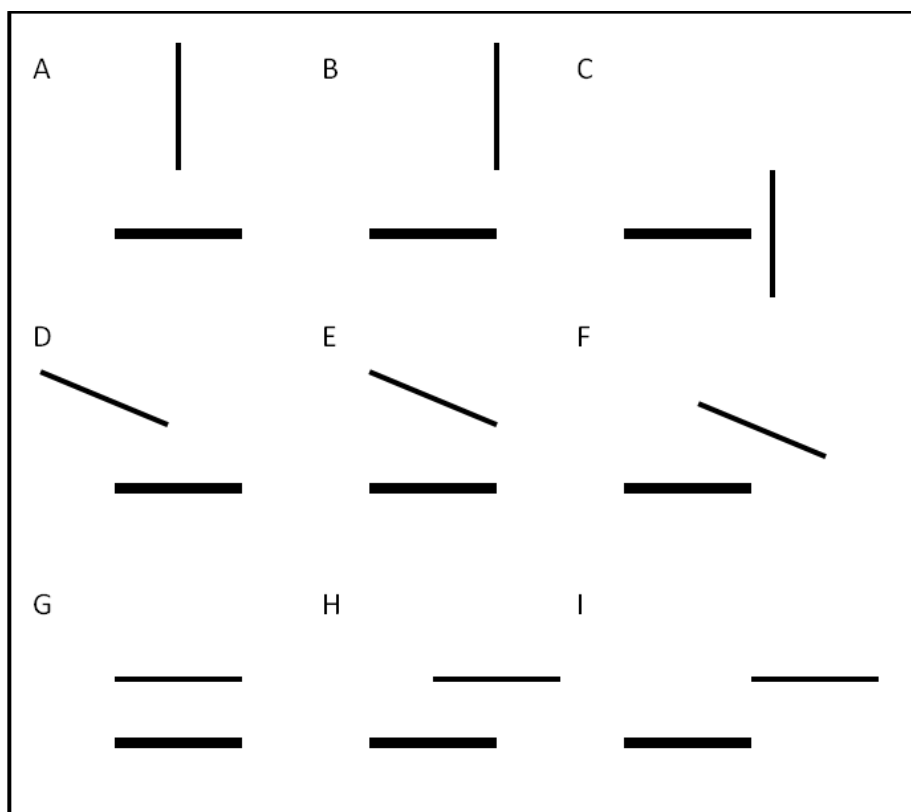


Figure 1.4.4: Interactions between two tryptophan residues observed in crystal structures of proteins. The thick and thin black lines represent the indole ring looking along the long axis of the ring. A) face – edge, B) offset – edge, C) edge – face, D) face – face, E) offset – tilted, F) edge – tilted, G) face – face, H) offset – face, I) edge – edge interaction. The most commonly observed interaction is the face – edge interaction as described in A. This figure has been adapted from, Chakrabarti and Bhattacharyya, 2007, figure 14.

Table 1.4.2: Hydrogen bonds formed by tryptophan residues in dimers AB and CD of CF3

Chain	Tryptophan residue	Portion	Distance (Å)	Hydrogen bond bonding partner	Angle
A	9	N _ε	3.55	π Trp ₁₁ C	100.5° (CD ₁ Trp ₁₁ C – π Trp ₁₁ C – N _ε Trp ₉ A)
	11	Indole ring	3.28	N _ε Trp ₉ B	83.7° (CD ₁ Trp ₁₁ A – π Trp ₁₁ A – N _ε Trp ₉ B)
			3.34	N _ε Trp ₉ C	95.9° (CD ₁ Trp ₁₁ A – π Trp ₁₁ A – N _ε Trp ₉ C)
		N _ε	2.82	lipid ₂	
	13	-	-	-	-
	15	-	-	-	-
B	9	N _ε	3.28	π Trp ₁₁ C	100.5° (CD ₁ Trp ₁₁ C – π Trp ₁₁ C – N _ε Trp ₉ A)
	11	-	-	-	-
	13	Indole ring	3.34	N _ε Trp ₁₁ D	79.0° (CD ₁ Trp ₁₃ B – π Trp ₁₃ B – N _ε Trp ₁₁ D)
	15	N _ε	3.04	OH Eta ₁₆ C	
C	9	N _ε	3.34	π Trp ₁₁ A	95.9° (CD ₁ Trp ₁₁ A – π Trp ₁₁ A – N _ε Trp ₉ C)
	11	N _ε	3.03	O ₂ lipid ₁	
			3.39	O ₁ lipid ₁	
		Indole ring	3.28	N _ε Trp ₉ B	100.5° (CD ₁ Trp ₁₁ C – π Trp ₁₁ C – N _ε Trp ₉ A)
			3.55	N _ε Trp ₉ A	100.5° (CD ₁ Trp ₁₁ C – π Trp ₁₁ C – N _ε Trp ₉ A)
	13	N _ε	2.97	Water ₂	
	15	N _ε	2.87	lipid ₃	
D	9	N _ε	2.97	lipid ₁	
				π Trp ₁₁ C	
	11	N _ε	3.34	π Trp ₁₃ B	79.0° (CD ₁ Trp ₁₃ B – π Trp ₁₃ B – N _ε Trp ₁₁ D)
	13	-	-	-	-
	15	-	-	-	-

1.4.4 Comparison between *in meso* Grown Crystal Structures (CF 1-3) of gramicidin

As described in the introduction, gramicidin was used as an example peptide to prove the applicability of the *in meso* method to peptides and small proteins. So far, gramicidin has been crystallized in three different crystal forms (CF1(Höfer et al., 2010; Höfer et al., 2011), CF2, CF3) using the lipidic cubic phase consisting of lipid matrixes of different thicknesses and with different locations of the double bond. CF1 was crystallized using lipids 7.7, 8.8, 9.9, 9.7 MAG. CF2 was obtained from 11.7 MAG. CF3 was also crystallized from 11.7 MAG. The major precipitant in the crystallization of CF1 and CF2 was PEG. PEG was also identified in the electron density. There, it was identified through the characteristic hydrogen bonding pattern that it formed with tryptophan residues. Long and narrow lobes of electron density that did not show this feature were left uninterpreted, because they could be either PEG or lipid molecules since the crystallization cocktail contained both. For this reason lipid molecules were not identified in CF1 and CF2 but could well be present. CF3 on the other hand crystallized without PEG. Lipid molecules were identified and modeled into the electron density.

The unit cell matrices were different for each crystal form. The number of dimers in the asymmetric unit varied. CF1 contained three dimers in the asymmetric unit, whereas forms 2 and 3 each consisted of two dimers in the asymmetric unit.

1.4.4.1 DSDH Conformation Observed in all Three Crystal Forms

The conformation found in all three crystal forms was the closed DSDH conformation, in which the dimers are linked together in an antiparallel fashion. The dimers are similar with an average r.m.s.d. value of 1 Å (table 1.4.3). Closer examination of these structures shows that the middle portion (residues 4 to 12) is similar and the structures only become different towards the N and C termini. This is due to slight unwinding of the backbone caused by different crystal packing. The

Table 1.4.3: R.m.s.d. values between main chain backbone atoms of the closed DSDH conformation of gramicidin dimers from *in meso* and organic solvent structures

Dimer / crystal structure		CF3		CF2		CF1			1AL4	1ALZ	1ALX
		AB	CD	AB	CD	AB	CD	EF			
CF3	AB	-	-	-	-	-	-	-	-	-	-
	CD	1.2	-	-	-	-	-	-	-	-	-
CF2	AB	1.0	-	-	-	-	-	-	-	-	-
	CD	1.5	1.8	1.5	-	-	-	-	-	-	-
CF1	AB	0.5	1.2	0.8	1.3	-	-	-	-	-	-
	CD	1.2	0.5	1.1	1.9	1.2	-	-	-	-	-
	EF	1.3	0.5	1.1	1.9	1.3	0.1	-	-	-	-
1AL4		0.5	1.2	1.0	1.4	0.5	1.3	1.3	-	-	-
1ALZ		0.5	1.2	0.9	1.4	0.5	1.3	1.3	0.3	-	-
1ALX		0.5	1.2	0.9	1.2	0.5	1.3	1.4	0.7	0.7	-

number of hydrogen bonds holding the backbones together ranges from 26 to 30 per dimer. The varying number of hydrogen bonds between dimers arises from altered hydrogen bonding in the N and C termini of each dimer. Most hydrogen bond donors or acceptors, not involved in holding backbones together, form crystal contacts or form hydrogen bonds with a ligand. In some cases however, hydrogen bond donor or acceptor atoms are not involved in any bonding. The main differences in these crystal structures (CF1-3) are found in their orientation of side chains particularly the tryptophan side chains.

1.4.4.2 Crystal Packing

All three crystal forms are arranged in layers with the long axis of one gramicidin dimer oriented approximately normal to the layer plane (figure 1.4.5). This kind of packing is referred to as Type I packing (Michel, 1982) and is typically found for crystals grown *in meso* (Caffrey, 2009). In each crystal form the layer has a thickness of approximately 32 Å and one unit cell axis is running perpendicular to the layers. The unit cell consists of either one or two layers. The layers are held together by either 15 (CF1), 7 (CF2) or 8 (CF3) hydrogen bonds per asymmetric unit (between two connecting asymmetric units). The crystal contacts are formed by residues close to the N or C termini.

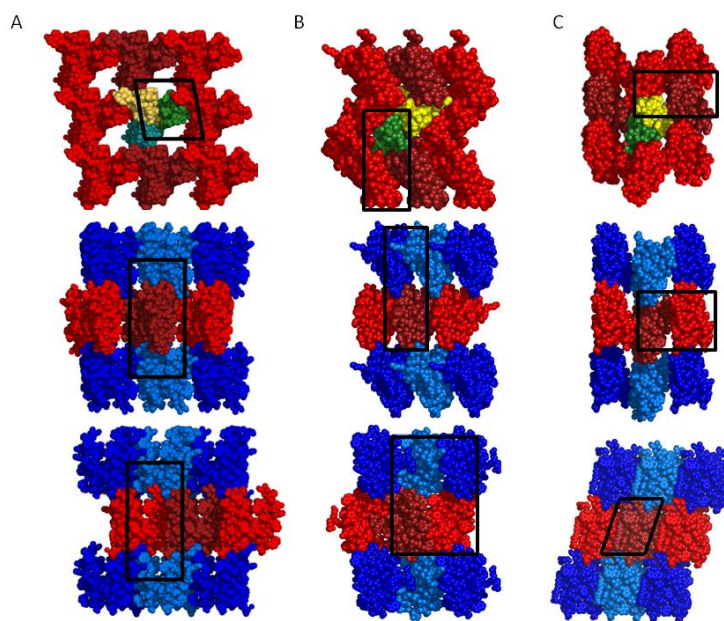


Figure 1.4.5: Layered (Type I) packing observed in crystals of gramicidin in different crystal forms (CF 1, A (Höfer et al., 2010; Höfer et al., 2011); CF2 (this chapter), B; CF3, C (this chapter (PDB ID 3ZQ8)) grown in the lipidic mesophases. Individual dimers are colored yellow, green, and cyan (top row). Alternate layers are colored red (light and dark) and blue (light and dark) to highlight Type I packing (middle and bottom rows). The unit cell is boxed in black. This figure has been adapted from Höfer et al., 2010; Höfer et al., 2011.

1.4.4.3 Crystallographic Water Molecules

The three *in meso* crystal structures of gramicidin (CF1, CF2, CF3) have only very few crystallographically visible water molecules. CF1 and CF2 have each two water molecules bound in the asymmetric unit and CF3 has one water molecule bound. It is not surprising to find such a low number of bound water molecules since the peptide is very hydrophobic and tends to exclude water molecules. Crystal structures of identical conformation (1AL4, 1ALX, 1ALZ) of gramicidin with crystals grown from organic solvents do not have water molecules modeled into electron density either. Another interesting aspect in the three *in meso* crystal forms is the location of the water molecules which are oriented close to the ends of individual crystal layers.

1.4.4.4 Solvent Content of *in meso* Grown Gramicidin Crystals

Crystal forms CF1, CF2, and CF3 have a solvent content of 52%, 38.8%, and 40.6%, respectively. The crystal forms originating from organic solvents range in their solvent content from 24% to 38.8% with an average of approximately 29% (figure 1.4.6; table 1.4.4). It appears that the solvent content from gramicidin crystals grown using organic solvents is lower than what has been observed for gramicidin crystals grown *in meso*. Investigation of linear peptides in the molecular weight range of gramicidin that were also crystallized from organic solvents revealed a solvent content even lower than for gramicidin crystals grown from organic solvents (table 1.4.4). A possible explanation could be that organic solvents are small volatile molecules that can easily escape the crystal lattice and allow tighter packing. Lipid and PEG are non-volatile and fairly big molecules that cannot be removed as easily.

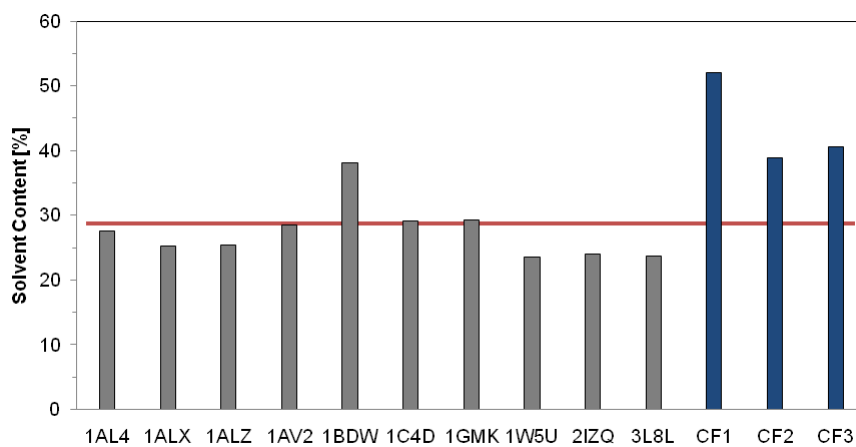


Figure 1.4.6: Solvent content (%) of gramicidin crystals grown from organic solvents (PDB IDs) (grey) and from *in meso* crystallization (CF1-3) (blue); the red line indicates the average solvent content in crystals grown from organic solvents. The information has been taken from the relevant PDB IDs.

Table 1.4.4: Crystal form, solvent content, Matthews coefficient (V_m), space group and unit cell dimensions of gramicidin crystals grown from various organic solvents

CF	PDB ID	Organic solvent, salt	Matthews coefficient (Solvent content (%))	Space group	a (Å)	b (Å)	c (Å)	α , β , γ , ($^\circ$)	Molprobitity clashscore (percentile (%))
A	1AL4	Propanol	1.69 (27.5)	P2 ₁ 2 ₁ 2 ₁	32.4	32.5	24.2	90	22.7 (2)
A	1ALZ	Ethanol	1.65 (25.4)	P2 ₁ 2 ₁ 2 ₁	31.6	32.4	24.2	90	21.3 (2)
B	1ALX	Methanol	1.64 (25.2)	P2 ₁	14.9	26	31.9	90	33.8 (0)
C	1AV2	Methanol, CaCl ₂	1.72 (28.5)	P2 ₁ 2 ₁ 2 ₁	31.1	31.9	52.1	90	20.7 (7)
C	1W5U	Ethanol, RbCl	1.61 (23.6)	P2 ₁ 2 ₁ 2 ₁	30.1	31.1	51.7	90	38.3 (0)
C	2IZQ	Methanol, KI	1.62 (24.0)	P2 ₁ 2 ₁ 2 ₁	30.1	31.3	51.6	90	24.9 (1)
C	3L8L	Methanol, NaI	1.61 (23.7)	P2 ₁ 2 ₁ 2 ₁	29.9	31.3	51.7	90	50.2 (0)
D	1BDW	Acetic acid	1.99 (38.1)	P2 ₁ 2 ₁ 2 ₁	20.6	27.9	52	90	25.3 (8)
E	1C4D	Methanol, CsCl	1.73 (29.1)	P2 ₁ 2 ₁ 2 ₁	32.1	52.1	31.2	90	428.8 (0)
E	1GMK	Methanol, KSCN	1.74 (29.2)	P2 ₁ 2 ₁ 2 ₁	32.1	51.8	31.4	90	93.6 (0)

1.5 Discussion

1.5.1 Crystallization of Membrane Peptides from Lipidic Cubic Phases (LCP) of Different Lipids

Gramicidin has been crystallized from a number of different LCP-forming lipids (7.7, 8.8, 9.9, 11.7 MAG). In CF1 and CF2 the major precipitant was PEG. Under these conditions the LCP had converted to a sponge phase. Grabe et al., 2003 suggested that crystallization of membrane proteins with fewer than 4 trans-membranal helices from the lipidic cubic Pn3m phase is not possible, due to their inability to flatten the curvature of the lipidic cubic phase enough to continue successful nucleation and crystal growth (Grabe et al., 2003). The theoretical calculations were based on a pure cubic phase system as described in the original finding for bacteriorhodopsin. In the case of the gramicidin crystals, the precipitant PEG triggers a swelling of the cubic phase and transformation to the so-called sponge phase (Cherezov et al., 2006; Engstroem, 1998). The latter has enlarged aqueous channels and an irregularly packed, less curved bilayer that has different physico-chemical parameters and therefore tends to better facilitate crystallization than pure cubic phase. The dynamics of the lipid in an *in meso* system was not fully covered by the theoretical study mentioned above.

Furthermore, these results encourage the use of the *in meso* method for the wide range of membrane peptides whose structures could previously only be explored by crystallization from organic solvents or by NMR. Each method has its own shortcomings; organic solvents may affect the conformation and NMR may require ^{13}C and ^{15}N labeled species, which can be costly to produce.

1.5.2 Crystallization Mechanism

1.5.2.1 Three Hosting Lipids, One Crystal Form in Case of CF1

An important finding of this work is that mesophases formed from three distinctly different MAGs go on to create identical crystals of gramicidin by the *in meso* method. Thus, at the level of crystal growth rate and final structure there was no

measurable difference between the three hosting lipids. What made this study possible in the first place was the prior knowledge that all three lipids form the cubic phase when fully hydrated at or close to 20 °C. However, the bilayer thickness of the cubic phase is different in each, with 7.7 and 9.9 MAG having the thinnest and thickest bilayers, respectively. The expectation was that this significant difference in microstructure would influence the state of the gramicidin in the mesophase prior to crystallogenesis and, in turn, reflect itself as a different form in the resulting crystal. In this way then it was hoped to access crystallographically the HSH form of gramicidin which, to date, has only been characterized structurally by NMR.

The finding that all three lipids produced identical crystal structures was unexpected. To what then can we ascribe this result? One possibility is that the conformation of the peptide is set by the organic solvent (TFE) and the protocol used to prepare samples for reconstitution (scenario 1, table 1.5.1). Of note however is the controversial nature of any such solvent effect (Killian et al., 1988; Sawyer et al., 1990). Presumably then if it is the DSDH form that gets reconstituted into the mesophase bilayer initially it does not, under current *in meso* crystallization conditions, equilibrate to other, perhaps more energetically preferred forms. As such, it is trapped in the mesophase as a DSDH which subsequently shows up in the final crystal.

Another explanation for how the same crystalline gramicidin conformation emerges is that the peptide actually reconstitutes into the cubic phase bilayer in the same DSDH conformation despite the large differences in mesophase microstructure (scenario 2, table 1.5.1). The DSDH form goes on then to nucleate during *in meso* crystallogenesis giving rise to identical final crystal structures in a manner that is independent of the hosting lipid. This is perhaps surprising given the large difference (6.5 Å comparing 7.7 MAG and 9.9 MAG) in the bilayer thickness of the parent cubic phases which might be expected to influence the pitch of the DSDH, if not the overall conformational state of the peptide. But such conformational differences, if they exist in the starting mesophase, do not show up in the final crystal. This suggests that the DSDH form crystallized is a particularly stable conformation.

It is also possible that the three different MAGs used in this study generate distinct mesophases that support different gramicidin conformations as a prelude to crystallogenesis. In this scenario (scenario 3 and 4, table 1.5.1) however, the DSDH form is the most stable conformation. Thus, during the processes that give rise to nucleation and crystal growth, those other forms convert uniformly to the DSDH conformation, and it is this that appears in the final crystal. As a prelude to crystal formation the extant model includes an intermediate state where the proteins are locally concentrated and partially ordered in a lamellar domain (Caffrey, 2008). It is possible that the conversion to the DSDH happens in this more fluid environment (scenario 5, table 1.5.1).

This discussion highlights the need for follow up studies to establish the effect that different solvents and dispersants, and indeed reconstitution protocols, have on the conformation that exists when gramicidin is incorporated into the lipidic cubic phase of MAGs. This should then be related to the form that ultimately crystallizes by the *in meso* method. Assorted spectroscopic, scattering, and chromatographic methods might be taken advantage of here.

