



**HAL**  
open science

# The function of a membrane protein : studies of structure and dynamics by NMR

Vilius Kurauskas

► **To cite this version:**

Vilius Kurauskas. The function of a membrane protein : studies of structure and dynamics by NMR. Structural Biology [q-bio.BM]. Université Grenoble Alpes, 2017. English. NNT : 2017GREAV005 . tel-01685239

**HAL Id: tel-01685239**

**<https://theses.hal.science/tel-01685239>**

Submitted on 16 Jan 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **THÈSE**

Pour obtenir le grade de

### **DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES**

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 25 mai 2016

Présentée par

**Vilius KURAUSKAS**

Thèse dirigée par **Beate BERSCH (CSV) CNRS**, et  
codirigée par **Paul SCHANDA (EDCSV) CEA**

préparée au sein du **Institut de Biologie Structurale**  
dans l'**École Doctorale Chimie et Science du Vivant**

## **Fonction d'une protéine membranaire: étude structurale et dynamique par RMN**

Thèse soutenue publiquement le **18 janvier 2017**,  
devant le jury composé de :

**Madame BEATE BERSCH**

CHARGE DE RECHERCHE, CNRS DELEGATION ALPES, Examinateur

**Monsieur EDMUND R.S. KUNJI**

PROFESSEUR, MRC – ROYAUME-UNI, Rapporteur

**Monsieur ALAIN MILON**

PROFESSEUR, UNIVERSITE TOULOUSE 3, Rapporteur

**Madame EVA PEBAY-PEYROULA**

PROFESSEUR, UNIVERSITE GRENOBLE ALPES, Président

**Madame STEPHANIE RAVAUD**

MAITRE DE CONFERENCES, UNIVERSITE GRENOBLE ALPES,  
Examinateur

**Monsieur MARTIN TOLLINGER**

DIRECTEUR DE RECHERCHE, UNIVERSITE D'INNSBRUCK-  
AUTRICHE, Examinateur





# Acknowledgements

The work presented in the following pages would have been impossible without direct and indirect contributions by a plethora of people. I am thankful beyond the words to my wonderful supervisors, Beate and Paul. They gave me an opportunity to spend three amazing years in the beautiful city of Grenoble, working in a team of many great people and great scientists. Throughout the course of this difficult project there was not a single moment when I would not feel a strong support from both of them, and any time I would stumble on a scientific problem they would be around with insightful solutions. Thank you, Beate, for believing in my abilities more than I did myself, for the patience with my not always easy personality, for all the advice I received from you and for the help in navigating the dark corridors of French bureaucracy. Thank you Paul for sharing your scientific enthusiasm, for all the Brobdingnagian effort in teaching me NMR, for allowing to bother you every ten minutes and patiently listening to my not always comprehensible blabber and most importantly for taking such a great care in my growth as a scientist and as a person.

I am also extremely thankful to my colleagues, some of who contributed substantially in advancing the project, while the others helped me to build many unforgettable memories in Grenoble. Thank you Audrey, Sergi, Peixiang, Roberto, Rida, Nicolas, Elisa, Hugo, Jia-Ying, Katharina, Tomáš, Diego, Orso, Caroline, Zsofia, Emmanuelle... I am especially thankful to my fellow office dwellers Peixiang and Audrey for teaching me the subtleties of membrane protein biochemistry, and for all the work they did on mitochondrial carriers. I could not forget to thank Nicolas for his lively mood that made these long days assigning ORC1 quite tolerable. Thank you fraulein Weinhaupl and Dr. Fraga for listening to all my existential babble and for becoming one of the best friends I have made in Grenoble. And Katharina, thank you for being a fellow PhD student to whom I was able to look up to in respect of scientific standards.

I am thankful for the whole NMR group, especially to Adrien, Sergi and Bernhard for making all the spectrometers so user-friendly, and I would like to thank Bernhard and Cedric for all the interesting conversations and for their belief in my ability, that was often more important for me than they have realised.

Thanks to the collaborators from University of Salento, Loredana Capobianco and Paola Lunetti for performing liposome transport assays on mitochondrial carriers.

I would like to thank my family, especially my brother Valentas for their support during my PhD. And lastly I would like to thank our beer club, including Joyce, Tomáš, Charles, Marko, Raleb and Quantin, which was instrumental in dealing with frustrations I have encountered during the project.



# Abstract

The use of detergents is often unavoidable in the structural studies of membrane proteins. Dodecylphosphocholine (DPC) is one of the most commonly used detergents for such studies in solution state NMR spectroscopy. The effect of detergent on structure and dynamics remains an important and poorly understood question. In this study we have investigated millisecond dynamics, substrate binding and structural features of three different yeast proteins from mitochondrial carrier family (GGC1, ORC1 and AAC3) in DPC micelles. We have detected millisecond dynamics, which are asymmetrically distributed across the structure. Contrary to previous claims, we show that these dynamics are unrelated to function, as they are not affected by the substitutions which abolish mitochondrial carrier transport in proteoliposomes. Furthermore, we could show that the very well-defined substrate specificity of these proteins in membranes is abolished when they are reconstituted in DPC, questioning their functionality. Structural investigations have revealed that both tertiary and secondary structures of these carriers are perturbed in DPC micelles, with some TM helices showing substantial solvent exposure. We have concluded from these observations that DPC detergent strongly perturbs these, and likely other mitochondrial carriers by rendering them very flexible. Our findings point to a possibly general effect of this detergent on membrane proteins, as we discuss with examples of previously studied membrane proteins.