Table 1.5.1: Possible scenarios for the way in which gramicidin crystallizes in mesophases formed with different MAGs. This table has been taken from Höfer *et al.*, 2011)

Steps in <i>in meso</i> Crystallization	Scenario / conformation ^a				
	1	2	3	4	5
Solubilization/Reconstitution	DSDH	MIXED	MIXED	MIXED	MIXED
In mesophase	DSDH	DSDH	MIXED	MIXED	MIXED
In mesophase + precipitant	DSDH	DSDH	DSDH	MIXED	MIXED
Phase separation	DSDH	DSDH	DSDH	DSDH	MIXED
Final crystal	DSDH	DSDH	DSDH	DSDH	DSDH

^aDSDH refers to the specific double stranded intertwined double helix conformation of gramicidin that appears in final crystals. MIXED indicates that gramicidin can adopt different conformations.

1.5.2.2 A Role for PEG

In addition to having the same structure in crystals grown from mesophases created using different lipids, the composition of the gramicidin crystal was almost identical in each case. Of particular note is the PEG component which was mapped to four regions of electron density. Two of these, represented by PEG-B and PEG-B₀, are located at the end of the DSDH where the planar sheets of gramicidin come together in the Type I packing arrangement. This corresponds to the interfacial region of the lipid bilayer in the mesophase from which crystals grew and, as a consequence, are not unexpected. However, the third and longest PEG molecule, PEG-A, resides in an elongated state with the bulk of its mass at the center of and in the plane containing the DSDH. It passes in between gramicidin dimers and hydrogen bonds with the N_ε of tryptophan side chains extending from the long axis of the DSDH dimers. To understand how this long stretch of polymer might find its way to this location in the crystal lattice, we must first consider the means by which crystals form by the *in meso* method.

As noted, a proposal has been advanced for how *in meso* crystallogenesis takes place at a molecular level (Caffrey, 2008; Caffrey, 2009). It begins typically with an isolated biological membrane which is treated with detergent to solubilize the target protein. The protein-detergent complex is purified by standard wet biochemical methods that usually involve a number of chromatographic steps. Homogenizing with a MAG affects reconstitution of the purified protein into the bilayer of the bicontinuous cubic phase. The protein retains its native conformation and activity and is free to move within the plane of the cubic phase bilayer. A precipitant is added to the mesophase which triggers a phase separation. Under conditions leading to crystallization, one of the separated phases is lamellar and becomes enriched in protein. The locally high concentration of protein (that may or may not include detergent and native membrane lipid), in conjunction with an appropriate bathing solution composition and bilayer microstructure, act to facilitate nucleation and crystal growth. The growing crystal is fed by proteins moving through a lamellar portal to the face of the crystal from the bulk mesophase which acts as a protein

reservoir. Experimental evidence exists in support of aspects of this hypothesis (Caffrey, 2008; Caffrey, 2009; Cherezov and Caffrey, 2007).

In the case of gramicidin, there is no solubilizing detergent, and given the purity of the peptide as purchased, it is unlikely to harbor any bound lipid. Thus, combining the dry lipid/gramicidin mix with water at 20 °C spontaneously generates the cubic phase with the peptide reconstituted in its bilayer (Liu and Caffrey, 2005). What probably happens then as soon as the precipitant makes contact with the gramicidin-laden mesophase is that it experiences an osmotic shock (Caffrey, 2008). This is so because the mesophase itself is porous and permeated by water-filled channels, while the precipitant is a concentrated solution of high molecular weight PEG. The osmotic shock itself may trigger a local and transient dehydration of the mesophase bolus which, if severe enough, could stabilize the lamellar phase (Qiu and Caffrey, 2000). With time, the precipitant components make their way into the recesses of the mesophase bolus and the osmotic shock dissipates. As the concentration of PEG in the mesophase rises it induces a cubic-to-sponge phase transition. The sponge phase is characterized by enlarged aqueous channels and a less ordered mesophase characterized by diffuse X-ray scattering (Cherezov et al., 2006). However, the sponge phase retains the character of being bicontinuous and thus, can support crystal growth by the mechanism outlined above for the *in meso* method. The proposal that the sponge phase forms during crystallogenesis is supported by the observation that, with time in contact with the precipitant, the mesophase adopted a more fluid character reminiscent of the sponge phase. In fact, it is from the fluid sponge phase that crystals were harvested.

PEG is a water-soluble, linear polymer composed of repeating ethylene glycol (-CH₂-CH₂-O-) units. Presumably, it enters the mesophase by way of its aqueous channels which are continuous with the bulk medium, in this case, the precipitant solution. The channels, which start out quite narrow—just tens of angströms wide—are surrounded by a highly curved bilayer of MAG. Thus, the wall of the channel is coated with polar lipid head groups made up of glycerols and ester moieties. With distance from the core of the aqueous channel, the polar interface gives way to the apolar hydrocarbon

of the bilayer created by the acyl chains of the MAG. In this environment, the PEG polymers reptate through the channels to the core of the mesophase bolus in a process driven by diffusion and mass action. Within the confines of the mesophase, the PEG molecules are expected to partition preferentially at the interface that separates the aqueous core of the channel and the apolar interior of the lipid bilayer. There, the apolar methylenes of the ethylene glycol units orient toward the hydrophobic interior of the bilayer. Its ether oxygen points toward the core of the aqueous channel where it can hydrogen bond with glycerol hydroxyls and water at the interface. In so doing, PEG has the effect of expanding the area per lipid molecule at the interfacial region, thereby lessening curvature and stabilizing the more swollen sponge phase variant of the mesophase. The current crystallization study was carried out at a MAG / gramicidin mole ratio of 20:1. Assuming that the peptide and lipid are uniformly distributed within the bilayer of the mesophase upon reconstitution then, on average, each gramicidin DSDH dimer will be surrounded by 1-2 annuli of lipid molecules. As PEG polymers contact and enter the mesophase, they will very quickly encounter and begin to interact with gramicidin. The interactions between gramicidin and PEG, revealed in the *in meso* crystal structure (figure 1.4.1), likely reflect the kind of interactions that occur in the mesophase prior to crystallization. Thus, the ether oxygens of PEG at the aqueous/ bilayer interface establish hydrogen bond contact with the Ne of tryptophan side chains extending from the DSDH dimers at that same interface. (This scenario assumes that gramicidin has already adopted the DSDH conformation. We cannot however discount the possibility that PEG triggers a transition to the double-stranded form.) In the DSDH, eight tryptophans from the two gramicidin monomers are arranged relatively uniformly along the long axis of the dimer. Tryptophan 15 is closest to the interface and likely is the first tryptophan to be encountered by PEG as it moves into the aqueous channels of the mesophase. Because of their close proximity to one another in the PEG chain, it is possible that adjacent ethylene glycol units bring DSDH dimers together. In so doing, the local concentration of gramicidin rises further, facilitating nucleation and crystal growth.

Tryptophan 13, 11, and 9 are arrayed on the DSDH dimer along and around its long axis in the direction away from the interface deeper into the hydrocarbon core of the bilayer. Thus, PEG has the opportunity to move from tryptophan 15 at the interface and to penetrate more deeply into the apolar recesses of the bilayer where it can interact with other tryptophans. Presumably, tryptophan 13 is the first candidate donor followed, in turn, by tryptophan 11 and 9. In fact, the polymer may choose to traverse the entire bilayer hopping from tryptophan 15 on one monomer in one leaflet of the bilayer all the way across via the intervening six tryptophans to tryptophan 15 on another monomer in the adjacent leaflet of the bilayer.

Hydrogen bonding to the N ϵ of buried tryptophans would enable the relatively polar ether oxygen of PEG to partition into the apolar interior of the bilayer. At the same time, the methylenes of individual ethylene glycol units would likely orient in such a way as to interact favorably with similar moieties of the lipid acyl chains which abound in the bilayer interior. In this way then it is possible to explain how PEG, as a water-soluble polymer, can make its way into the low dielectric interior of the mesophase bilayer facilitated by the presence of the transmembrane gramicidin dimer. The fact that PEG-A is found in the crystal with 60 % of its 16 modeled ethylene glycol units at the midplane of the layered gramicidin molecules undoubtedly reflects a set of polymer-peptide interactions that are particularly favorable energetically.

In a separate study, we demonstrated that high concentrations of gramicidin in the cubic phase of 9.9 MAG induced formation of the inverted hexagonal (H_{II}) phase (Li et al., 2011). This was explained by a hydrophobic mismatch between the presumed HSSH form of the peptide *in meso* and the bilayer in which it resided. It was speculated further that locally high concentrations of the peptide that must exist prior to nucleation and crystal growth by the *in meso* method would work against crystal growth by favoring H_{II} phase formation. This led to the suggestion that additives and/or hosting lipids which stabilize a flatter, less curved aqueous/apolar interface would facilitate crystallization. This is entirely consistent with the results presented in the current study. PEG, the common precipitant ingredient in all three hosting lipids

examined, induced a sponge phase with these same characteristics and it is in its presence that structure-quality crystals of gramicidin formed. It is possible then by flattening out the bilayer interface of the cubic mesophase that PEG creates an environment that preferentially stabilizes the DSDH form of gramicidin. The latter has a more cylindrical shape in contrast to the HSHH conformation which is hourglass shaped. A flatter, more lamellar bilayer is likely to accommodate a cylindrical object with its long axis parallel to the bilayer normal more so than one that necks down toward its midsection. Thus, it may be that by stabilizing the sponge phase the PEG component of the precipitant favors conversion to the DSDH form which ends up in the final crystal (Scenario 4, table 1.5.1). Any difference in the original conformation that might exist upon reconstitution may well disappear upon equilibration of the hosting mesophase with a PEG-containing precipitant.

This reasoning suggests that, in addition to examining the conformational state of gramicidin in the lipidic cubic phase as a prelude to crystallogenesis (Liu and Caffrey, 2005), it is equally important to make such measurements after precipitant has been added. Thus, the gramicidin might exist in different states depending on the hosting lipid but then transition to another upon treatment with PEG (Scenario 3, table 1.5.1). The presence of PEG, and the sponge phase it induces, may selectively stabilize the DSDH form. This can only be established by following the conformational state, and indeed mesophase characteristics, from initial reconstitution into the lipidic cubic phase all the way to sponge phase formation on to nucleation and crystal growth.

1.5.2.3 An Alternative Crystallization Mechanism?

The above discussion is based on the premise that gramicidin crystallizes from the lipidic mesophase by the so-called *in meso* mechanism as already described. However, since gramicidin is soluble in pure PEG, we must be open to the possibility that gramicidin is crystallizing from the lipidic mesophase by some other mechanism that involves dissolution in PEG. The following observations are presented that argue against this possibility. The current *in meso* crystallization trials were performed

using water based precipitant solutions containing high molecular weight PEGs at 20 to 30 %(w/w). Separately, we have established that gramicidin has a very low solubility (micromolar solubility, table 1.3.1) in the PEG-containing precipitant solutions that produced diffraction-grade crystals. It is hard to visualize how such an apolar peptide as gramicidin would partition preferentially out of the bilayer of the mesophase into the bathing precipitant solution and to crystallize therein. Second, all crystals were observed to grow within the confines of the mesophase, not in the precipitant solution. Third, gramicidin crystals produced in this study exhibit Type I or layered packing which is entirely consistent with the *in meso* crystallization mechanism.

1.5.2.4 Crystallizing “Small Proteins” *In Meso* is Worthy of Consideration

Arguments have been advanced as to why the original cubic phase method would not work with proteins having fewer than five transmembrane helices (Grabe et al., 2003). The first has to do with the thermodynamic driving force which accrues from having proteins aggregate at regions of minimum curvature the center of the so-called monkey saddle—within the unit cell of the cubic phase. In so doing, the aggregate stabilizes and helps propagate a local lamellar phase where the hydrophobic mismatch between the apolar surface around the midsection of the protein and the bilayer interior is lessened. Calculations performed using the cubic-Pn3m phase formed by 9.9 MAG at 20 °C show that the elastic energy component of the chemical potential driving crystallization is dependent on membrane protein radius (in the plane of the bilayer) to the fourth power. Thus, proteins with a tight cluster of four or less idealized transmembrane R-helices will not release enough energy upon aggregating to compensate for the energy costs associated with nucleation and crystal growth. The second factor working against nucleation in this theoretical model is the energy barrier that must be crossed as proteins move between monkey saddle centers separated from one another by horse saddles where curvature is maximized. This

barrier serves to slow the overall process of protein aggregation and ultimately crystal growth.

In this and a related (Höfer et al., 2010) study, we show that gramicidin crystallizes *in meso* and that the crystals formed using three different hosting lipids, including 9.9 MAG, are of structure-quality. Gramicidin is not a protein; it is a peptide with 15 residues. Nonetheless, it has many of the attributes of a small protein and has been referred to as a mini-protein (Andersen et al., 2005). Under current experimental conditions, it is long enough, at 35 Å, to cross the membrane, not as an idealized α -helix, but as a DSDH with a radius in the plane of the membrane of ~ 18 Å. According to the model calculations of Grabe et al., 2003 referred to above this DSDH should not crystallize *in meso* because its cross-sectional area, and thus its perimeter, in the plane of the membrane is simply too small. The disparity between the theoretical analysis of Grabe et al., 2003 and our finding that gramicidin, as a miniprotein, does indeed crystallize *in meso* can be explained as follows. To begin with, the calculations are based on an idealized cubic-Pn3m phase formed by 9.9 MAG at 20 °C. Under these conditions, the bilayer of the cubic phase is indeed highly curved and the energetics described by Grabe et al., 2003 apply. However, the conditions under which gramicidin crystallized *in meso* included a precipitant with a high PEG concentration. As noted, PEG stabilizes the sponge phase which, like the cubic mesophase from which it forms, is bicontinuous. Accordingly, it can support crystallization in a manner entirely consistent with the proposed *in meso* mechanism but with energy barriers that are less than those expected for the simple cubic-Pn3m phase. In the sponge phase overall curvature is considerably reduced. Thus, the magnitude of the elastic energy term referred to above that counters the crystallization of smaller proteins is lessened. Additionally, the aqueous channels of the sponge phase are considerably enlarged compared to the default cubic-Pn3m phase, and the sponge phase itself is less ordered (as judged by small-angle X-ray scattering (Cherezov et al., 2006)) and more fluid. Thus, the horse saddle regions, which in the cubic-Pn3m phase have high curvature energy, are less curved. Accordingly, the energy barrier experienced by proteins moving throughout the

sponge phase is less. This, in turn, enables a more rapid onset of nucleation and crystallogenesis.

On thermodynamic and kinetic grounds therefore, crystallization of small proteins should be possible by the *in meso* method when the conditions are favorable. The screening process is designed to sample a wide variety of chemical and environmental conditions that adjust the character of the hosting mesophase and the target protein. When the proper conditions are found, crystals should form.

1.5.2.5 Functional Significance of the DSDH

Regardless of the lipid used to create the hosting mesophase for *in meso* crystallization of gramicidin, the DSDH was the only conformation observed in this study. Further, the pitch of the helix was sufficiently large that ions could not pass through the core of the dimer. Accordingly, it was the non-conducting or closed form of the dimer that was found.

One of the assumptions to explain this result is that the crystals responsible for these structures nucleated from lamellar domains of gramicidin within the mesophase and, by extension, represent growth from a membrane mimetic. This suggests therefore that the DSDH form of the peptide found in the crystal is a conformation that might exist in a biological membrane. Energetically, this is not expected to be a preferred conformation because several of the tryptophan indole side chains are buried in the bilayer interior. However, the energy cost of so positioning indoles is context dependent (White and von Heijne, 2008). In this case, the four tryptophans at alternating positions in the C-terminal half of the peptide have highly apolar leucine or valine residues as nearest neighbors. To some degree, these would facilitate indoles residing in an apolar environment. It is possible too that the DSDHs associate to form multimers. Through cation- π interactions (Gallivan and Dougherty, 1999) between tryptophans of adjacent dimers, stabilization of the intertwined form in a bilayer setting could be envisioned. The other alternative is that the DSDH conformation was induced to form and was stabilized by PEG in the precipitant solution, as described.

1.5.3 Do Tryptophan Interactions Stabilize Aggregates?

Gramicidin has been crystallized from the lipid matrix of monovaccein (11.7 MAG) using the *in meso* method. The conformation observed is the closed DSDH conformation as described by Burkhart, 1998 (1AL4, 1ALX, 1ALZ). This conformation is not associated with ion transport, but has been observed in membranes (Dzikovski et al., 2011; Kelkar and Chattopadhyay, 2007). CF3 crystallized in type I crystal arrangement with its typically stacked crystal layers. According to the proposed mechanism of *in meso* crystallization the protein concentrates in lamellar like areas of the lipidic cubic phase (Caffrey, 2011). In these regions the protein forms aggregates. If the aggregate is uniformly organized, it is a two-dimensional crystal in a membrane bilayer. During crystal growth the two-dimensional crystal grows and layers are stacked on top of each other. In this crystal structure (CF3) partial lipid molecules were modeled into electron density. The location of the lipids is approximately perpendicular to a crystal layer. The lipids headgroups are located at the interface of the crystal layer. Furthermore, the lipid headgroups form hydrogen bonds with the outer ring tryptophan side chains of gramicidin. In a membrane-like setting both lipid headgroups and outer ring tryptophans would be located at the membrane interface anchoring gramicidin in the membrane.

The most striking feature about this crystal structure is the hydrogen bonds that formed between the inner ring tryptophan residues of gramicidin. The spatially proximal tryptophan residues are aligned in a T shape arrangement, the most commonly observed interaction between tryptophan residues that are in close proximity (Chakrabarti and Bhattacharyya, 2007). In the DSDH conformation the inner ring tryptophan residues (Trp₉, Trp₁₁) are buried in the hydrophobic moiety of the lipid bilayer. This is energetically not as favorable as having them exposed in the bilayer interface as observed for the HSH conformation. By forming the observed hydrogen bonds the structure becomes energetically more favorable.

Aggregation in gramicidin has been observed for both conformations (Diociaiuti et al., 2002; Kelkar and Chattopadhyay, 2007). During the aggregation process

gramicidin remains in the membrane and dimers move closer together. This brings tryptophan residues in close proximity and allows them to interact through hydrogen bonding as observed in this structure. Aggregation therefore compensates for the unfavorable energetic situation. Atomic force microscopy suggests that the HSHH conformation aggregates in hexameric units that form a donut shaped object (Diociaiuti et al., 2002). So far, information about the shape or minimum number of dimers in DSDH conformation required to form an aggregate is not known. Crystallographic data from this work suggests that the minimal building block is a dimer of dimers as found in the asymmetric unit.

The setup of this experiment was such that the concentration of gramicidin to lipid was at 1:20. This causes a high concentration of gramicidin in the membrane. The likelihood of two dimers being in close proximity is very high. It is therefore possible to assume that through statistical fluctuations dimers come close in space and formation of aggregates is likely. The lipid used in this experiment was monovaccenin (11.7 MAG), which has 18 carbon atoms in the acyl chain and a membrane thickness of 32 Å like monoolein. This is approximately the length of a dimer in DSDH conformation. The HSHH and the DSDH conformation are always in equilibrium. This equilibrium is shifted towards the DSDH conformation with increasing bilayer thickness. This has been observed for bilayers in a lamellar phase, a predominantly flat bilayer. By comparison, lipidic cubic phases are highly curved bilayers and may alter the equilibrium between DSDH and HSHH. CD studies for gramicidin in lipidic cubic phase of 9.9 MAG, 9.7 MAG, and 11.9 MAG suggest that the peptide's predominant conformation is DSDH (Liu and Caffrey, 2005). It is therefore likely that gramicidin formed aggregates in DSDH conformation before it was exposed to precipitant solution. The precipitant then helped stabilize this particular arrangement of interactions. It is concluded that the crystal form described here is physiologically relevant conformation and not artificial in the crystallization. This is because this conformation has been observed in bilayers and the tryptophan-tryptophan interactions are not unique to this structure. They have been observed in a number of protein structures.

1.5.4 Similar Origin of all *in meso* Gramicidin Crystal Forms

Comparing all three crystal forms described in this paper, it becomes apparent, that they all show the layered type I crystal packing that is typically found for *in meso* crystals. The second common feature is the arrangement of two dimers in the shape of a V. Since this arrangement is similar in all three crystal forms it could be concluded that the formation of this V shaped arrangement preceded crystallization and is the stabilizing arrangement in an aggregate of gramicidin molecules. The aggregate is then stabilized by the addition of a precipitant solution. The major precipitant PEG as found for CF1 and CF2 may wiggle into this stable aggregate and form hydrogen bonds between the side chain of tryptophan residues and the oxygens in PEG.

The tighter packing observed in organic solvent crystals may be explained by the much smaller and more volatile nature of the organic solvent molecules. These can escape from the crystal more easily than molecules like lipids and PEG that are bulkier and are not volatile. It does not explain the V shaped packing observed in all crystal forms originating from *in meso* crystals. It can be speculated that the increased solvent content for *in meso* crystals compared to organic solvent crystals is caused by an already V shaped arrangement in the aggregate in the bilayer. It may be the most energetically stable formation in which the inner ring tryptophan residues are all satisfied through hydrogen bonding.

1.6 Conclusions and Outlook

Gramicidin D has been used here as a test peptide to show that standard *in meso* crystallization method can be used to produce diffraction quality crystals of hydrophobic membrane peptides. It was possible to obtain crystals of gramicidin from the lipidic cubic phase of 9.9 MAG (monoolein) that diffracted to 1.7 Å. The peptide was found in the DSDH conformation. PEG was a main precipitant and was identified in the crystal structure forming weak hydrogen bonds with predominantly tryptophan residues. In the past the crystallization of highly hydrophobic peptides has only been successful from organic solvents. They are hydrophobic in nature, but do not represent the natural lipid bilayer. Therefore, gramicidin structures solved using organic solvents may represent a non-native conformation. The findings indicate that the DSDH conformation may be a relevant natural conformation of gramicidin and not one artificially induced by organic solvents. This procedure also demonstrates that lipidic cubic phase crystallization method is suitable for small, hydrophobic peptides and potentially small membrane proteins.

The crystallization of gramicidin was further explored. The shape and size of the two major gramicidin conformations differ in different environments. The HSHH conformation is wider and shorter, whereas the DSDH conformation is longer and thinner. In the DSDH conformation the tryptophan residues are distributed evenly along the molecule. In the HSHH conformation tryptophan residues are found as a ring of tryptophan residues around the top and the bottom of the dimer. It was assumed that reconstituting gramicidin into a lipidic mesophase with bilayers of different thickness would preferentially stabilize one form over the other and subsequently yield crystals of gramicidin in different conformations. Three different bilayers, formed by either 7.7 MAG, 8.8 MAG, or 9.9 MAG were tested. The lipid acyl chain lengths varied from 14 to 18 carbons. Surprisingly, gramicidin crystallized in all three bilayers in the DSDH conformation as previously found in 9.9 MAG. In all cases PEG was the major precipitant and formed weak hydrogen bonds with

tryptophan residues. The highest resolution was found for crystals that were grown in 7.7 MAG and diffracted to 1.14 Å.

Two more crystal forms of gramicidin were solved in the DSDH conformation. In one crystal form (CF2) PEG was again the major precipitant and formed a very similar network of hydrogen bonds between tryptophan residues and PEG oxygens as was previously seen in CF1. The third crystal form, CF3 is different in that gramicidin crystallized without PEG as major precipitant. Again, gramicidin was found in the DSDH conformation. Tryptophan side chains that formed hydrogen bonds with PEG in CF1 and CF2 now formed hydrogen bonds between nearby tryptophans. In this structure tryptophan side chains are oriented such that the aromatic moiety (indole ring) is perpendicular to the N_{ϵ} . The distance between the N_{ϵ} nitrogen and the indole ring is between 3.3 and 3.6 Å. These hydrogen bonds allow gramicidin to aggregate in bilayers if in the DSDH conformation. All crystal forms crystallized in type I crystal packing.

Here the non-ion-bound DSDH conformation was observed in a variety of different crystal forms (CF1 – CF3). Of interest to the scientific community is particularly the HSH conformation. Although there is evidence by CD analysis that gramicidin D exists in both conformations (HSH, DSDH) in the lipidic cubic phase, the crystallization process favors the DSDH conformation.

The two monomers that form the gramicidin dimer can be covalently linked. Literature (Bamberg and Janko, 1977; Stankovic et al., 1989; Urry et al., 1971) has shown that this linkage forces gramicidin to adopt the HSH conformation. By forcing the HSH conformation through chemical linkage, crystallization should be repeated. Crystals that may be found should be in the HSH conformation.

1.7 References

- Andersen, O.S., Koeppel, R.E., 2nd, and Roux, B. (2005). Gramicidin channels. *IEEE Trans Nanobioscience* 4, 10-20.
- Anderson, D.M., Gruner, S.M., and Leibler, S. (1988). Geometrical aspects of the frustration in the cubic phases of lyotropic liquid crystals. *Proc Natl Acad Sci U S A* 85, 5364-5368.
- Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G.W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogra D* 67, 271-281.
- Bouchard, M., and Auger, M. (1993). Solvent history dependence of gramicidin-lipid interactions: a Raman and infrared spectroscopic study. *Biophys J* 65, 2484-2492.
- Briggs, J., and Caffrey, M. (1994a). The temperature-composition phase diagram and mesophase structure characterization of monopentadecenoin in water. *Biophys J* 67, 1594-1602.
- Briggs, J., and Caffrey, M. (1994b). The temperature-composition phase diagram of monomyristolein in water: equilibrium and metastability aspects. *Biophys J* 66, 573-587.
- Burkhart, B.M., and Duax, W.L. (1999). Gramicidin channel controversy - reply. *Nat Struct Biol* 6, 611-612.
- Burkhart, B.M., Gassman, R.M., Langs, D.A., Pangborn, W.A., and Duax, W.L. (1998a). Heterodimer formation and crystal nucleation of gramicidin D. *Biophys J* 75, 2135-2146.
- Burkhart, B.M., Li, N., Langs, D.A., Pangborn, W.A., and Duax, W.L. (1998b). The conducting form of gramicidin A is a right-handed double-stranded double helix. *Proc Natl Acad Sci U S A* 95, 12950-12955.
- Burley, S.K., and Petsko, G.A. (1986). Amino-aromatic interactions in proteins. *FEBS Lett* 203, 139-143.
- Caffrey, M. (2003). Membrane protein crystallization. *J Struct Biol* 142, 108-132.
- Caffrey, M. (2008). On the Mechanism of Membrane Protein Crystallization in Lipidic Mesophases†. *Crystal Growth & Design* 8, 4244-4254.

Caffrey, M. (2009). Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu Rev Biophys* 38, 29-51.

Caffrey, M. (2011). Crystallizing membrane proteins for structure-function studies using lipidic mesophases. *Biochem Soc Trans* 39, 725-732.

Caffrey, M., and Cherezov, V. (2009). Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc* 4, 706-731.

Caffrey, M., and Porter, C. (2010). Crystallizing membrane proteins for structure determination using lipidic mesophases. *J Vis Exp*.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Rance, M.S., and Skelton, N.J. (2007). *Protein NMR Spectroscopy* San Diego: Academic Press.

CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-763.

Chakrabarti, P., and Bhattacharyya, R. (2007). Geometry of nonbonded interactions involving planar groups in proteins. *Prog Biophys Mol Biol* 95, 83-137.

Chattopadhyay, A., Rawat, S.S., Greathouse, D.V., Kelkar, D.A., and Koeppe, R.E., 2nd (2008). Role of tryptophan residues in gramicidin channel organization and function. *Biophys J* 95, 166-175.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66, 12-21.

Chen, Y., Tucker, A., and Wallace, B.A. (1996a).

Chen, Y., Tucker, A., and Wallace, B.A. (1996b). Solution structure of a parallel left-handed double-helical gramicidin-A determined by 2D ¹H NMR. *J Mol Biol* 264, 757-769.

Chen, Y., Wallace, B. A. (1997). Solvent Effects on the Conformation and Far UV CD Spectra of Gramicidin. *Biopolymers* 42, 771-781.

Cheng, A., Hummel, B., Qiu, H., and Caffrey, M. (1998). A simple mechanical mixer for small viscous lipid-containing samples. *Chem Phys Lipids* 95, 11-21.

Cherezov, V. (2011a). Lipidic cubic phase technologies for membrane protein structural studies. *Curr Opin Struct Biol* 21, 559-566.

Cherezov, V. (2011b). Lipidic cubic phase technologies for membrane protein structural studies. *Current Opinion in Structural Biology In Press, Corrected Proof*.

Cherezov, V., and Caffrey, M. (2007). Membrane protein crystallization in lipidic mesophases. A mechanism study using X-ray microdiffraction. *Faraday Discuss 136*, 195-212; discussion 213-129.

Cherezov, V., Clogston, J., Misquitta, Y., Abdel-Gawad, W., and Caffrey, M. (2002). Membrane protein crystallization in meso: lipid type-tailoring of the cubic phase. *Biophys J 83*, 3393-3407.

Cherezov, V., Clogston, J., Papiz, M.Z., and Caffrey, M. (2006). Room to move: crystallizing membrane proteins in swollen lipidic mesophases. *J Mol Biol 357*, 1605-1618.

Cherezov, V., Peddi, A., Muthusubramaniam, L., Zheng, Y.F., and 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, D.M., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.J., Kuhn, P., Weis, W.I., Kobilka, B.K., *et al.* (2007). High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science 318*, 1258-1265.

Chung, H., and Caffrey, M. (1994). The neutral area surface of the cubic mesophase: location and properties. *Biophys J 66*, 377-381.

Chung, H., and Caffrey, M. (1995). Polymorphism, mesomorphism, and metastability of monoelaidin in excess water. *Biophys J 69*, 1951-1963.

Clogston, J., Rathman, J., Tomasko, D., Walker, H., and Caffrey, M. (2000). Phase behavior of a monoacylglycerol: (myverol 18-99K)/water system. *Chem Phys Lipids 107*, 191-220.

Cross, T.A., Arseniev, A., Cornell, B.A., Davis, J.H., Killian, J.A., Koeppe, R.E., 2nd, Nicholson, L.K., Separovic, F., and Wallace, B.A. (1999). Gramicidin channel controversy--revisited. *Nat Struct Biol 6*, 610-611; discussion 611-612.

de Godoy, C.M.G., and Cukierman, S. (2011). Membrane phosphate headgroups' modulation of permeation of alkaline cations in gramicidin channels. *J Phys Chem B 115*, 5026-5031.

Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1985). Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3[ångström] resolution. *Nature* *318*, 618-624.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System. De-Lano Scientific, San Carlos, CA, USA <http://www.pymol.org>.

Diociaiuti, M., Bordi, F., Motta, A., Carosi, A., Molinari, A., Arancia, G., and Coluzza, C. (2002). Aggregation of gramicidin A in phospholipid Langmuir-Blodgett monolayers. *Biophys J* *82*, 3198-3206.

Doyle, D.A., Morais Cabral, J., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., and MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity. *Science* *280*, 69-77.

Doyle, D.A., and Wallace, B.A. (1997). Crystal structure of the gramicidin/potassium thiocyanate complex. *J Mol Biol* *266*, 963-977.

Dzikovski, B.G., Borbat, P.P., and Freed, J.H. (2011). Channel and nonchannel forms of spin-labeled gramicidin in membranes and their equilibria. *J Phys Chem B* *115*, 176-185.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* *60*, 2126-2132.

Engstroem, S., Alfons, K., Rasmusson, M., and Ljusberg-Wahren, H. (1998). Solvent-induced sponge (L₃) phases in the solvent-monoolein-water system. *Colloid Polymer Science* *108*, 93-98.

Fischetti, R.F., Xu, S., Yoder, D.W., Becker, M., Nagarajan, V., Sanishvili, R., Hilgart, M.C., Stepanov, S., Makarov, O., and Smith, J.L. (2009). Mini-beam collimator enables microcrystallography experiments on standard beamlines. *J Synchrotron Radiat* *16*, 217-225.

Fisher, R., and Blumenthal, T. (1982). An interaction between gramicidin and the sigma subunit of RNA polymerase. *Proc Natl Acad Sci U S A* *79*, 1045-1048.

Gallivan, J.P., and Dougherty, D.A. (1999). Cation-pi interactions in structural biology. *Proc Natl Acad Sci U S A* *96*, 9459-9464.

Glowka, M.L., Olczak, A., Bojarska, J., Szczesio, M., Duax, W.L., Burkhart, B.M., Pangborn, W.A., Langs, D.A., and Wawrzak, Z. (2005). Structure of gramicidin D-

RbCl complex at atomic resolution from low-temperature synchrotron data: interactions of double-stranded gramicidin channel contents and cations with channel wall. *Acta Crystallogr D Biol Crystallogr* *61*, 433-441.

Grabe, M., Neu, J., Oster, G., and Nollert, P. (2003). Protein interactions and membrane geometry. *Biophys J* *84*, 854-868.

Grass, G., Otto, M., Fricke, B., Haney, C.J., Rensing, C., Nies, D.H., and Munkelt, D. (2005). FieF (YiiP) from *Escherichia coli* mediates decreased cellular accumulation of iron and relieves iron stress. *Arch Microbiol* *183*, 9-18.

Greathouse, D.V., Hinton, J.F., Kim, K.S., and Koeppe, R.E., 2nd (1994). Gramicidin A/short-chain phospholipid dispersions: chain length dependence of gramicidin conformation and lipid organization. *Biochemistry* *33*, 4291-4299.

Gu, H., Lum, K., Kim, J.H., Greathouse, D.V., Andersen, O.S., and Koeppe, R.E. (2011). The membrane interface dictates different anchor roles for "inner pair" and "outer pair" tryptophan indole rings in gramicidin channels. *Biochemistry* *50*, 4855-4866.

Henderson, R., and Unwin, P.N.T. (1975). Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* *257*, 28-32.

Höfer, N., Aragao, D., and Caffrey, M. (2010). Crystallizing transmembrane peptides in lipidic mesophases. *Biophys J* *99*, L23-25.

Höfer, N., Aragao, D., Lyons, J.A., and Caffrey, M. (2011). Membrane Protein Crystallization in Lipidic Mesophases. Hosting Lipid Effects on the Crystallization and Structure of a Transmembrane Peptide. *Crystal Growth & Design* *11*, 1182.

Hotchkiss, R., and Dubos, R. (1940). Fractionation of the bactericidal agent from cultures of a soil bacillus. *J Biol Chem* *132*, 791-792.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* *47 (Pt 2)*, 110-119.

Jones, T.L., Fu, R., Nielson, F., Cross, T.A., and Busath, D.D. (2010). Gramicidin channels are internally gated. *Biophys J* *98*, 1486-1493.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637.

Kantardjieff, K.A., and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* 12, 1865-1871.

Kelkar, D.A., and Chattopadhyay, A. (2007). Modulation of gramicidin channel conformation and organization by hydrophobic mismatch in saturated phosphatidylcholine bilayers. *Biochim Biophys Acta* 1768, 1103-1113.

Kessler, N., Schuhmann, H., Morneweg, S., Linne, U., and Marahiel, M.A. (2004). The linear pentadecapeptide gramicidin is assembled by four multimodular nonribosomal peptide synthetases that comprise 16 modules with 56 catalytic domains. *J Biol Chem* 279, 7413-7419.

Ketchum, R.R., Hu, W., and Cross, T.A. (1993). High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR. *Science* 261, 1457-1460.

Killian, J.A., Prasad, K.U., Hains, D., and Urry, D.W. (1988). The membrane as an environment of minimal interconversion. A circular dichroism study on the solvent dependence of the conformational behavior of gramicidin in diacylphosphatidylcholine model membranes. *Biochemistry* 27, 4848-4855.

Killian, J.A., and von Heijne, G. (2000). How proteins adapt to a membrane-water interface. *Trends Biochem Sci* 25, 429-434.

Landau, E.M., and Rosenbusch, J.P. (1996). Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci U S A* 93, 14532-14535.

Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography No. 26*.

Li, D., Lee, J., and Caffrey, M. (2011). Crystallizing Membrane Proteins in Lipidic Mesophases. A Host Lipid Screen. *Cryst Growth Des* 11, 530-537.

Li, Y., Andersen, O.S., and Roux, B. (2010). Energetics of double-ion occupancy in the gramicidin A channel. *J Phys Chem B* 114, 13881-13888.

Lin, T.H., Huang, H., Wei, H. A., Shiao, S. H., Chen, Y. C. (2005). The Effect of Temperature and Lipid on the Conformational Transition of Gramicidin A in Lipid Vesicles. *Biopolymers* 78, 179-186.

Liu, W., and Caffrey, M. (2005). Gramicidin structure and disposition in highly curved membranes. *J Struct Biol* 150, 23-40.

Liu, W., and Caffrey, M. (2006). Interactions of tryptophan, tryptophan peptides, and tryptophan alkyl esters at curved membrane interfaces. *Biochemistry* 45, 11713-11726.

LoGrasso, P.V., Moll, F., 3rd, and Cross, T.A. (1988). Solvent history dependence of gramicidin A conformations in hydrated lipid bilayers. *Biophys J* 54, 259-267.

Lomize, A.L., Orekhov, V., and Arsen'ev, A.S. (1992). Refinement of the spatial structure of the gramicidin A ion channel. *Bioorg Khim* 18, 182-200.

Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* 50, 437-450.

Lu, M., and Fu, D. (2007). Structure of the zinc transporter YiiP. *Science* 317, 1746-1748.

Lundbaek, J.A., Koeppe, R.E., 2nd, and Andersen, O.S. (2010). Amphiphile regulation of ion channel function by changes in the bilayer spring constant. *Proc Natl Acad Sci U S A* 107, 15427-15430.

Mandl, J., and Paulus, H. (1985). Effect of linear gramicidin on sporulation and intracellular ATP pools of *Bacillus brevis*. *Arch Microbiol* 143, 248-252.

Marahiel, M.A., Nakano, M.M., and Zuber, P. (1993). Regulation of peptide antibiotic production in *Bacillus*. *Mol Microbiol* 7, 631-636.

Markham, J.C., Gowen, J.A., Cross, T.A., and Busath, D.D. (2001). Comparison of gramicidin A and gramicidin M channel conductance dispersities. *Biochim Biophys Acta* 1513, 185-192.

Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491-497.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658-674.

MerckIndex (1996). Merck Index, pp. 4553.

Michel, H. (1982). Characterization and crystal packing of three-dimensional bacteriorhodopsin crystals. *EMBO J* *1*, 1267-1271.

Misquitta, L.V., Misquitta, Y., Cherezov, V., Slattery, O., Mohan, J.M., Hart, D., Zhalnina, M., Cramer, W.A., and Caffrey, M. (2004a). Membrane protein crystallization in lipidic mesophases with tailored bilayers. *Structure* *12*, 2113-2124.

Misquitta, Y., and 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, J.M., Wells, A.J., Hart, D.J., and Caffrey, M. (2004b). Rational design of lipid for membrane protein crystallization. *J Struct Biol* *148*, 169-175.

Misra, N., Martinez, J.A., Huang, S.C., Wang, Y., Stroeve, P., Grigoropoulos, C.P., and Noy, A. (2009). Bioelectronic silicon nanowire devices using functional membrane proteins. *Proc Natl Acad Sci U S A* *106*, 13780-13784.

Mo, Y., Cross, T.A., and Nerdal, W. (2004). Structural restraints and heterogeneous orientation of the gramicidin A channel closed state in lipid bilayers. *Biophys J* *86*, 2837-2845.

Modest, B., Marahiel, M.A., Pschorn, W., and Ristow, H. (1984). Peptide antibiotics and sporulation: induction of sporulation in asporogenous and peptide-negative mutants of *Bacillus brevis*. *J Gen Microbiol* *130*, 747-755.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* *53*, 240-255.

Nies, D.H., Nies, A., Chu, L., and Silver, S. (1989). Expression and nucleotide sequence of a plasmid-determined divalent cation efflux system from *Alcaligenes eutrophus*. *Proc Natl Acad Sci U S A* *86*, 7351-7355.

Nugent, T., and Jones, D.T. (2009). Transmembrane protein topology prediction using support vector machines. *BMC Bioinformatics* *10*, 159.

Olczak, A., Glowka, M.L., Szczesio, M., Bojarska, J., Duax, W.L., Burkhart, B.M., and Wawrzak, Z. (2007). Nonstoichiometric complex of gramicidin D with KI at 0.80 Å resolution. *Acta Crystallogr D Biol Crystallogr* *63*, 319-327.

Olczak, A., Glowka, M.L., Szczesio, M., Bojarska, J., Wawrzak, Z., and Duax, W.L. (2010). The first crystal structure of a gramicidin complex with sodium: high-resolution study of a nonstoichiometric gramicidin D-NaI complex. *Acta Crystallogr D Biol Crystallogr* 66, 874-880.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* 276, 307-326.

Paulus, H., Sarkar, N., Mukherjee, P.K., Langley, D., Ivanov, V.T., Shepel, E.N., and Veatch, W. (1979). Comparison of the effect of linear gramicidin analogues on bacterial sporulation, membrane permeability, and ribonucleic acid polymerase. *Biochemistry* 18, 4532-4536.

Pebay-Peyroula, E., Rummel, G., Rosenbusch, J.P., and Landau, E.M. (1997). X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science* 277, 1676-1681.

Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001). ARP/wARP and molecular replacement. *Acta Crystallogr D Biol Crystallogr* 57, 1445-1450.

Petsko, G.A., and Ringe, D. (2004). *Protein Structure and Function*. London: New Science Press.

Pope, C.G., Urban, B.W., and Haydon, D.A. (1982). The influence of n-alkanols and cholesterol on the duration and conductance of gramicidin single channels in monoolein bilayers. *Biochim Biophys Acta* 688, 279-283.

Qiu, H., and Caffrey, M. (1998). Lyotropic and Thermotropic Phase Behavior of Hydrated Monoacylglycerols: Structure Characterization of Monovaccenin†. *The Journal of Physical Chemistry B* 102, 4819-4829.

Qiu, H., and Caffrey, M. (2000). The phase diagram of the monoolein/water system: metastability and equilibrium aspects. *Biomaterials* 21, 223-234.

Ramachandran, G.N., and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv Protein Chem* 23, 283-438.

Raman, P., Cherezov, V., and Caffrey, M. (2006). The membrane protein data bank. *Cell Mol Life Sci* 63, 36-51.

Rasmussen, S.G., Devree, B.T., Zou, Y., Kruse, A.C., Chung, K.Y., Kobilka, T.S., Thian, F.S., Chae, P.S., Pardon, E., Calinski, D., *et al.* (2011). Crystal structure of the beta(2) adrenergic receptor-Gs protein complex. *Nature*.

Ristow, H., and Paulus, H. (1982). Induction of sporulation in *Bacillus brevis*. 1. Biochemical events and modulation of RNA synthesis during induction by tyrocidine. *Eur J Biochem* *129*, 395-401.

Roux, B., Allen, T., Berneche, S., and Im, W. (2004). Theoretical and computational models of biological ion channels. *Q Rev Biophys* *37*, 15-103.

Samanta, U., Pal, D., and Chakrabarti, P. (1999). Packing of aromatic rings against tryptophan residues in proteins. *Acta Crystallogr D Biol Crystallogr* *55*, 1421-1427.

Samanta, U., Pal, D., and Chakrabarti, P. (2000). Environment of tryptophan side chains in proteins. *Proteins* *38*, 288-300.

Sarges, R., and Witkop, B. (1965). Gramicidin. Vii. The Structure of Valine- and Isoleucine-Gramicidin B. *J Am Chem Soc* *87*, 2027-2030.

Sarkar, N., Langley, D., and Paulus, H. (1979). Studies on the mechanism and specificity of inhibition of ribonucleic acid polymerase by linear gramicidin. *Biochemistry* *18*, 4536-4541.

Sarkar N, P.H. (1972). Function of peptide antibiotics in sporulation. *Nat New Biol* *239*, 228-230.

Sawyer, D.B., Koeppe, R.E., 2nd, and Andersen, O.S. (1990). Gramicidin single-channel properties show no solvent-history dependence. *Biophys J* *57*, 515-523.

Schoenafinger, G., Schracke, N., Linne, U., and Marahiel, M.A. (2006). Formylation domain: an essential modifying enzyme for the nonribosomal biosynthesis of linear gramicidin. *J Am Chem Soc* *128*, 7406-7407.

Schracke, N., Linne, U., Mahlert, C., and Marahiel, M.A. (2005). Synthesis of linear gramicidin requires the cooperation of two independent reductases. *Biochemistry* *44*, 8507-8513.

Schulz, G.E. (2011). A new classification of membrane protein crystals. *J Mol Biol* *407*, 640-646.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* *31*, 3381-3385.

Sklenar, V., and Bax, A. (1987). Spin-echo water suppression for the generation of pure phase two-dimensional NMR spectra. *J Magn Reson* 74, 469.

Strieker, M., Tanovic, A., and Marahiel, M.A. (2010). Nonribosomal peptide synthetases: structures and dynamics. *Curr Opin Struct Biol* 20, 234-240.

Stroud, R.M. (2011). New tools in membrane protein determination. *F1000 Biol Rep* 3, 8.

Sun, H., Greathouse, D.V., Andersen, O.S., and Koeppe, R.E., 2nd (2008). The preference of tryptophan for membrane interfaces: insights from N-methylation of tryptophans in gramicidin channels. *J Biol Chem* 283, 22233-22243.

Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRY SOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* 28, 768.

Terwilliger, T.C. (2003). SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol* 374, 22-37.

Terwilliger, T.C., and Berendzen, J. (1999). Evaluation of macromolecular electron-density map quality using the correlation of local r.m.s. density. *Acta Crystallogr D Biol Crystallogr* 55, 1872-1877.

Townsley, L.E., Tucker, W.A., Sham, S., and Hinton, J.F. (2001). Structures of gramicidins A, B, and C incorporated into sodium dodecyl sulfate micelles. *Biochemistry* 40, 11676-11686.

Urry, D.W. (1971). The gramicidin A transmembrane channel: a proposed $\pi(L,D)$ helix. *Proc Natl Acad Sci U S A* 68, 672-676.

Urry, D.W., Goodall, M.C., Glickson, J.D., and Mayers, D.F. (1971). The gramicidin A transmembrane channel: characteristics of head-to-head dimerized (L,D) helices. *Proc Natl Acad Sci U S A* 68, 1907-1911.

Vaguine, A.A., Richelle, J., and Wodak, S.J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* 55, 191-205.

Wallace, B.A. (1983). Gramicidin A Adopts Distinctly Different Conformations in Membranes and in Organic Solvents. *Biopolymers* 22, 397-402.

Wallace, B.A., and Ravikumar, K. (1988). The gramicidin pore: crystal structure of a cesium complex. *Science* *241*, 182-187.

White, S.H., and von Heijne, G. (2008). How translocons select transmembrane helices. *Annu Rev Biophys* *37*, 23-42.

Yau, W.M., Wimley, W.C., Gawrisch, K., and White, S.H. (1998). The preference of tryptophan for membrane interfaces. *Biochemistry* *37*, 14713-14718.

Yu, C.H., and Pomes, R. (2003). Functional dynamics of ion channels: modulation of proton movement by conformational switches. *J Am Chem Soc* *125*, 13890-13894.

Chapter 2

Crystallization and Structure Determination of the Soluble Domain of C_{zr}B from *Thermus thermophilus*

The content of this chapter was published prior to this thesis in the following two publications:

Nicole Höfer, Olga Kolaj, Hui Li, Vadim Cherezov, Richard Gillilan, Gerard J. Wall and Martin Caffrey, Crystallization and preliminary X-ray diffraction analysis of a soluble domain of the putative zinc transporter CzcB from *Thermus thermophilus*, 2007, *Acta Crystallographica Section F*, F63, 673–677.

V. Cherezov, N. Höfer, D.M.E. Szebenyi, O. Kolaj, J.G. Wall, R. Gillilan, V. Srinivasan, C.P. Jaroniec, and M. Caffrey, Insights into the Mode of Action of a Putative Zinc Transporter CzcB in *Thermus thermophilus*, 2008, *Structure* 16, 1378–1388.

Text, figures and tables have been copied.

2.1 Introduction to CzcB

Zinc ions play an essential role in the catalysis of enzymes and they are important in the stabilization of protein structure (Alberts et al., 1998). They are involved in the catalysis of all 6 major enzyme groups. Zinc ions can help facilitate reactions in the following ways. In carboxypeptidase and alcohol dehydrogenase, the electrophilic character of zinc ions stabilizes the negative charges in the reaction intermediate. In catalytic sites of enzymes zinc ions are exposed to solvent molecules. In carbonic anhydrase catalytic zinc ions ionize bound solvent to nucleophilic hydroxyls. Zinc ions also stabilize protein structures as has been observed in zinc finger DNA binding domains (Alberts et al., 1998).

Zinc is most commonly ligated by histidine, glutamate, aspartate, and cysteine. In each case zinc ions prefer to be coordinated by four, five or six ligands (Alberts et al., 1998). The observed coordination geometry for four ligands is tetragonal, for five ligands square-based pyramidal or trigonal bipyramidal, and for six ligands is octahedral. The coordination geometries are predominantly determined by electrostatics. The distance between the zinc ion and the coordinating residue ranges from 2 to 2.35 Å.

Due to zinc ions' character as Lewis acid and electrophile, zinc ion concentration needs to be controlled in the cell. *Escherichia coli* is the most extensively studied organism concerning the homeostasis of zinc ions. It has a zinc quota (number of zinc atoms per cell) of just under a million, a value that is very tightly regulated (Outten and O'Halloran, 2001). The most remarkable fact is that essentially all of the cellular zinc exists in a bound state with less than a single zinc atom on average remaining free in the cytosol. It appears therefore that zinc ions are shuttled about in the cell in the bound form. In order to maintain homeostasis the import and export of zinc ions into the cell is tightly controlled by a number of protein families. The Cation Diffusion Facilitator family (CDF) is such a family. Members of this family have been found in all domains of life (Haney et al., 2005; Montanini et al., 2007). – YjiP (FieF) from *Escherichia coli* (*E. coli*) (Grass et al., 2005), CzcB in *Thermus thermophilus* (Spada et al., 2002), CzcD from *Ralstonia eutropha*-like

Figure 2.1.1: Sequence alignment of CzcB from *Thermus thermophilus* with CDF proteins from other organisms that are involved in zinc transport. Identical and homologous residues are highlighted in magenta and yellow, respectively. Red dots mark residues involved in direct binding to zinc in CzcB_{sf}. Secondary structure elements are indicated as α helical (bars) and β strand (arrows) using the zinc-CzcB_{sf} structure. Residues involved in apo-CzcB_{sf} dimerization are highlighted in green. Residues that make additional contacts subsequent to zinc binding are highlighted in blue. Numbers used to identify residues in the soluble fragment CzcB_{sf} that are described in the text are shown at the bottom of each block. For reference, residue number 1 in CzcB_{sf} corresponds to residue number 198 in the full-length protein and is marked with an open circle. The side chains of the conserved apolar residues (Ile15, Leu48, Leu70, Ile81) are all located in the hydrophobic core of CzcB_{sf}. The conserved Glu65 has its side chain on the exposed surface of the protomer. Transmembrane helices are labeled TM1–TM6 based on homology modeling using the known structure of YiiP (Lu and Fu, 2007). This figure has been taken from Cherezov et al., 2008.

CH34 (Munkelt et al., 2004; Nies et al., 1989), PA3963 in *Pseudomonas aeruginosa*, Zrc1 from yeast *Saccharomyces cerevisiae* (Conklin et al., 1992), AtZat from *Arabidopsis thaliana* (Bloss et al., 2002), ZnTX (X represents a number for order of discovery) from mammals (Haney et al., 2005). In Gram-negative bacteria, these proteins transport zinc and other ions like cobalt and manganese from the cytoplasm into the periplasmic space in a proton/cation antiporter mode (Haney et al., 2005).

Members of the CDF family share common features like dimerization, histidine-rich N and C termini, two domains – a cytosolic domain and a transmembrane domain. YiiP from *E. coli* has been solved in 2007 by Lu and Fu to a resolution of 3.6 Å and later to 2.9 Å (figure 2.1.2) (Lu and Fu, 2007; Lu and Fu 2009). The protein is a homodimer with each protomer having a molecular weight of 33 kDa. Each protomer has a transmembrane domain and a cytosolic domain. The transmembrane domain per protomer consists of 6 transmembrane helices. The hydrophilic C-terminal domain is located in the cytosol. The overall shape of the dimer resembles a Y shape (Lu et al.,

2009). The transmembrane domain consists of two portions per protomer; a four-helix-bundle formed by transmembrane helices 1, 2, 4, and 5 and a pair of transmembrane helices 3 and 6. Transmembrane helices 3 and 6 in each protomer form hydrophobic contacts with each other to stabilize the dimer. The transmembrane domain also shows a large cavity. A short loop connects transmembrane helix 6 with the cytosolic domain. The cytosolic domain adopts a metallochaperone-like fold. The majority of the dimer contacts originate from the two cytosolic domains.

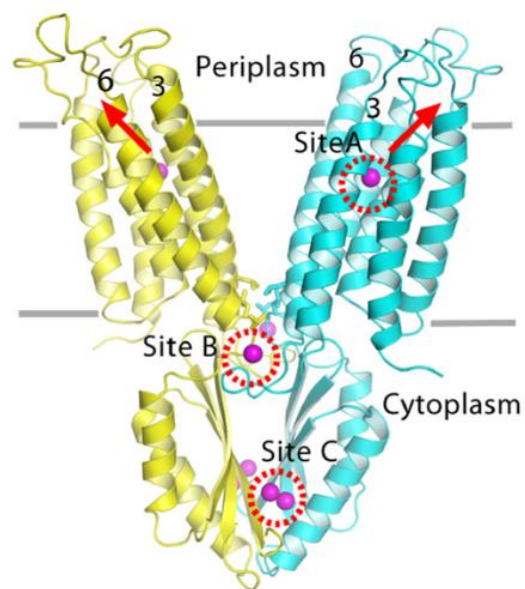


Figure 2.1.2: Ribbon representation of YiiP viewed from the membrane plane. The dimer is represented by the two protomers in yellow and cyan. Magenta spheres represent bound zinc ions. Zinc binding sites are marked in red circles. Grey lines indicate the possible membrane boundaries. Red arrows indicate the directions of Zn(II) exit from the two active-sites. Conserved hydrophobic residues involved in the TMD-TMD contacts are shown as sticks. This figure has been taken from figure 1 from Lu and Fu, 2009.

Each protomer binds 3 zinc ions, described in figure 2.1.2 as sites A, B, and C (Lu et al., 2009). Site A is considered the active site for zinc transport. This site is located in the center of the transmembrane domain. Zinc ions are stabilized in a tetragonal fashion by residues aspartate 45 and aspartate 49 from transmembrane helix 2, and histidine 153 and aspartate 157 from transmembrane helix 5. Site B and site C are both cytoplasmic binding sites. Site B is formed by a loop between transmembrane helices 2 and 3. Site C is formed between the cytosolic domains and is considered to actively stabilize the dimer interface. Zinc ions are coordinated in a binuclear fashion. Aspartate 285 bridges zinc ions 3.8 Å apart. Of interest is that the zinc ion coordinating residues show are constrained through hydrogen bonds with neighboring residues forming an extensive network of outer shell interactions that stabilize the dimer interaction.

A recent study by (Coudray et al, 2012) proposes a mechanism by which YjiP may function as an antiporter (figure 2.1.3). In this study, different structures were obtained using two different methods. The first structure came from traditional X-ray diffraction, carried out on crystals that were grown in the presence of high concentrations of zinc ions. The second structure came from first carrying out cryo-electron-microscopy, and using molecular dynamics. The crystal used for this second structure was grown with very little zinc present. It is presumed that the two conformations observed in this study were determined by the presence of zinc ions, or the lack thereof. The two methods yielded structures that varied significantly, particularly in the orientation of a bundle of four alpha helices. It appears that the tilting of this bundle relative to helices 3 and 6 may switch the access to the zinc transport site from one side of the membrane to the other. The cavity leading from the periplasmic side of the membrane to the zinc binding sites has previously been suggested as a pathway for binding and releasing zinc ions (Lu and Fu, 2007). This cavity to the periplasmic side collapses upon the tilting of the four-helix bundle, and a new cavity is formed leading to the cytoplasmic side of the membrane. The pivoting of the four-helix bundle appears to occur in each monomer, rather than in only one monomer per dimer. This switching mechanism has been called the “Alternating Access Mechanism”. The energy required to induce the conformational change

described is proposed (Klinenberg, 2005) to come from the binding energy of the substrate.

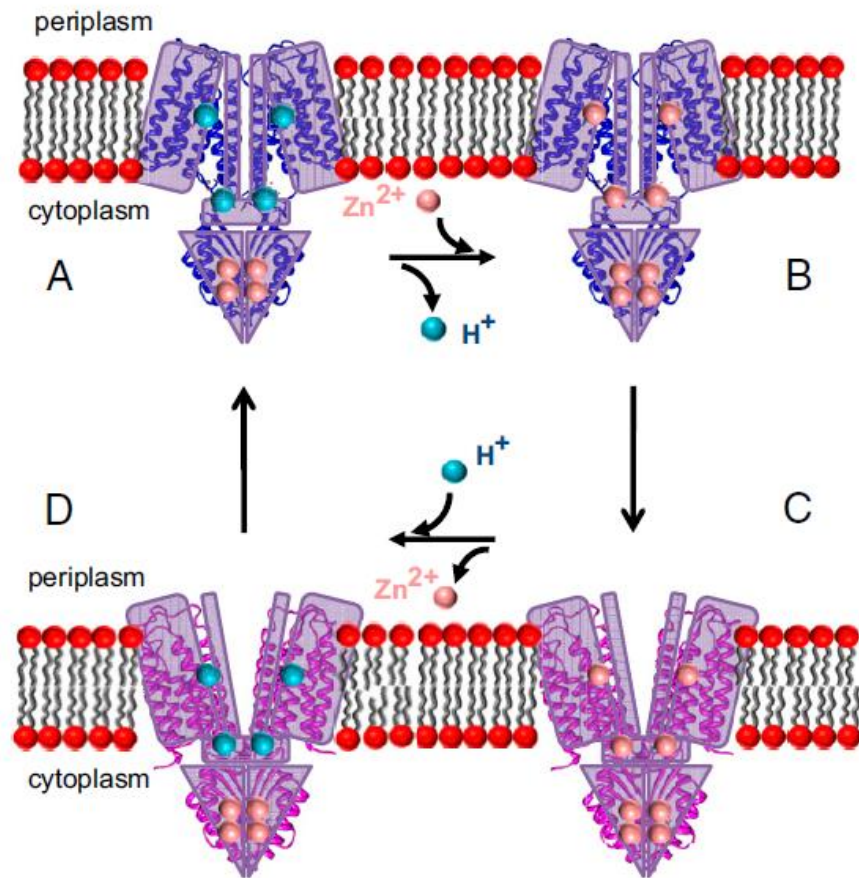


Figure 2.1.3: Alternating access mechanism for $\text{Zn}^{2+}/\text{H}^{+}$ antiport by YiiP. This mechanism involves inward-facing (A and B) and outward-facing (C and D) conformations, both of which are able to bind Zn^{2+} or H^{+} . The proton motive force provides a driving force for exporting Zn^{2+} from the cytoplasm. The conformation in A corresponds to that observed in the helical tubes, whereas the conformation in C corresponds to the X-ray structure solved in the presence of Zn^{2+} . This figure has been taken from figure 4 from Coudray et al., 2012.

The model of this alternating access mechanism may be applicable to more members of the CDF family. A similar mechanism has been suggested by Russel and Soulimane (Russell and Soulimane, 2012) to explain cation transport through MmCDF3, a CDF member that lacks a cytoplasmic domain. In this case it is specified that the initial binding sites of zinc must be the N-termini of the protein. It is proposed that the binding with the N-termini may trigger a structural shift similar to the pivoting of the four-helix bundle in YiiP. While the overall structures of YiiP and MmCDF3 are not very much alike, there are some similarities. For example, the active binding sites are tetrahedral in both proteins.

In this chapter CzrB from *Thermus thermophilus* is the focus. It shares the separation of the protein into a transmembrane domain and a cytosolic C-terminal domain. The C-terminus is 92 amino acids long (10.2 kDa) and it appears to form a potentially independent extra-membranal, cytosolic domain (Spada et al., 2002). Through sequencing analysis of the *czrB* gene, it was revealed that the CzrB gene has an ATG consensus sequence at methionine 200, which is closely preceded by a putative ribosome-binding site. The C-terminal domain could potentially exist as part of the full-length protein or as a soluble independent protein (sf-CzrB). *E. coli* strains containing the *CzrB* gene show an enhanced tolerance for zinc. Furthermore, if either the *CzrB* or *sf-CzrB* gene is present in *E. coli*, it has been found that cultures can grow to higher densities in bacteriophage and recombinant protein production (Spada et al., 2002). This suggested an ill-defined chaperone-like function for the corresponding gene products. Here the purification, crystallization and structure determination of the cytosolic domain of CzrB is approached. The tendency of the water-soluble fragment to exist as dimers in solution has been verified in the current study by a variety of solution-based methods including small-angle X-ray scattering, ¹H nuclear magnetic resonance (NMR) spectroscopy, and size-exclusion chromatography.

2.2 Materials and Methods

2.2.1 Materials

BL21 cells and Dnase (Lot 1339311) were purchased from Invitrogen (Carlsbad, CA). Ampicillin (Lot 060573), Luria Bertani broth (Lot 020845), sodium chloride (Lot 060877), and glacial acetic acid (Lot 041410) were obtained from Fisher Scientific (Pittsburgh, PA). Isopropyl β -D-1-thiogalactopyranoside (Lot 11106300), lysozyme (Lot 10683700), and dithiothreitol (Lot 1067 4000) were from American Bioanalytical (Natick, MA). Rnase (Lot 124101816), Ni-NTA agarose (Lot 11230934), Factor Xa removal resin (Lot 12194722) were supplied by Qiagen (Valencia, CA). Tris buffer (Lot 49240B) was obtained from BioRad (Hercules, CA). Triton X-100 (Lot 106F-0052), calcium chloride (Lot 68H06341), and sodium azide (Lot 118H2514) were bought from Sigma (St. Louis, MO). Imidazole (Lot 33271711 494), sodium formate (Lot 1103535 13304122), glycerol (Lot 40959/1 293), polyethylene glycol monomethyl ether 2000 (Lot 449313/1 21504221), and dipotassium hydrogen phosphate (Lot 429860/1 20502) were from Fluka (Buchs, Switzerland). Amicon Ultra-4 concentrators (5,000 MWCO) were purchased from Millipore (Bedford, MA). Dialysis tubing (3,500 MWCO; Lot 3223607) was from Spectra / Por (Rancho Dominguez, CA). Factor Xa protease (Lot N67059-2) was obtained from Novagen (EMD Chemicals, Inc., San Diego, CA). Bradford Albumin Standard (2 mg/mL in 0.9 % aqueous NaCl solution containing sodium azide; Lot GH97262) and Coomassie PlusTM Protein assay (Lot FH71183) were from Pierce (Rockford, IL). For crystallization trials VDXm plates, Hampton Screen I and II (Lots 04169940 and 05039921), polyethylene glycol 4000 (Lot 260522) were obtained from Hampton Research (Aliso Viejo, CA). Zinc chloride (Lot B18339) and potassium phosphate monobasic (Lot D22083) were purchased from Baker Analyzed (Phillipsburg, PA). Ammonium sulfate (Lot 124-1316) was from Jenneile Chemical Company (Cincinnati, OH) and deuterium oxide (99.96 % D₂O; Lot 21768LO) was from Aldrich (Milwaukee, MN). Water, with a resistivity of >18 M Ω cm, was purified by using a Milli-Q Water System (Millipore, Bedford, MA). The system

consists of a carbon filter cartridge, two ion exchange filter cartridges, an organic removal cartridge, and a final filter (Sterile Millipore, millipak 40, lot F6BN22683).

2.2.2 Cloning and Expression

The DNA fragment that contains the C-terminal 92 amino acids long soluble fragment of CzrB was cloned into pIVEX2.4d plasmid (Roche, Basel, Switzerland) via a *NotI-PstI* insert (Spada *et al.*, 2002). A proteolytically cleavable N-terminal 6x His-tag was incorporated. The final fragment consisted of an expressed peptide of 111 amino acids. 19 amino acids originate from the tag sequence and 92 residues from the original peptide sequence. Factor Xa recognizes the amino acid sequence **IEGR** (in bold) and cleaves the C-terminus of (indicated by arrow), leaving a fragment of 94 amino acids (underlined):

MSGSHHHHHSSG**IEGR**[↓]GRMDEGLPPEEVERIRAFLQERIRGRALEVHDLKTRRAGPRSFLEFHLVVRGDTPVEEAHRLCDELERALAQAFPGLQATIHVEPEGERKRTNP.

In this report, there are two species of sf-CzrB, which are referred to as the 6x His-tag sf-CzrB and sf-CzrB, free of its His-tag.

In preparation for over-expression the plasmid (pIVEX2.4d) was incorporated into BL21 cells (Invitrogen, Calsbad, CA, USA) by electroporation. The protein was expressed in Luria-Bertani (LB) broth containing 100 µg/mL ampicillin (American Bioanalytical, Natick, MA, USA). The main culture (2 L) was inoculated with an overnight culture (37 °C, 240 rpm agitation frequency) in the ratio of 50:1 (vol/vol) and grown at 37 °C and 240 rpm agitation frequency (New Brunswick Scientific, C24 Incubator / Shaker, Edison, NJ, USA) to an optical density of 0.5 at 600 nm (Agilent 8453 UV/Vis spectrometer, Agilent Technologies, Santa Clara, CA, USA). The over-expression of sf-CzrB was induced by addition of β-D-thiogalactopyranoside (IPTG) (American Bioanalytical, Natick, MA, USA) to a final concentration of 0.5 mM in the 2 L culture and by lowering both the temperature to 25 °C and the frequency of agitation to 160 rpm. The culture was grown for an additional four hours and then harvested by centrifugation at 8,000 g (Sorvall® RC6™, Kendro Laboratory Equipment, Ramsey, MN, USA) for 10 min at 4 °C. The wet pellet of 8.5 g from a

2 L culture was stored at -80°C .

2.2.3 Protein Purification

The frozen cell pellet (8.5 g) was thawed on ice for 15 min and resuspended in 5 mL cold (4°C) buffer (50 mM Tris-HCl pH 8.0, 300 mM sodium chloride, 10 mM imidazole per 1 g of wet cell pellet. Furthermore, about 5 mg of lysozyme (American Bioanalytical, Natick, MA, USA) per 1 g wet cell pellet and Triton X-100 to 0.1 % (v/v) were added. The mixture was stirred for 1 h on ice. The cells were disrupted by sonication (Sonicator 3000, Misonix, Inc., Farmingdale, NY, USA) on ice with 3 pulses each of 30 s duration at 30 W separated by 2 min intervals. 750 Units Dnase I (Invitrogen, Lot: 1339311, Calsbad, CA, USA) and 0.5 mg Rnase A (Qiagen, Lot: 124101816, Valencia, CA, USA) were added and the mixture was stirred for another hour on ice. The debris was removed by centrifugation (Sorvall® RC6™, Kendro Laboratory Equipment, Ramsey, MN, USA) for 30 min at 47,800 g and 4°C . The supernatant (40 mL) was applied to a 1 mL bed-volume home-made gravity Ni-NTA column of a 1 cm diameter (Ni-NTA agarose, Qiagen, Lot: 11230934, Carlsbad, CA, USA), which was pre-equilibrated with 25 mL of 50 mM Tris-HCl pH 7.5, 150 mM sodium chloride, 20 mM imidazole. The protein was eluted with a constant volume of 15 mL per step, in five steps from 100 mM imidazole to 500 mM imidazole, with an increase of imidazole concentration by 100 mM between each step. Fractions containing 100, 200 and 300 mM imidazole also contained protein and were combined. The sample was dialyzed (dialysis tubing, 3500 MWCO; Lot 3223607, Spectra/Por, Rancho Dominguez, CA, USA) three times against 500 mL of 20 mM Tris-HCl pH 8.0 for 4 h at 4°C . The protein was then concentrated by ultrafiltration (Amicon Ultra-4, 5000 MWCO, Millipore, Bedford, MA, USA) to 10 - 30 mg/mL and used for crystallization trials.

In order to remove the 6x His-tag, the sample was concentrated to 5 mL by ultrafiltration (Amicon Ultra-4, 5000 MWCO) and then washed by ultrafiltration (Amicon Ultra-4, 5000 MWCO Amicon) 4 times with 10 mL of Factor Xa cleavage buffer consisting of 50 mM Tris-HCl pH 6.5, 100 mM sodium chloride, and 5 mM calcium chloride. Factor Xa Protease (Novagen, Lot: N67059-2, EMD Chemicals,

Inc., San Diego, CA, USA) was added to the 6x His-tag sf-CzrB (1U / 3.125 mg/mL 6x His-tag sf-CzrB per milliliter) and incubated with gentle stirring at 20 °C overnight. Precipitate that had formed over night was removed by centrifugation (Sorvall® RC6™, Kendro Laboratory Equipment, Ramsey, MN, USA) for 30 min at 47,800 g and 4 °C. The cleavage enzyme was removed by addition of the Factor Xa removal resin (Qiagen, Lot: 12194722, Valencia, CA, USA) (50 µL resin binds 4 U of factor Xa). After incubation for 10 min at 20 °C with gentle shaking, the resin was removed by centrifugation at 1000 g for 5 min at 20 °C. The supernatant was decanted and its pH adjusted from 6.5 to 8.0 by addition of 1 M Tris-HCl at pH 8.0. Furthermore, the mixture was applied to a gravity Ni-NTA column (as described above). The flow-through containing the cleaved protein (sf-CzrB) was collected, and the cleaved 6x His-tag and uncleaved protein (6x His-tag sf-CzrB) were bound to the column. The cleaved protein was concentrated by ultrafiltration (Amicon Ultra-4, 5000 MWCO) from a volume of approximately 15 mL to about 1.5 mL. This was applied to a gel-filtration column, which was equilibrated with 20 mM Tris-HCl pH 8.0 (volume: 120 – 124 mL, HiLoad™ 16/60, Superdex™ 75, preparative grade, Amersham Biosciences, GE, Piscataway, NJ, USA) to remove remaining impurities. The 6x His-tag free protein was then concentrated by ultrafiltration (Amicon Ultra-4, 5000 MWCO, Millipore, Bedford, MA, USA) to 10 - 30 mg/mL and used for crystallization trials. The yield of the protein was approximately 25 to 35 mg per 1 L culture. The protein concentration was determined by the Bradford assay (Bradford, 1976) using bovine serum albumin (albumin standard; 2 mg/ml in 0.9% aqueous NaCl solution containing sodium azide; Lot GH97262, Pierce, Rockford, USA) and lyophilized salt-free dissolved sf-CzrB as standards. The absorbance response at 595 nm in the range up to 0.25 mg/ml protein was similar for both standards. To evaluate progress during purification and cleavage the protein was checked by SDS page after each step (figure 2.2.1).

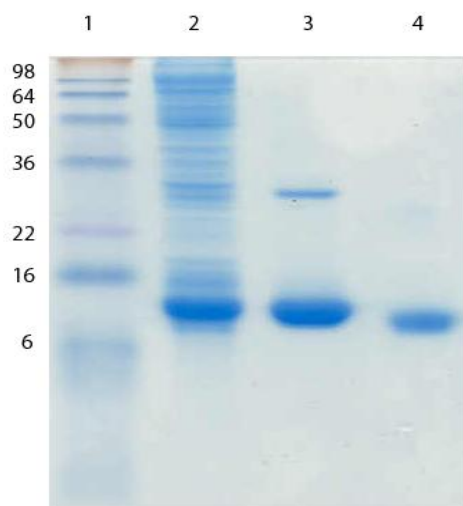


Figure 2.2.1: SDS-PAGE of sf-CzrB at different stages during preparation. Lane 1, molecular-weight standard markers (molecular weights are shown on the left in kDa). Lane 2, soluble protein prepared from *E. coli* BL21 (DE3) cells 4 h post induction with IPTG. Lane 3, 6x His-sf-CzrB following Ni-NTA column purification. Lane 4, sf-CzrB following gel-permeation chromatography. This figure has been taken from Höfer *et al.*, 2007.

2.2.4 Crystallization

2.2.4.1 6x His-tag sf-CzrB Crystallization

Initial crystallization trials were performed with the 6x His-tag sf-CzrB protein in 20 mM Tris-HCl pH 8.0. Hanging drops in VDXm plates (Hampton Research, Aliso Viejo, CA, USA) were set up against Hampton Screen I and II (Hampton, Lot 04169940 and 05039921, Aliso Viejo, CA, USA) using 18 mm silanized circular glass cover slides at 20 °C and 4 °C (1 μ L of protein solution at 12 mg/mL was added to 1 μ L of precipitant over reservoir with 400 μ L of precipitant solution). Crystal hits were detected in 4 M sodium formate, 0.1 M sodium acetate pH 4.6 at 20 °C and were further optimized to 3.0 – 3.5 M sodium formate, 0.1 M sodium acetate pH 4.6. The best crystals reached 150 x 40 x 40 μ m³ in size after 4 days, but typically had visual growth defects (figure 2.2.2(A)). To improve crystal quality, seeding was used. A grown crystal was removed from the crystallization drop and washed in about 5 μ L

precipitant solution. It was transferred again into a drop of 10 μL precipitant solution and the crystal was crushed with a spatula. This seeding solution was transferred into a tube and diluted 1:10 (vol/vol) with precipitant solution. The seeding solution was stored at 20 °C. To set up seeding drops 1 μL of a protein solution at or below 6 mg/mL was mixed with 1 μL of precipitant solution over a reservoir of 400 μL . The drop was equilibrated for two days at 20 °C and then 0.1 μL of 1:10 diluted seeding solution was added. Crystals would grow without visual growth defects within a week after seeding and reached a maximal size of 150 x 75 x 30 μm^3 (figure 2.2.2(B)). Before freezing, crystals were transferred into a drop containing reservoir solution supplemented with 10 %(v/v) glycerol for few seconds and then plunged directly into liquid nitrogen.

2.2.4.2 sf-CzrB Crystallization

Initial crystal conditions for the 6x His-tag free protein (sf-CzrB) were found by screening sf-CzrB against precipitant solutions from the Hampton crystal screen I and II (Hampton Research, Lot 04169940 and 05039921). Hanging drop crystallization was performed in VDXm Plates (Hampton Research); a drop consisted of 1 μL protein solution (15 mg/mL in 20 mM Tris-HCl at pH 8.0) and 1 μL precipitant solution suspended above 300 μL precipitant solution. The well was covered with siliconized glass slides from Hampton Research. Two initial conditions (Crystal screen I, Number 20, 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, 25 %(w/v) polyethylene glycol (PEG) 4000; Crystal Screen II, Number 13, 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, 30 %(w/v) polyethylene glycol monomethyl ether (PEG-MME) 2000 resulted in crystalline material and were optimized by changing the PEG-MME 2000 concentration from 12.5 – 25 %(v/v) to grow crystals of sufficient size and diffraction power. Crystals usually grew within one week to an average size of 300 x 150 x 150 μm^3 . The best diffracting crystals were found under conditions: 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, 15 %(w/v) PEG 4000 and 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, 17.5 %(w/v) PEG-MME 2000 (figure 2.2.2(C)). To yield cryo-conditions for the His-tag free protein the precipitant solution in the reservoir was replaced by either 0.1 M

sodium acetate pH 5.6, 0.2 M ammonium sulfate and 25.5 %(w/v) PEG-MME 2000 or 0.1 M sodium acetate pH 5.6, 0.2 M ammonium sulfate and 25.5 %(w/v) PEG 4000 depending on the initial condition. Drops belonging to the changed reservoir conditions were permitted to adjust to the new conditions for two days at 20 °C and the crystals were harvested and plunged directly into liquid nitrogen.

In preparation for small-angle X-ray scattering (SAXS) measurements, 2.5 µl of a 1.0 M aqueous zinc chloride solution was added to 100 µl sf-CzrB (28 mg/ml) in 50 mM Tris buffer pH 7.5. The initial turbidity produced by the zinc chloride disappeared upon gentle shaking at 4 °C and the clear solution was used for SAXS measurements, which will be reported separately. After five days of storage at 4 °C a shower of protein crystals formed in an unused portion of the SAXS sample (~70 µl stored in a sealed 500 ml tube). Well formed crystals, which reached maximum dimensions of 50 x 50 x 300 µm³ (figure 2.2.2(D)), were harvested from the batch, mounted in immersion oil (Hampton Research HR3-611, type A) and flash-cooled directly in the cryostream immediately prior to data collection.

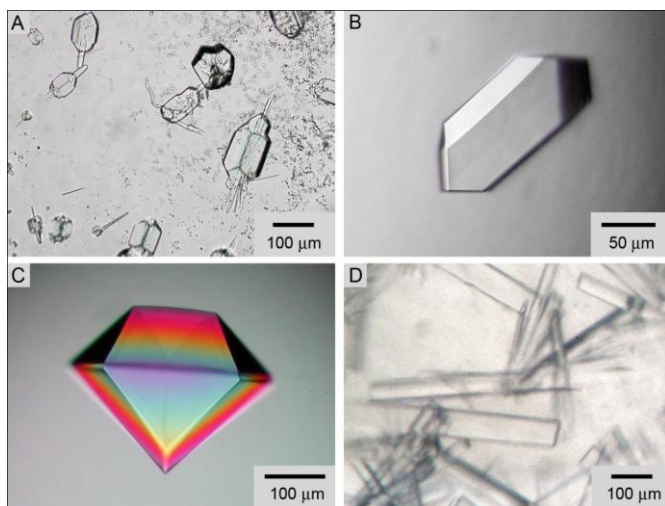


Figure 2.2.2: Crystals of sf-CzrB with and without the 6x His-tag grown with and without added zinc. (A) 6x His-tag sf-CzrB. (B) 6x His-tag sf-CzrB crystals produced by seeding. (C) sf-CzrB. (D) Co-crystals of sf-CzrB and Zn²⁺. The image in (C) was recorded using polarized light. This figure has been taken from Höfer *et al.*, 2007.

2.2.5 X-Ray Diffraction, Structure Determination, and Refinement

X-ray data were collected on the F1 (apo-CzrB_{sf}) and F2 beamlines (Zn-CzrB_{sf}) at the Cornell High Energy Synchrotron Source (CHESS, Ithaca, NY, USA). On F1, a 100 μm diameter 0.9124 \AA beam was used together with an ADSC Quantum-4 detector positioned 180 mm from the crystal. On F2, data were collected using a 20 μm diameter beam produced by single-bounce capillary optics and an ADSC Quantum-210 detector placed 134 mm from the sample. Two crystals were used to obtain full data sets at the zinc edge (1.2827 \AA), the zinc peak (1.2818 \AA), and the zinc high-energy remote (1.2567 \AA). Data on the Zn-CzrB_{sf} crystals were integrated with MOSFLM (Leslie, 1992) and scaled with SCALA (CCP4, 1994). Multiple wavelength anomalous diffraction (MAD) phasing was carried out at 2.2 \AA with SOLVE (Terwilliger and Berendzen, 1999) using the signal from the bound zinc, followed by density modification and automatic chain tracing by RESOLVE (Terwilliger, 2003). Three zinc ions were located and an interpretable map was produced; the iterative build procedure of RESOLVE placed almost the entire polypeptide chain, with unambiguous assignment of residues to the CzrB_{sf} sequence. Refinement was performed using REFMAC (Murshudov et al., 1997), Coot (Emsley and Cowtan, 2004), and O (Jones et al., 1991); for the later stages all available data, to 1.8 \AA resolution, were used. Five translation, libration, screw (TLS) groups were used for the last few refinement steps. The Zn₃ ion is apparently of low occupancy; it was located during refinement and its identity was verified by an anomalous difference Fourier map. Waters were placed automatically by Coot, with a few manual adjustments. The final R and R_{free} values were 0.207 and 0.250, respectively. Details of the structure determination are summarized in table 2.2.1. The apo-CzrB_{sf} crystal diffraction data were integrated and scaled using HKL2000 (Otwinowski, 1997). MOLREP (CCP4, 1994) was used to place the Zn-CzrB_{sf} model without zinc ions in the apo unit cell. ARP/ wARP (Perrakis et al., 2001) was then invoked and about 90% of the model was built automatically. A few residues in the loops were manually built using O and were input into ARP/wARP to complete the model. During the last stages of model building, water molecules were added using ARP/wARP. A final restrained refinement with REFMAC using TLS parameters

brought the R and R_{free} values to 0.187 and 0.224, respectively. The quality of both structures was examined with MolProbity (Lovell et al., 2003). The total MolProbity score placed the Zn-CzrB_{sf} structure in the 95th percentile (score 1.45) and the apo-CzrB_{sf} model in the 97th percentile (score 1.31). The Ramachandran plot revealed that the Zn-CzrB_{sf} structure has 98.9% and 1.1% of its residues occupying the most favored and additionally allowed regions, respectively. The corresponding values for the apo form are 98.8% and 1.2%. Coordinates and structure factors have been deposited in the Protein Data Bank under ID code 3BYP for apo-CzrB_{sf} and 3BYR for zinc-CzrB_{sf}.

Table 2.2.1: Data Collection and Processing Statistics. This table has been taken from Cherezov et al., 2008.

	Zinc-CzrB _{sf}		Apo-CzrB _{sf}	
	Data Collection			
Data Set	Crystal 2, Set x2a	Crystal 1, Set xtallc	Crystal 1, Set x2c	
Wavelength (Å)	1.2827	1.2818	1.2567	0.9124
Resolution (Å)	30-1.80 (1.90-1.80) ^a	30-1.90 (2.00-1.90)	30-1.80 (1.90-1.80) ^a	50-1.70 (1.76-1.70) ^a
Reflections (total/unique)	56,909/ 7,849	46,123/ 6,667	50,341/ 7,811	270,944/ 21,179
Rmerge	0.072 (0.389) ^a	0.126 (0.529) ^a	0.099 (0.644) ^a	0.042 (0.393) ^a
I/δ(I)	20.1 (4.6) ^a	16.0 (3.7) ^a	14.8 (2.6) ^a	54.8 (4.4) ^a
Completeness (%)	98.2 (96.4) ^a	97.5 (93.2) ^a	98.1 (96.3) ^a	98.9 (92.4) ^a
Redundancy	7.3 (7.4) ^a	6.9 (6.6) ^a	6.4 (6.6) ^a	12.8 (6.2) ^a
Space group, unit cell	C222 ₁ ; 50.8x74.5x 43.9 Å ³ 90° 90° 90°		P3 ₂ 12; 60.33x 60.33x 92.14 Å ³ 90° 90° 120°	

Structure Solution

Resolution/ reflections used for MAD	30-2.2 Å /4,414
Zn sites	3 (relative occupancies 1.0, 0.9, 0.4)
Z score, figure of merit	13.4, 0.56
Residues built by RESOLVE	86; 72 placed in sequence
R/Rfree ^b after RESOLVE	0.331/0.336

Refinement

Resolution (Å)	30-1.80 (1.85-1.80) ^a	50-1.70(1.74-1.70) ^a
Final model	Residues 3–91, 4 Zn, 99 HOH	Residues 6–87 (A, B chains), 3 SO ₄ , 301 HOH
Missing structure	Residues 1–2, 92–94; N of Met3; side chains of Met3, Glu86	Residues 1–5, 88– 94; side chains of Glu10, Glu21
Multiple conformations	Glu10, Arg16, Arg61, Arg68	Gln A20, Gln B20, Arg B22, Asp B32, Glu B65, Arg B68
Model atoms	836	1,650
Reflections (total/test set)	7,397/378	20,053/1,086
R/Rfree ^b	0.207/0.250 (0.254/0.365) ^a	0.187/0.224 (0.225/0.319) ^a
R.m.s.d (bond lengths/ angles)	0.008 Å /1.109°	0.012 Å /1.202°
TLS groups	5	2

^a Numbers in parentheses refer to the highest resolution shell.

^b R_{free} was calculated using 5% of data randomly excluded from refinement.

2.2.6 Circular Dichroism

Secondary structure analysis was carried out using 6His-CzrB_{sf} eluted from a Ni-NTA column in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 200 mM imidazole. The buffer was exchanged into 50 mM sodium phosphate, 200 mM NaCl (pH 7.4) by ultrafiltration (2 ml Centricon, 3 kDa cutoff) at 5000 g and 4 °C. A 20 ml sample with a concentration of 3.3 mg protein/ml was transferred into a quartz cuvette with a path length of 0.1 mm. CD data were recorded at 20 °C using an AVIV-202 circular dichroism spectrometer (AVIV Biomedical, Lakewood, NJ, USA) from 280 to 180 nm with 1 nm steps and 2 s integration time per point. A CD spectrum of the corresponding buffer was recorded under the same conditions and subtracted from that of the protein solution. CD data were analyzed by three popular methods from the CDPro program suite (Sreerama and Woody, 2000) using a reference set of 43 soluble proteins.

2.2.7 SAXS

2.2.7.1 Data Collection

Solution scattering data were collected on beamline G1 at CHESS. A custom 1024 x 1024 (69.78 μm) pixel CCD detector fabricated by the Gruner group (Cornell University, Ithaca, NY, USA) was used to measure sample scattering profiles. Two-dimensional images were integrated by Data Squeeze 2.07 (Ddatasqueeze Software, Wayne, PA, USA) to give one-dimensional intensity profiles as a function of the momentum transfer $S = 4\pi\sin\theta/\lambda$ (where 2θ is the scattering angle). Measurements were taken at 20 °C with a sample-to-detector distance of 880 mm. With a calibrated wavelength of 1.236 Å (10.03 keV), scattering profiles covered an S range from 0.025 to 0.327 Å⁻¹. The incident X-ray beam was collimated to a spot size measuring 0.5 x 0.5 mm², which was significantly smaller than the opening of the sample cells. Commercial sample cells holding approximately 12 ml (ALine, Redondo Beach, CA, USA) were fitted with 25 mm thick scratch-free mica windows (Attwater, Lancashire, UK). Silver behenate powder (The Gem Dugout, State College, PA, USA) was used to locate the beam center and to calibrate the sample-to-detector distance. CzrB_{sf} in 50 mM Tris-HCl (pH 7.5) was concentrated to 28 mg/ml. All samples were

centrifuged at 10,000 x g for 5 min prior to loading the sample cell. Matching buffer samples were used for solvent subtraction. The zinc-bound protein was prepared by gently rocking 100 ml of 28 mg protein/ml solution with 2.5 ml 1.0 M ZnCl₂ solution at 4 °C for 5 min. To assess radiation induced aggregation and possible denaturation, samples were exposed twice for 5, 20, and 120 s, comparing Guinier plots of successive curves for nonlinearity at low angle. Protein concentrations of 28 mg/ml and 2.8 mg/ml were also compared to assess concentration dependence. Best data for the apo form of the protein were obtained by combining the low-angle portion of the first 20 s exposure ($S < 0.1 \text{ \AA}^{-1}$) with the wide-angle portion ($S > 0.1 \text{ \AA}^{-1}$) of the first 120 s exposure, both at 2.8 mg protein/ml. The zinc-bound protein solution became cloudy after the first 120 s exposure and showed signs of increased sensitivity to radiation damage. Accordingly, the scattering profile was obtained by combining a 5 s exposure at low angles ($S < 0.1 \text{ \AA}^{-1}$) with a 120 s exposure at wider angles ($S > 0.1 \text{ \AA}^{-1}$) using a sample solution that was 3.5 mg protein/ml.

2.2.7.2 Data Analysis

Pair-distance distribution functions were calculated from scattering profiles using the GNOM program (Svergun, 1992). The maximum diameter of the particle (D_{\max}) was adjusted in GNOM to obtain the best goodness-of-fit parameter (0.827 “good” for the apo state; 0.93 “excellent” for the Zn-bound state). The distribution function fell naturally to near zero beyond D_{\max} but was constrained to zero for the final solution. Radius of gyration (RG) is also computed by GNOM. Expected scattering profiles were calculated from the crystal structures using the CRY SOL program (Svergun et al., 1995). Experimental profiles were fed into CRY SOL at the final stage of computation for automatic superposition. Low-resolution envelopes were obtained from the pair-distance distribution functions using the DAMMIN (Svergun, 1999) program in fast mode. No symmetry was imposed on the solution, so the symmetric appearance of the envelopes in both cases reflects the underlying symmetry of the complex. Envelopes were visualized with PyMOL (DeLano, 2002) using atomic radii set to the dummy atom packing radius determined by DAMMIN.

2.2.8 Size-Exclusion Chromatography

Size-exclusion chromatography was used to assess oligomer formation. For this purpose, the mobility of 10–30 mg C_{zr}B_{sf} on a column (volume 120 ml, HiLoad 16/60, Superdex 75, preparative grade; Amersham Biosciences, GE) in a mobile phase consisting of 50 mM Tris-HCl (pH 8.0) at a flow rate of 1 ml/min and at 24 °C was monitored. The column was calibrated using the following molecular weight marker proteins: myoglobin (16.9 kDa; Sigma), cytochrome c (13.4 kDa; Sigma), hyaluronidase (55 kDa; Sigma), conalbumin (77 kDa; Sigma), and the B1 immunoglobulin binding domain of protein G (GB1, 6.2 kDa) (for details on protein expression and purification, see (Nadaud et al., 2007)).

2.2.9 NMR Spectroscopy

Protein for use in the NMR studies was column-purified as described above, and an Amicon Ultra-15 5000 molecular weight cutoff device (Millipore, Billerica, MA, USA) was used to exchange the buffer and concentrate the protein. The final NMR sample consisted of C_{zr}B_{sf} at a concentration of ~1 mM in an aqueous solution containing 50 mM sodium phosphate (pH 6.5), 5 mM dithiothreitol, 7% (v/v) deuterium oxide (99.96% D₂O; Aldrich, Milwaukee, WI, USA), and 0.02% (w/v) sodium azide in a total volume of 0.5 ml. For the NMR measurements the sample was transferred to a 5 mm sample tube (Wilmad- Labglass, Buena, NJ, USA). NMR experiments were carried out at 25 °C using a Bruker spectrometer (Bruker, Billerica, MA, USA) operating at 600 MHz ¹H frequency and equipped with a triple-resonance three-axis pulsed-field gradient probe optimized for proton detection. The 1-1 pulse echo sequence (Sklenar and Bax, 1987) was used to qualitatively assess the structural integrity of C_{zr}B_{sf} in solution, and to determine the average transverse relaxation time of the amide protons.

2.2.10 Multiple Alignment and Homology Modeling

Multiple sequence alignment of full-length C_{zr}B with its homologs from different organisms was performed using the M-Coffee web server (Moretti et al., 2007).

Despite low strict homology, the alignment shows good consistency and an overall alignment score of 90. The alignment along with the recently solved YiiP structure (Protein Data Bank ID code 2QFI; (Lu and Fu, 2007)) were submitted to the SWISS-MODEL web server (Schwede et al., 2003) to build a full-length CzrB homology model. The model corresponds to the zinc-bound form of CzrB. The cytoplasmic domain of the CzrB model was structurally aligned and replaced with the solved high-resolution zinc-CzrB_{sf} structure. The zinc-CzrB dimer was generated by applying crystallographic symmetry in the zinc-CzrB_{sf} structure to the full-length zinc-CzrB model. To produce a full-length model of the apo-CzrB dimer, we assumed that the orientation of the transmembrane part of CzrB (CzrB_{tm}) in the membrane remains the same regardless of the zinc loading state. Accordingly, the dimeric form of zinc-CzrB_{sf} in the zinc-CzrB model was replaced by the apo-CzrB_{sf} dimer domain, preserving the membrane orientation of the CzrB_{tm} domains. This manipulation resulted in a moving apart of the two CzrB_{tm} domains by ~18 Å and a changing in the angle at the presumed flexible linker between CzrB_{sf} and CzrB_{tm} by ~20°.

2.3 Results

2.3.1 Crystallization

The cytosolic soluble domain of the membrane protein CzcB has been successfully over-expressed and purified (figure 2.2.1, lane 4). Initial crystals had substantial growth defects (figure 2.2.2(A)) and diffracted poorly. Two strategies were developed to overcome the insufficient diffraction power of these crystals. In the first approach seeding was tried and resulted in optically perfect crystals (figure 2.2.2(B)) but unchanged diffraction. The second approach, the removal of the 6x His-tag turned out to be more successful (figure 2.2.2(C)) and resulted in crystals that diffracted up to 1.7 Å. The Matthews coefficient (Kantardjieff and Rupp, 2003; Matthews, 1968) is consistent with 9 or 8 monomers per asymmetric unit for the 6x His-tag sf-CzcB protein crystal. The solvent content is estimated to be 42% or 48%, respectively. Only 2 monomers per asymmetric unit and an estimated solvent content of 48.2% have been suggested for the 6x His-tag free protein (sf-CzcB) crystal.

2.3.2 Crystal Structures

The crystal structures of the apo and zinc forms of CzcB_{sf} have been solved to 1.7 and 1.8 Å resolution, respectively. Refinement statistics are presented in table 2.2.1. The apo form of the protein crystallized as a dimer in the asymmetric unit (figure 2.3.1). With zinc bound, the asymmetric unit contains a single protein molecule. However, application of crystallographic symmetry produces a dimer (figure 2.3.1). Our view is that both forms of the protein exist as dimers under physiological conditions. Support for this conclusion is presented below. Henceforth in this paper they will be treated as dimers. The apo form of CzcB_{sf} includes residues 6–87 in each protomer; the zinc form includes three additional N- and four additional C-terminal residues. The apo and zinc models include three sulfate and four zinc ions, respectively. One of each of the sulfate and zinc ions (Zn4) sits on a special position with occupancy of 0.5. About 100 and 300 structured waters, respectively, have been placed in the zinc and apo forms of the protein.

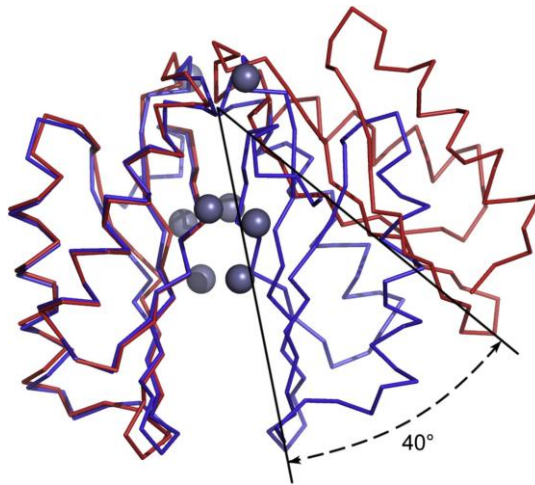


Figure 2.3.1: Structural Rearrangements in CzrB_{sf} upon Zinc Binding C_α traces of apo- (red) and zinc-CzrB_{sf} (blue). Zinc ions are shown as bluish gray spheres. The protomers on the left were overlain to highlight the change in relative orientation of the second protomer upon zinc binding. The figure shows that metal binding triggers a dramatic reorientation of the two monomers with respect to one another, as illustrated by the 40 ° angular shift. This figure has been taken from Cherezov et al., 2008.

2.3.2.1 Apo-CzrB_{sf}

The overall shape of the protomer within the apo homodimer is that of an ellipsoid with a flattened face paralleling its long axis (figures 2.3.1 and 2.3.2). The protomer is highly structured with a distinctly polar surface and a hydrophobic core (figure 2.3.3). It consists of three β strands and two α helices. The former exists as a mixed β sheet with the central strand running parallel and antiparallel to its neighbor on either side. It creates the planar surface of the flattened ellipsoid. The rounded part of the domain is made up of the two helices that run antiparallel to one another and are slightly displaced from each other along their long axes. They sit on top of the β sheet. The connectivity of the different elements in the protomer from the N to the C terminus is $\alpha_1\beta_1\beta_2\alpha_2\beta_3$ (figure 2.3.2). The N terminus of CzrB_{sf} connects to the transmembrane domain presumably by an extension of its C-terminal helix.

The apo form of the protein exists as a dimer in the asymmetric unit (figure 2.3.1). Contact between the monomers along one end of the flattened ellipsoid involves particularly the loop connecting β_2 and α_2 and residues toward the C terminus (hydrogen bonds [two of each per dimer]): Glu57(N)-Gly52(O), 3.0 Å; Val56(N)-Thr54(O), 2.8 Å; hydrophobic contacts involve Val50, Pro55, Val56, Ala59, His60, Val83, and Pro85. This gives the dimer a distinct chevron or inverted V-shaped appearance with the membrane- anchored ends of each protomer splayed apart (figure 2.3.1). Facing one another from either arm of the V are the flattened β sheets of the two protomers.

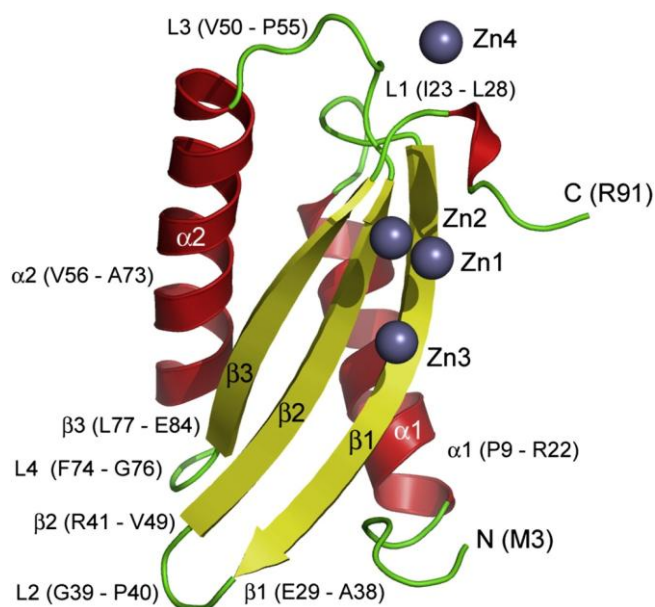


Figure 2.3.2: Cartoon representation of secondary structure elements in the Zinc-CzrB_{sf} model. Color coding is as follows: a helix, red; b strand, yellow; loops and disordered regions, green; zinc ions, bluish gray. Parts of the protein referred to in the text are labeled. The one-letter code is used to identify amino acids. This figure has been taken from Cherezov et al., 2008.

2.3.2.2 Zinc-CzrB_{sf}

The zinc form of the protein is also considered to exist as a homodimer (figure 2.3.1), with very minor changes in protomer secondary and tertiary structure associated with metal binding. A superposition of the apo-CzrB_{sf} and Zn-CzrB_{sf} models (figure 2.3.1) reveals a root-mean-square deviation (r.m.s.d) of 0.7–0.8 Å over 82 common C_α atoms; the r.m.s.d. between protomers in the apo dimer is 0.45 Å. The largest zinc-induced changes occur in the backbone of the β₁-β₂ and β₂-α₂ loops (r.m.s.d, 1–2 Å). In the vicinity of the zinc binding sites the backbone is mostly unaltered compared to the apo form (r.m.s.d < 1 Å). However, the side chains of residues that interact with zinc reorient to provide better coordination (figure 2.3.4). Further, with zinc bound, the density for several additional residues at both the C and N termini are revealed, some of which participate in crystal contacts. The extended C terminus contains a 3₁₀ helical turn. The most dramatic effect of zinc binding is that the two protomers that are splayed apart in the apo form snap together along the flat surfaces of the opposing ellipsoids (figure 2.3.1). The effect comes about as a result of an action that brings the two arms of the inverted V together with a hinge-like motion about the contact surface at the β₂-α₂ loop region. The net effect is to bring the N terminus ends of the two protomers into very close proximity. The distance between N termini in the apo dimer is 33.5 Å. The distance between corresponding residues (Gly6) in the zinc form is 15.2 Å. Because these are connected to the transmembrane domain, it is likely that this large zinc-triggered motion will impact the relative orientation of the cytosolic and transmembrane domains, as discussed below.

The snapping shut of the V upon metal ion binding has the effect that the contact surface area between the two protomers goes from ~400 Å² in the absence of zinc to ~1050 Å² in its presence. The additional contact area is stabilized by hydrogen bonds (two of each per dimer) (Glu45[OE1]-Thr80[OG1], 2.9 Å; His82[ND1]-Ile81[O], 2.8 Å), salt bridges (two of each per dimer) (Arg36[NH1]-Glu5[OE2], 3.0 Å; Arg36[NH2]-Glu5[OE1], 2.6 Å), and by the coordination of a zinc ion (Zn²⁺, two per dimer).

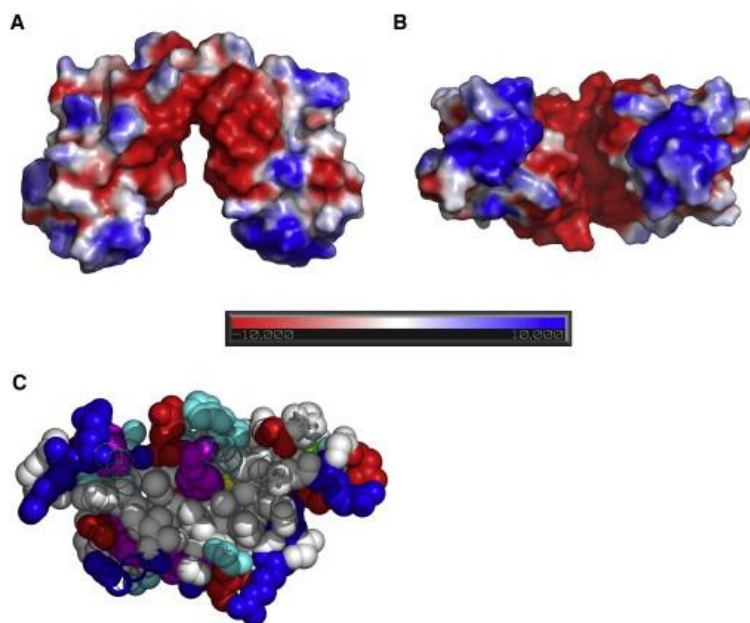


Figure 2.3.3: Surface potential and hydrophobic core of apo-CzrB_{sf}

The surface potential (A and B) (negative, red; positive, blue; neutral, white) was calculated using the Adaptive Poisson-Boltzmann Solver (APBS; <http://apbs.sourceforge.net/>) program. The pronounced negative charge in the cleft between the two protomers likely contributes to keeping the apo form of the dimer in the “open” splayed-apart conformation. The negatively charged cleft is viewed from the side (parallel to the membrane surface) in (A) and from below (normal to the membrane surface) in (B). A cross-section through one of the apo-CzrB_{sf} monomers in (C) clearly shows the hydrophobic core of this cytosolic fragment and polar surface. Residues are colored according to side-chain type: hydrophobic, white; polar, cyan; aromatic, magenta; positive, blue; negative, red. This figure has been taken from Cherezov et al., 2008.

2.3.2.3 Zinc Coordination Sites

Each protomer is associated with four zinc ions (Zn1–Zn4). The first three exist at or close to the interface between protomers in the dimer and may be physiologically relevant. Zn4 is likely only present in the crystal. It sits at a special position shared by two symmetry-related protomers from different dimers. It is coordinated by two

Glu57(OE1) from different dimers and by four waters, and is 2.4 Å distant from Glu57(OE1).

Zn1 is tetraordinated by His47(ND1) (bond length, 2.0 Å), His31(NE2) (2.1 Å), and Glu84(OE2) (1.9 Å), all from the same protomer, and by HOH517 (2.0 Å) (figure 2.3.4). Liganding atoms in the primary coordination sphere around Zn1 are arranged tetrahedrally.

Zn2 is hexacoordinated by His82(NE2) (2.1 Å) and Glu84(OE1) (2.3 Å) in protomer A, by His60(NE2) (2.2 Å) in protomer B, and by three water molecules: HOH585(O) (2.3 Å), HOH527(O) (2.3 Å), and HOH589(O) (2.5 Å) (figure 2.2.4). The geometric arrangement of the six ligand atoms with respect to zinc is octahedral. All residues that are involved in coordinating Zn2 from protomer B are part of helix α_2 , whereas those from protomer A originate in strand β_3 . Given that Zn2 is clearly coordinated by residues in both monomers, it presumably is part of the glue that binds the two protomers together in the dimer. Its symmetry partner does the same thing on the other side of the dimer, effectively pinning the monomers together. Zn3 is likely to be tetraordinated. It is clearly coordinated by His47(NE2) (2.4 Å), Asp32(OD2) (2.0 Å), and HOH516(O) (2.5 Å) (figure 2.2.4). A fourth ligand (possibly water of weak density), however, is not apparent. The lack of density for a fourth ligand can be explained by weaker interactions and higher disorder at this site. In fact, the B factor of Zn3 at 80 Å² is substantially higher than that of the other two Zn sites (17 and 18 Å² for Zn1 and Zn2, respectively). In vivo, this site is likely unoccupied at lower concentrations of zinc. Zn1 and Zn2 are separated by a distance of 4.9 Å and are bridged by the carboxyl oxygens of Glu84. Zn1 and Zn3 are 5.9 Å apart and are bridged by the imidazole of His47. The distances and angles of liganding atoms in the primary coordination spheres of Zn1, Zn2, and Zn3 are close to ideal and are in agreement with zinc serving a structural role in the zinc-CzrB_{sf} dimer (Alberts et al., 1998). The geometry of zinc coordinating residues is close to ideal and does not appear to be distorted by interactions with zinc (table 2.3.1).

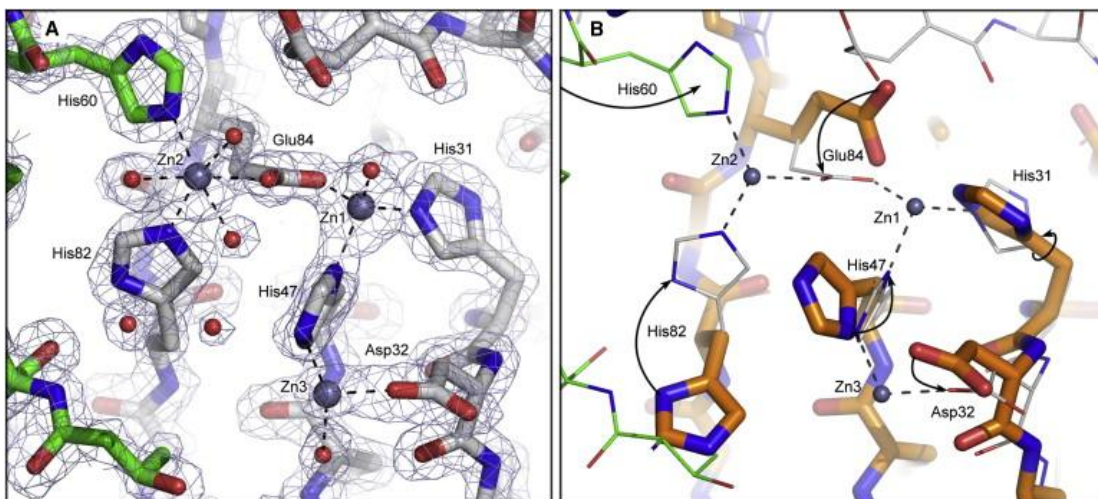


Figure 2.3.4: Zinc coordination at the dimer interface in CzrB_{sf}

(A) Zinc-CzrB_{sf} protomers in different colors. (B) An overlay of zinc and apo protomers. The two protomers in zinc-CzrB_{sf} are distinguished in (A) by having carbon atoms colored gray and green, respectively. Amino acid oxygens and nitrogens are shown as red and blue sticks, respectively. Zinc ions are represented as bluish gray spheres. In (A), water oxygens are shown as red spheres and electron density ($2F_o - F_c$) is contoured at 1.5 σ . The overlay in (B) shows apo-CzrB_{sf} with thick sticks and orange carbon atoms, and zinc-CzrB_{sf} with thin sticks and gray carbon atoms. Arrows highlight the changes side chains undergo upon binding zinc. This figure has been taken from Cherezov et al., 2008.

Table 2.3.1: Geometry of Zinc-Binding Residues in C_{zr}B_{sf} analyzed by MolProbity (Lovell et al., 2003) This table has been taken from Cherezov et al., 2008.

Residue	State, Chain	High B (Å ²)	Ramachandran Φ (°), ψ (°)	Rotamer chi1 (°), chi2 (°), chi3 (°)	Cβ Deviation (Å)
His 31	apo, chain A	33.3	Favored (9.14%) -145.3, 176.8	3.6% 76, 116.9	0.113
	apo, chain B	27.0	Favored (26.25%) -152.5, 166.9	22.5% 71.2, 91.9	0.092
	zinc-	4.3	Favored (5.07%) -148.4, -176.8	3.1 % 66.1, 237.5	0.118
Asp 32	apo, chain A	36.8	Favored (25.58%) 55.1, 41.5	22.2% 203.9, 13.5	0.08
	apo, chain B	27.3	Favored (8.84%) 56.6, 52.9	33.5% 193.4, 29.7	0.121
	zinc-	8.4	Favored (4.73%) 43.4, 54.4	28.1% 196.6, 30.3	0.141
His 47	apo, chain A	30.7	Favored (46.53%) -97.0, 132.8	19.4% 279.7, 60.1	0.053
	apo, chain B	27.1	Favored (43.99%) -96.8, 134.3	20.1% 288.7, 239.7	0.021
	zinc-	6.9	Favored (58.52%) -105.5, 129.7	26.7% 297.5, 189.4	0.07
His 60	apo, chain A	31.2	Favored (91.91%) -62.6, -46.0	28.9% 173.5, 246.5	0.027
	apo, chain B	30.9	Favored (91.66%) -62.2, -46.2	57.2% 168.1, 65.8	0.066
	zinc-	6.2	Favored (75.28%) -69.7, -33.9	51.4% 188.5, 259.8	0.056
His 82	apo, chain A	29.9	Favored (27.38%) -107.4, 109.8	88.7% 304.1, 296.9	0.066
	apo, chain B	32.3	Favored (31.75%) -103.6, 112.5	85.2% 307.8, 285	0.044
	zinc-	10.6	Favored (36.86%) -104.8, 114.3	58% 170.5, 61.7	0.043
Glu 84	apo, chain A	35.8	Favored (22.98%) -131.5, 132.6	86% 178.7, 182, 175.4	0.107
	apo, chain B	35.1	Favored (28.55%) -145.5, 143.3	11.2% 56.2, 186.8, 52.2	0.012
	zinc-	11.8	Favored (68.57%) -134.6, 155.3	63.7% 297.1, 184.8, 131.9	0.033

2.3.3 Characteristics of the Protein in Solution

The crystal structure data suggest that both the apo and zinc forms of the CzrB_{sf} protein exist as homodimers. It is possible that the constraints imposed by the crystal lattice artificially give rise to dimer formation. Because the CzrB_{sf} part of the full length transporter is expected to be extramembranal and to be exposed to the cytosol, it was considered worthwhile to evaluate the tendency of the soluble fragment to oligomerize and to characterize its structure in aqueous solution. Accordingly, its solution properties have been quantified by size-exclusion chromatography, small-angle X-ray scattering, ¹H NMR spectroscopy, and circular dichroism, as follows.

2.3.3.1 Size-Exclusion Chromatography

The hydrodynamic volume and state of oligomerization of apo-CzrB_{sf} in solution were investigated using size-exclusion chromatography. Five water-soluble proteins, whose molecular weights bracketed that of the test protein, were employed as standards. The data (figure 2.3.5(A)) show that apo-CzrB_{sf} elutes with an apparent molecular weight of 21.6 kDa. The calculated molecular weight of the protein, based on its amino acid composition, is 10.8 kDa. These data are consistent with the apo form of the protein existing in solution as a dimer.

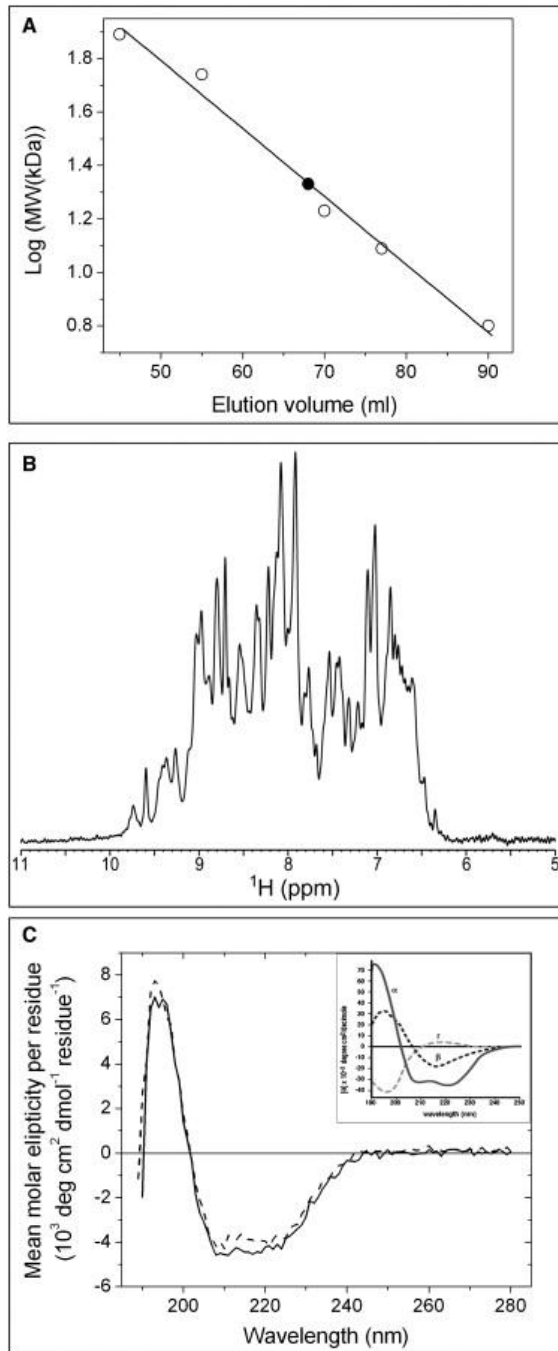


Figure 2.3.5: Characterization of the solution properties of CzrB_{sf}

(A) Size-exclusion chromatography. Apo-CzrB_{sf} eluted at an included volume of 68 ml (filled circle) corresponding to a molecular weight of 21.6 kDa. The identity and molecular weight of the standard proteins used to calibrate the column (open circles) follow: conalbumin (77 kDa), hyaluronidase (55 kDa), myoglobin (16.9 kDa), cytochrome c (12.4 kDa), and GB1 (6.2 kDa). The line of best fit to the data is shown as a solid line. (B) ¹H nuclear magnetic resonance. The amide and aromatic region of the ¹H NMR spectrum of ~1 mM apo-CzrB_{sf} at 25 °C and pH 6.5 is shown. (C) Circular dichroism spectra for secondary structure analysis of the Histagged protein are shown before (solid line) and after lyophilization (dashed line). The inset shows for comparison the circular dichroic behavior of proteins exclusively in the

a-helical (a), antiparallel b sheet (b), and random-coil (r) conformations (Petsko and Ringe, 2004). This figure has been taken from Cherezov et al., 2008.

2.3.3.2 Small-Angle X-Ray Scattering

Small-angle X-ray scattering (SAXS) of protein solutions can yield information about the oligomeric state of the solute and can also provide low-resolution envelopes and other structural information (Svergun, 1999). Accordingly, the SAXS behavior of the apo and zinc forms of C_{zr}B_{sf} in solution was investigated and the corresponding scattering profiles are presented in figure 2.3.6. For comparison, the expected profiles calculated using the X-ray crystallographic models in their dimeric configurations are included in the figure. The experimental solution scattering data are in excellent agreement with the predicted curves and support the view that both apo and zinc forms of the protein exist as dimers in solution. At low resolution ($S < 0.1 \text{ \AA}^{-1}$), the apo- and zinc-C_{zr}B_{sf} scattering profiles overlap. However, they diverge significantly in the vicinity of $S = 0.2 \text{ \AA}^{-1}$. This is consistent with a change in dimer shape upon zinc binding. The radius of gyration for apo-C_{zr}B_{sf} calculated from the scattering data is 18.2 Å. The corresponding value for the zinc form of the protein is 15.7 Å. Similarly, the maximum diameters, D_{max} , which optimize the match of the observed and calculated profiles, are 60 and 55 Å, respectively, for the apo and zinc forms of C_{zr}B_{sf}. These data are consistent with the formation of a more compact dimeric structure upon binding zinc. Low-resolution envelopes constructed from the scattering curves using DAMMIN (Svergun, 1999) are shown in figure 2.3.6 (inset) along with the corresponding envelopes from the crystal structures. These results support the contention that the crystal structures indeed reflect both the configuration and oligomeric state of the apo and zinc forms of the protein in solution.

2.3.3.3 ¹H NMR Spectroscopy

A small section from a one-dimensional 1-1 echo ¹H NMR spectrum of apo-C_{zr}B_{sf} corresponding to the amide and aromatic proton signals is shown in figure 2.3.5(B). The dispersion of the amide proton resonances in the ~ 8–10 ppm region indicates that the protein is folded under the experimental conditions used. Moreover, a series of 1-1 echo ¹H spectra with increasing echo delays was recorded and used to determine the

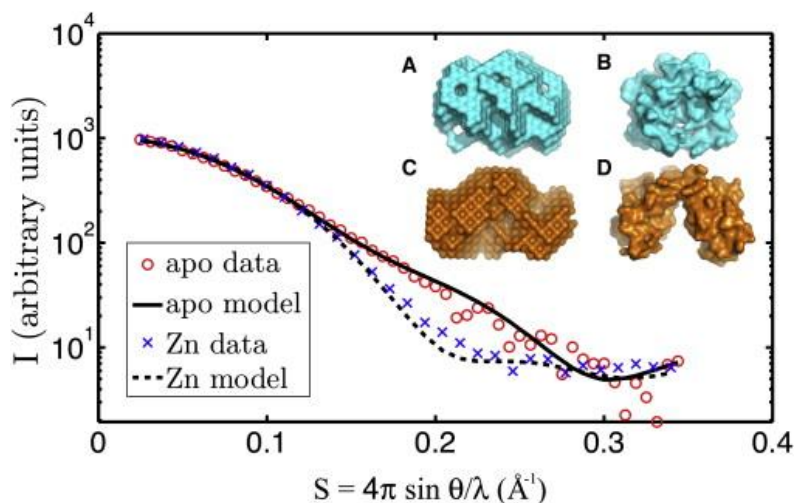


Figure 2.3.6: Small-angle X-Ray scattering from CzrB_{sf} in solution with and without added zinc.

The scattering data are shown as individual data points (with zinc, blue x; without added zinc, red circle). Predicted scattering profiles are shown as solid or dashed lines along with the corresponding zinc- (inset [A]) and apo-CzrB_{sf} (inset [C]) models. For comparison, the crystal structures of zinc and apo-CzrB_{sf} are shown in insets (B) and (D), respectively. This figure has been taken from Cherezov et al., 2008.

average transverse relaxation time, T_2 , of the amide protons (H^N). As discussed by Anglister (Anglister et al., 1993), the average $H^N T_2$ can be used to estimate the effective protein rotational correlation time (τ_c) according to $\tau_c \approx 1/(5T_2)$ ns, where T_2 is the average $H^N T_2$ value in seconds. For the ~ 1 mM apo-CzrB_{sf} solution at 25 °C the average amide proton T_2 was 13.3 ms, which gives an effective of $\tau_c \sim 15$ ns. The rotational correlation time of a protein is determined by both molecular size and shape, and the estimated τ_c value of ~ 15 ns is characteristic of proteins having molecular weights in the ~ 20 – 30 kDa range (Cavanagh et al., 2007). This clearly indicates that apo-CzrB_{sf} oligomerizes in solution, most likely into dimeric form. Additional NMR spectra (data not shown), recorded at 25 °C and pH 6.5 as a function of protein concentration (between ~ 0.25 and 1 mM protein) and ionic strength (between 0 and 100 mM NaCl), revealed effectively identical H^N resonance

frequencies and average T_2 values, indicating the high stability of CzrB_{sf} dimers in solution.

2.3.3.4 Circular Dichroism

Circular dichroism in the UV region provides useful information on the secondary structure make-up of a protein in solution. Apo-CzrB_{sf} with the His tag in place was subjected to this form of analysis and the data are shown in figure 2.3.5(C). The spectrum reveals the characteristic negative peaks in ellipticity in the vicinity of 208 and 222 nm associated with α helices and at 217 nm associated with β strands. Deconvolution of the spectrum using two different algorithms in the CDPro program suite provided an average probable secondary structure composition of 18% α helix, 30% β strand, 25% turn, and 27% disordered region (table 2.3.2). The corresponding values obtained based on the X-ray structure are 35%, 30%, 12%, and 23%, respectively (DSSP) (Kabsch and Sander, 1983). The disparities should be viewed in light of the fact that the CD values are based on a polypeptide containing 111 residues, the N-terminal segment of which includes a His tag and is presumably disordered. In contrast, the crystal structure is based on an analysis of 82 “ordered” residues in the apo form of CzrB_{sf}. Lyophilization of the protein had little effect on its secondary structure as revealed by CD analysis (figure 2.3.5(C)).

Table 2.3.2: Results of Analysis of CD Spectra from Apo-CzrB_{sf} Before/After Lyophilization Using the CDPro Suite (Sreerama and Woody, 2000). This table has been taken in modified form from Cherezov et al., 2008.

Method	α -helices	β -strands	turns	disordered
CDSSTR	0.163 / 0.071	0.306 / 0.35	0.26 / 0.242	0.269 / 0.329
CONTINLL	0.196 / 0.177	0.297 / 0.307	0.237 / 0.229	0.269 / 0.288

2.4 Discussion

The crystal structure of the cytoplasmic domain of the putative zinc transporter CzrB in *Thermus* determined in this study is consistent with the protein functioning *in vivo* as a dimer. This was corroborated by solution studies with CzrB_{sf} that involved size exclusion chromatography, ¹H NMR, and small-angle X-ray scattering. In what follows, the zinc binding characteristics of the cytoplasmic domain are examined and a proposal is advanced for how CzrB_{sf} may function as a metallochaperone and as an independent regulator of CzrB zinc transport activity.

2.4.1 Zinc Binding

We anticipate that three of the four zinc ions observed in the zinc-CzrB_{sf} structure are physiologically relevant (figure 2.3.1). The fourth, Zn4, likely plays a role in crystallogenesis and crystal packing. The Zn1 and Zn2 sites are homologous with the two zinc sites in the cytosolic domain of the *E. coli* zinc transporter YiiP (figure 2.4.1) (Lu and Fu, 2007). As in YiiP, they straddle the three centrally located β strands of the domain, with Zn2 acting to bridge protomers in the dimer. The two zinc ions are 4.2 Å apart in YiiP. The corresponding separation in CzrB_{sf} is 4.9 Å. The zinc coordination is similar between the two proteins, but there are notable differences. For example, the carboxyl of the highly conserved Glu84 in CzrB_{sf} is coordinated to Zn1 and Zn2. The corresponding residue in YiiP is Asp285, where just one of the carboxyl oxygen atoms appears to engage in bidentate coordination to bridge the two zinc ions. One other notable difference arises at His47, where the imidazole side chain of this conserved residue bridges Zn2 and Zn3 in CzrB_{sf}. In YiiP the corresponding residue is Ile245. We speculate that the disparity between CzrB_{sf} and YiiP at this residue and in this region of the structure arises from a three residue frame shift in the region between loops L2 and L3 that includes strand β 2 in YiiP (figure 2.4.2). The misalignment is perhaps understandable, given that the YiiP structure is reported at 3.8 Å.

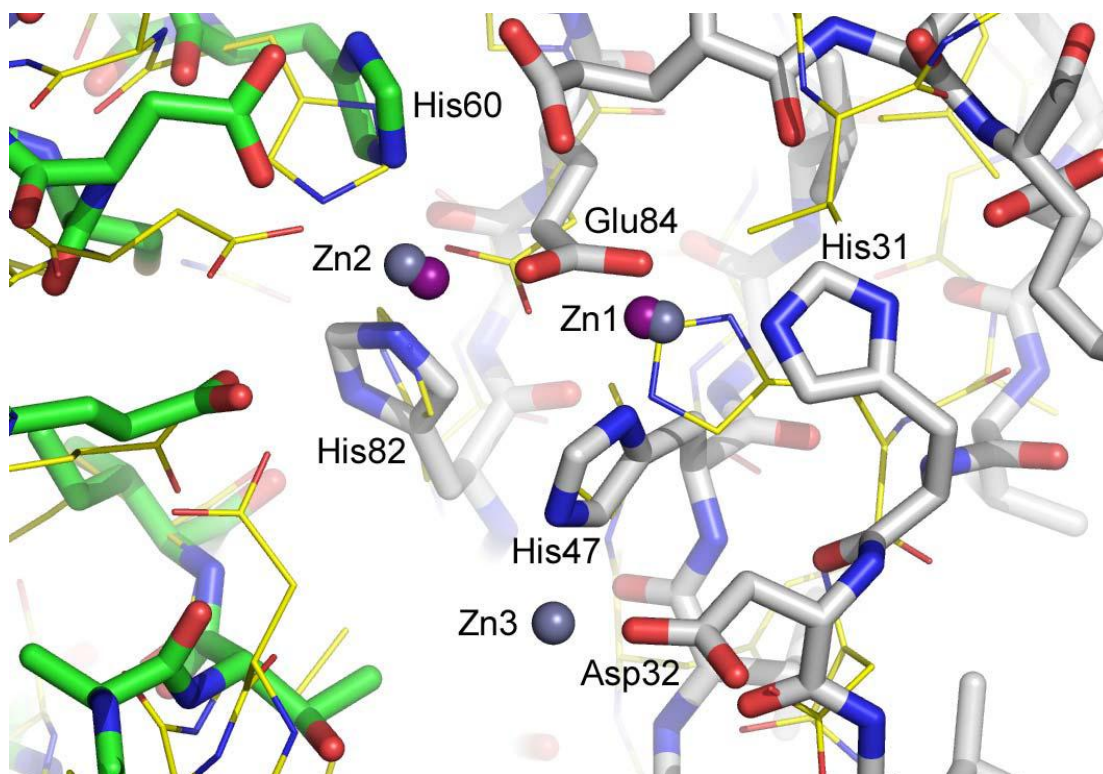


Figure 2.4.1: Overlay of zinc-CzrB_{sf} (thick sticks, gray and green carbons) and YiiP (thin sticks, yellow carbons) at homologous zinc binding sites.

The same view is shown as in figure 2.3.4. Zinc ions in CzrB_{sf} and in YiiP are colored bluish gray and purple, respectively. The distance between common zinc ions in CzrB_{sf} and in YiiP is 4.9 Å and 4.2 Å, respectively. Notable differences between the two structures include i) Glu84 in CzrB_{sf} which corresponds to Asp285 in YiiP and ii) His47 in CzrB_{sf} which corresponds to Ile245 in YiiP. The latter may be accounted for by a 3-residue frame shift along strand β2 in YiiP as described in the text and in figure 2.4.2. This figure has been taken from Cherezov et al., 2008.

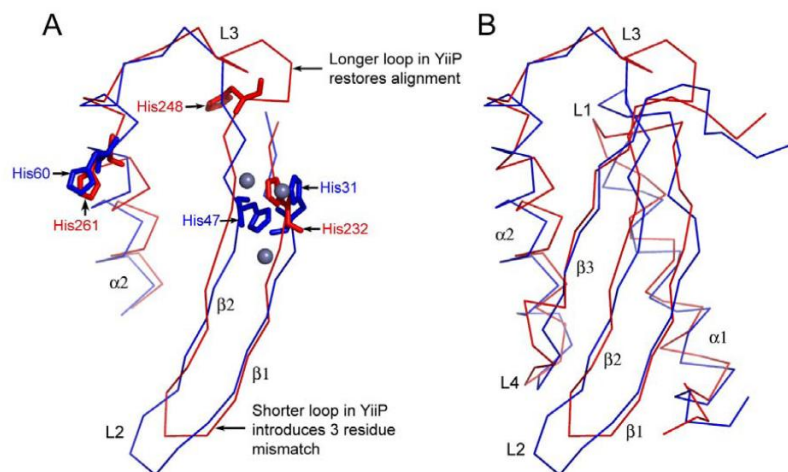


Figure 2.4.2: An overlay of the C α trace of zinc-CzrB_{sf} (blue) and YiiP (red) in the region that includes strands β 1 and β 2 and loops L2 and L3 (A) and for the full-length Zinc-CzrB_{sf} (B).

The zinc-coordinating residues His31, His47 and His60 in CzrB_{sf} and sequence homologous residues in YiiP are shown in stick representation. Zinc atoms in zinc-CzrB_{sf} are shown as bluish gray spheres. The structural coincidence of residues CzrB_{sf} homologous to His31 and His60 in YiiP is obvious. However, His47 in CzrB_{sf} and its sequence homologue His248 in YiiP are displaced along the β 2 strand by \sim 3 residues. Interestingly, L2 is shorter whilst L3 is longer by about 3 residues in YiiP compared to CzrB_{sf}. This suggests that the sequence was incorrectly threaded into the electron density of YiiP beginning at L2. The frame shift of 3 residues that ensued was likely corrected subsequently by a second misthreading at L3. This figure has been taken from Cherezov et al., 2008.

In the absence of zinc, the surface potential of CzrB_{sf} in the cleft between the two protomers is highly negative (figures 2.3.2A and 2.3.2B). By charge repulsion, this will contribute to holding the two protomers apart and to stabilizing the inverted V shape of the apo form of the protein. The negative potential arises primarily from the side chains of Glu45, Glu67, Glu84, Asp32, and Asp64. Upon zinc binding the side-chain carboxyls of these residues form hydrogen bonds (Glu45-Thr80), coordinate directly with zinc (Asp32, Glu84), or do so via a water molecule (Glu67, Asp64). The large negative surface potential between protomers in the apo-CzrB_{sf} dimer will serve

to direct zinc into the cleft region as intracellular concentrations of the metal ion rise. We speculate that the Zn1 site will be the first to fill because three of its coordinating ligands (His31, His47, Glu84) already exist on the surface of apo-CzrB_{sf} with side chains oriented in such a way that only small changes are required to properly coordinate Zn1, as illustrated in figure 2.3.4. The rotation of the imidazole ring of His47 to ligand Zn1 may in turn facilitate the repositioning of Asp32 to coordinate with the second zinc, Zn3, via its side-chain carboxyl. It may also facilitate the relatively large movement of the His82 imidazole, which contributes to creating a binding pocket for Zn2. The latter is stabilized also by the carboxyl of the Glu84 side chain, already suitably positioned following the locking in of Zn1. With this “partial pocket” in place, the stage is set for the neighboring protomer to rotate into position and to bring its surface-exposed His60 in on top of the anchored Zn2. Along with three solvent molecules, this then completes the hexacoordination of the protomer-bridging Zn2. The net effect of this sequential zinc binding is to snap the open end of the dimer shut and to change the relative orientation of the cytosolic and transmembranal domains with possible consequences for interactions with metallochaperones and zinc transport, as outlined below.

2.4.2 Homology Modeling and Mode of Action

YiiP is a Zn²⁺/H⁺ antiporter in *E. coli*. The zinc-bound form of the full-length transporter was solved recently to a resolution of 3.8 Å (Lu and Fu, 2007). CzrB is a homolog with an overall sequence identity of 32.6%. The corresponding values for the cytoplasmic and transmembranal domains are 31.5% and 33.2%, respectively. Despite the relatively low sequence homology in the cytoplasmic domains, the structural homology between them is impressive. Thus, for example, zinc-CzrB_{sf} overlays the cytoplasmic domain of YiiP with a C_α-r.m.s.d of 1.8 Å over 79 residues (figure 2.4.2). The transmembranal domain of YiiP is made up of a bundle of six tightly packed helices. Five of these extend across the membrane and encircle a sixth short helix in a way that creates two cavities at either end of the bundle. One cavity is exposed to the periplasm (and to the outer leaflet of the membrane) and contains a zinc ion at its base. The second is exposed to the cytosol, and access to it is

presumably by way of a space that exists between the cytoplasmic and transmembrane domains. It has been proposed that a metallochaperone delivers zinc to this intracellular cavity. The ion then makes its way across the membrane into the extracellular cavity where effectively it is out of the cell (Lu and Fu, 2007). The model shows the YiiP dimer in the form of a Y. The stem of the Y is the cytoplasmic dimer and the two arms represent the transmembrane domains that diverge from one another as they angle across the bilayer. In the current study, we have solved the structure of the apo and zinc-bound forms of CzrB_{sf}. The YiiP structure, which is of the full-length transporter, is only available in the zinc-bound state. By homology modeling to the latter, we set about exploring the changes in the full-length form of CzrB that might accompany zinc binding to its cytoplasmic domain, CzrB_{sf}, and that may provide clues to the mode of action of the protein as a cation facilitator transporter. Accordingly, a structural homology model of the transmembrane domain of CzrB (CzrB_{tm}) was created using the corresponding domain of YiiP (YiiP_{tm}). This was spliced to the high-resolution structure of CzrB_{sf} in both its apo and zinc-bound states, with a view to producing a picture of the full-length protein. In this modeling exercise, the orientation of the transmembrane domain in the bilayer “observed” with YiiP (Lu and Fu, 2007) was preserved in the CzrB_{sf}-CzrB_{tm}/YiiP_{tm} “chimera.” As noted, in the apo form of CzrB_{sf} the protomers of the dimer are splayed apart. They snap together into close proximity upon zinc binding. When this ion-binding event happens in the chimera, the angle and the space between the cytoplasmic and transmembrane domains changes dramatically (shaded blue in figure 2.4.3). This is the region where the holo-metallochaperone (green in figure 2.4.3) is proposed to interact with the transporter and to transfer its cargo of metal ions to the intracellular cavity for translocation across the membrane. In this model, zinc binding to the CzrB_{sf} triggers the conformational change that enables chaperone docking and zinc delivery for transport through the transmembrane part of the protein. This proposal requires that the orientation of the transmembrane domain of the full-length transporter in the membrane remains fixed and independent of zinc status. It also requires that zinc binding trigger a clamping together of the cytosolic domains in the dimer. The net effect is to alter the angle between the cytosolic and transmembrane

domains in a way that enables interaction with a zinc-bearing chaperone. High-resolution X-ray structures of the full-length CzrB with and without zinc will go a long way toward evaluating this proposal. It is interesting to consider the possibility that CzrB_{sf}, which may be expressed in the cell as an independent protein, itself acts as a metallochaperone. Further, given that the full-length protein would appear to function as a homodimer of form CzrB-CzrB, the monomeric form of CzrB_{sf} could destabilize the active full-length species by competing with the partnered cytoplasmic domains. This would lead to an inactive heterodimer of form CzrB-CzrB_{sf} (figure 2.4.4). In this hybrid, it is possible for the two CzrB_{sf}s to snap shut upon zinc binding but without any attendant change in the relative orientations of the cytoplasmic and transmembrane domains in the full-length monomer. As a result, no docking site is created for the metallochaperone to bind and to deliver its cargo. Regardless of zinc status, then, the CzrB-CzrB_{sf} complex would be incompetent as a transporter. In this way, CzrB_{sf} could regulate the transport activity of the protein.

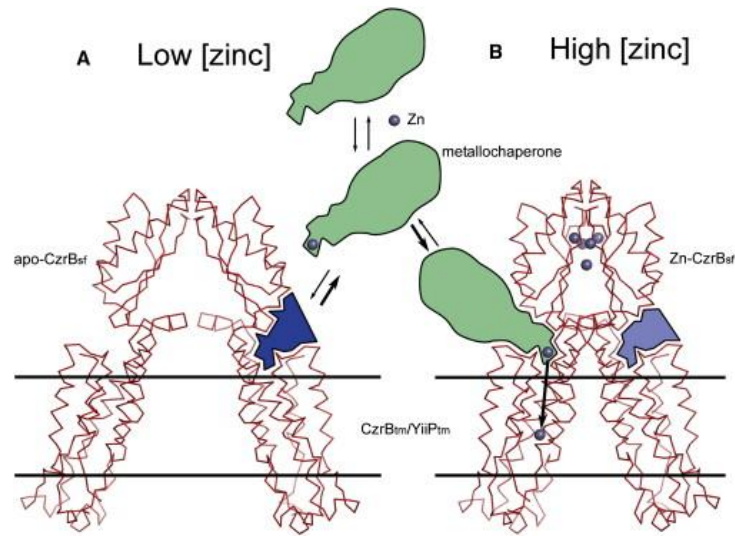


Figure 2.4.3: A model for CzrB action as a zinc transporter

The proposed mechanism is based on the changes induced in the dimer conformation of CzrB_{sf} upon zinc binding (figure 2.3.1) combined with homology modeling from YiiP (Lu and Fu, 2007) to the CzrB transmembrane domains. The structures shown are chimeras of apo-CzrB_{sf} and CzrB_{tm}/YiiP_{tm} (A) and zinc-CzrB_{sf} and CzrB_{tm}/YiiP_{tm} (B) as described in the text. At low zinc levels, the relative orientation of the CzrB_{sf} and CzrB_{tm} domains is such that the zinc-loaded metallochaperone (green) fails to dock productively (see site represented by purple in [A]) with CzrB to deliver its cargo of metal. Binding of zinc to CzrB_{sf} causes the dimer to snap shut, bringing the two transmembrane domains into close proximity and creating sites (represented by the blue area in [B]) conducive to zinc-metallochaperone docking and to zinc ion delivery. This figure has been taken from Cherezov et al., 2008.

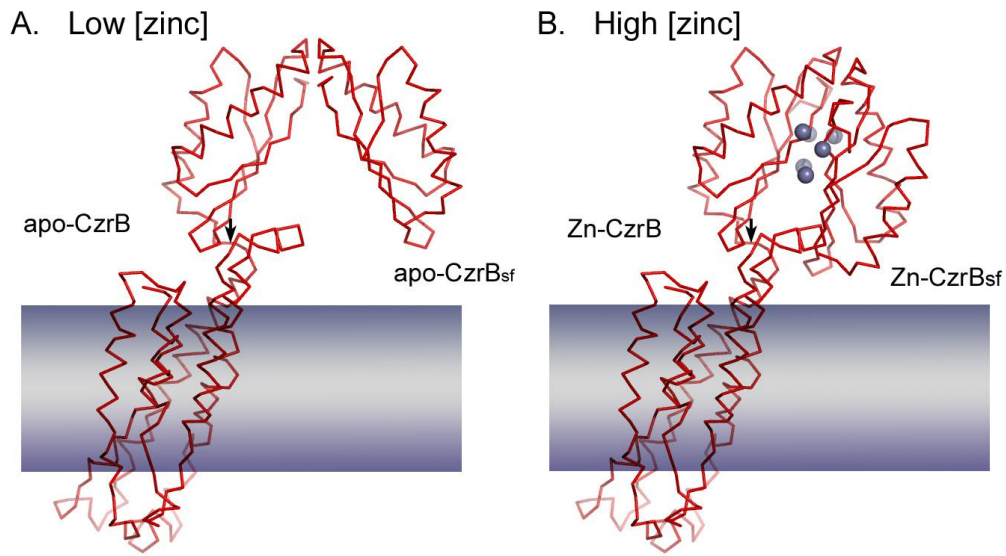


Figure 2.4.4: A model to explain how CzrB_{sf} can render CzrB inactive as a zinc transporter. This figure has been taken from Cherezov et al., 2008.

Through CzrB-CzrB_{sf} heterodimer formation based on the structure of the apo-CzrB_{sf} dimer in figure 2.3.1 and the model in figure 2.4.3 we propose that CzrB_{sf} displaces one full-length CzrB protomer from the apo form of the dimer and produces the heterodimer shown in A. Upon zinc binding to the apoheterodimer (B) the CzrB_{sf} is entirely free to rotate into position next to its partner in the full-length monomer without the need for a bending motion at the hinge (arrow) between the cytosolic and transmembrane domains of the latter. Failure to rotate at the hinge means that the site for metallochaperone docking (vicinity of arrow, see figure 2.4.3) does not materialize and the heterodimer is rendered inactive as a zinc transporter. In the heterodimer therefore zinc binding is effectively uncoupled from the creation of a chaperone docking site. Zinc ions are shown as bluish gray spheres. The membrane is represented as a shaded rectangle the upper surface of which faces the cytoplasm.

2.5 Conclusions and Outlook

The crystal structure of the cytoplasmic domain of the putative zinc transporter CzrB from *Thermus thermophilus* has been solved. Structures, determined to resolutions of 1.7 and 1.8 Å in the apo and zinc-bound forms, respectively, are consistent with the protein functioning *in vivo* as a dimer. Support for dimer formation was provided by solution studies involving size-exclusion chromatography, ¹H NMR, and small-angle X-ray scattering. The overall structure of the CzrB_{sf} monomer, which consists of two antiparallel α helices sitting atop a three-stranded mixed β sheet, is stabilized by an apolar core and is quite insensitive to zinc binding. However, the conformation of the dimer changes dramatically upon zinc binding. In the apo form, the two monomers associate at one end of the molecule and are splayed apart at the other. With zinc loading, the monomers snap together along their full length. The zinc-bound form of the protein includes four zinc ions, one of which plays a role in crystallogenesis. The other three are located in an anionic cleft at the monomer-monomer interface and are considered to be physiologically relevant. Zinc coordination involves highly conserved histidine, glutamate, and aspartate residues and water molecules. Full length variants of CzrB in the apo and zinc-loaded states were generated by homology modeling with the known Zn²⁺/H⁺ antiporter YiiP. The model suggests a way in which zinc binding to the cytoplasmic fragment creates a docking site to which a metallochaperone can bind for delivery and transport of its zinc cargo. Because the cytoplasmic domain may exist in the cell as an independent, soluble protein, a proposal is advanced that it functions as a metallochaperone and that it regulates the zinc-transporting activity of the full-length protein. The latter requires that zinc binding becomes uncoupled from the creation of a metallochaperone-docking site on CzrB. A more complete evaluation of these assorted proposals awaits a high-resolution structure of the full-length protein in its apo and zinc-bound forms.

2.6 References

- Alberts, I.L., Nadassy, K., and Wodak, S.J. (1998). Analysis of zinc binding sites in protein crystal structures. *Protein Sci* 7, 1700-1716.
- Anglister, J., Grzesiek, S., Ren, H., Klee, C.B., and Bax, A. (1993). Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS. *J Biomol NMR* 3, 121-126.
- Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G.W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogra D* 67, 271-281.
- Bloss, T., Clemens, S., and Nies, D.H. (2002). Characterization of the ZAT1p zinc transporter from *Arabidopsis thaliana* in microbial model organisms and reconstituted proteoliposomes. *Planta* 214, 783-791.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72, 248-254.
- Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Rance, M.S., and Skelton, N.J. (2007). *Protein NMR Spectroscopy* San Diego: Academic Press.
- Coudraya, N., Valvob, S., Hua, M., Lasalaa, R., Kima, C., Vinka, M., Zhoud, M., Provasic, D., Filizolac, M., Taoe, J., Fange, J., Penczeke, P.A., Ubarretxena-Belandiac, I., Stokes, D.L. (2012). Inward-facing conformation of the zinc transporter YiiP revealed by cryoelectron microscopy. *P.N.A.S.* In press.
- CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-763.
- Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66, 12-21.
- Conklin, D.S., McMaster, J.A., Culbertson, M.R., and Kung, C. (1992). COT1, a gene involved in cobalt accumulation in *Saccharomyces cerevisiae*. *Mol Cell Biol* 12, 3678-3688.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System. De-Lano Scientific, San Carlos, CA, USA <http://www.pymol.org>.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.

Grass, G., Otto, M., Fricke, B., Haney, C.J., Rensing, C., Nies, D.H., and Munkelt, D. (2005). FieF (YiiP) from *Escherichia coli* mediates decreased cellular accumulation of iron and relieves iron stress. *Arch Microbiol* 183, 9-18.

Haney, C.J., Grass, G., Franke, S., and Rensing, C. (2005). New developments in the understanding of the cation diffusion facilitator family. *J Ind Microbiol Biotechnol* 32, 215-226.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47 (Pt 2), 110-119.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637.

Klingenberg, M., (2005). Ligand-protein interaction in biomembrane carriers: The induced transition fit of transport catalysis. *Biochemistry* 44(24):8563–8570.

Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography No. 26.

Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* 50, 437-450.

Lu, M., Chai, J., and Fu, D. (2009). Structural basis for autoregulation of the zinc transporter YiiP. *Nat Struct Mol Biol* 16, 1063-1067.

Lu, M., and Fu, D. (2007). Structure of the zinc transporter YiiP. *Science* 317, 1746-1748.

Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491-497.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658-674.

Montanini, B., Blaudez, D., Jeandroz, S., Sanders, D., and Chalot, M. (2007). Phylogenetic and functional analysis of the Cation Diffusion Facilitator (CDF) family: improved signature and prediction of substrate specificity. *BMC Genomics* 8, 107.

Moretti, S., Armougom, F., Wallace, I.M., Higgins, D.G., Jongeneel, C.V., and Notredame, C. (2007). The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Res* 35, W645-648.

Munkelt, D., Grass, G., and Nies, D.H. (2004). The chromosomally encoded cation diffusion facilitator proteins DmeF and FieF from *Wautersia metallidurans* CH34 are transporters of broad metal specificity. *J Bacteriol* 186, 8036-8043.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53, 240-255.

Nadaud, P.S., Helmus, J.J., Hofer, N., and Jaroniec, C.P. (2007). Long-range structural restraints in spin-labeled proteins probed by solid-state nuclear magnetic resonance spectroscopy. *J Am Chem Soc* 129, 7502-7503.

Nies, D.H., Nies, A., Chu, L., and Silver, S. (1989). Expression and nucleotide sequence of a plasmid-determined divalent cation efflux system from *Alcaligenes eutrophus*. *Proc Natl Acad Sci U S A* 86, 7351-7355.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* 276, 307-326.

Outten, C.E., and O'Halloran, T.V. (2001). Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* 292, 2488-2492.

Paulus, H., Sarkar, N., Mukherjee, P.K., Langley, D., Ivanov, V.T., Shepel, E.N., and Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001). ARP/wARP and molecular replacement. *Acta Crystallogr D Biol Crystallogr* 57, 1445-1450.

Petsko, G.A., and Ringe, D. (2004). *Protein Structure and Function*. London: New Science Press.

Ramachandran, G.N., and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv Protein Chem* 23, 283-438.

Russell, D., and Soulimane, T. (2012). Evidence for zinc and cadmium binding in a CDF transporter lacking the cytoplasmic domain. *FEBS Lett* 586, 4332-4338.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 31, 3381-3385.

Sklenar, V., and Bax, A. (1987). Spin-echo water suppression for the generation of pure phase two-dimensional NMR spectra. *J Magn Reson* 74, 469.

Spada, S., Pembroke, J.T., and Wall, J.G. (2002). Isolation of a novel *Thermus thermophilus* metal efflux protein that improves *Escherichia coli* growth under stress conditions. *Extremophiles* 6, 301-308.

Sreerama, N., and Woody, R.W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287, 252-260.

Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRY SOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* 28, 768.

Svergun, D.I. (1992). Determination of the regularization parameter in indirect transform methods using perceptual criteria. *J Appl Crystallogr* 25, 495–503.

Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76, 2879–2886.

Terwilliger, T.C. (2003). SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol* 374, 22-37.

Terwilliger, T.C., and Berendzen, J. (1999). Evaluation of macromolecular electron-density map quality using the correlation of local r.m.s. density. *Acta Crystallogr D Biol Crystallogr* 55, 1872-1877.