In the second part we have addressed a fundamental question of protein dynamics: how do proteins move inside crystals? We have investigated  $\mu$ s dynamics in a crystalline ubiquitin to gain the insight on the impact of the crystalline lattice on such motions, using solid-state NMR and  $\mu$ s long MD simulations of explicit crystal arrangements. Interestingly a local dynamic exchange process on a  $\mu$ s time scale is still present in crystals. However, by comparing different crystal forms we establish that the thermodynamics of the exchanging states and their interconversion rate constants are significantly altered by the crystal contacts. Furthermore, we detect overall "rocking" motion of molecules in the crystal, occurring on a tens-of- $\mu$ s time scale, and provide evidence that overall and local motion are coupled. We discuss the implications of  $\mu$ s dynamics on the data quality in X-ray diffraction experiments.



# Abstract in French

Cette thèse est composée de deux parties : la première partie, qui est aussi la plus importante, est dédiée aux études des propriétés structurales et dynamiques des transporteurs mitochondriaux, reconstitués dans des micelles de DPC (n-dodecylphosphocholine). L'intérêt d'effectuer cette étude était double : des études fonctionnelles suggéraient que ces transporteurs mitochondriaux peuplent deux états bien distincts lorsqu'ils effectuent le transport des substrats à travers la membrane mitochondriale. Le passage d'un état à l'autre (ou interconversion) se produit à l'échelle de la milliseconde, ce qui est une gamme de temps idéale pour la caractérisation par dispersion de relaxation de type CPMG (Carr-Purcell-Meiboom-Gill). Une analyse quantitative de ce type de données permet d'obtenir une information précise sur les populations des deux états en échange, la vitesse d'échange (ou interconversion) ainsi que de l'information structurale sur l'état mineur, ce qui est souvent inaccessible par d'autres techniques d'analyse.

Nous avons donc cherché à apporter des détails mécanistiques sur l'interconversion fonctionnelle de trois transporteur mitochondriaux différents, tous présents chez la levure : les transporteurs de l'ornithine (ORC1), du GTP/GDP (GGC1) et de l'ATP/ADP (AAC3). Nous avons effectué des expériences de type CPMG avec des conditions expérimentales très différentes et avons analysé comment l'échange dépend des substrats et ligands connus. Nous avons également étudié comment des mutations, dans les résidus impliqués dans le mécanisme de transport, modifient la flexibilité de ces transporteurs.

Nos données ont démontré que de l'échange à l'échelle de la milliseconde est effectivement observé pour les trois transporteurs étudiés dans ce travail (ORC1, GGC1 et AAC3). De plus, nous avons noté des similarités significatives entre les propriétés d'échange de ces trois protéines. Par exemple, dans toutes les protéines étudiées, l'échange est distribué d'une manière asymétrique dans la structure et l'analyse des données expérimentales donne des paramètres quantitatifs très similaires : nous observons un état mineur avec des populations de 1-2% pour les trois transporteurs, et une vitesse d'interconversion de l'ordre de  $1000-2000 \text{ s}^{-1}$ . Cependant, nous n'avons pas pu détecter d'effet significatif des différentes mutations introduites. Ces mêmes mutations inactivent pourtant les protéines, comme démontré par des tests de transport effectués dans des liposomes. Il en résulte que la dynamique que nous observons n'est très probablement pas corrélée avec le cycle fonctionnel de ces protéines. Elle est au contraire très certainement provoquée par des perturbations structurales induites par le détergent, le DPC.

Ceci relève une autre question intéressante et peu comprise : comment la structure et la fonction d'une protéine membranaire sont influencées par le détergent ? Le DPC est un détergent fréquemment utilisé dans des études par RMN des protéines membranaires. La majeure partie des structures de protéines membranaires en solution a été résolue en utilisant ce même détergent. Il apparaît donc très important de comprendre comment ce détergent peut perturber les protéines.

Afin de comprendre ces perturbations nous avons étudié l'interaction des ces trois transporteurs avec différents composés par des analyses de déplacements chimiques induits lors d'expériences de titrations. Nous avons trouvé que ORC1 reconstitué dans du DPC n'est pas capable d'interagir avec son substrat natif, le 1-ornithine. De l'autre côté, GGC1 pouvait clairement lier le GTP. Les

transporteurs mitochondriaux sont connus pour leurs spécificités envers leurs substrats (par exemple, GGC1 lie le GTP mais pas l'ATP). Nous avons donc voulu savoir si cette spécificité est retenue dans des micelles de DPC. Les résultats de nos expériences RMN ont montré que cette spécificité n'est pas conservée : GGC1 pouvait fixer aussi bien l'ATP que le GTP ainsi que le CATR, un inhibiteur bien caractérisé des transporteurs de l'ATP/ADP, mais pas du GTP/GDP. L'ATP et le CATR n'interagissent normalement pas ou très faiblement avec GGC1 dans des membranes natives. Nous avons fait des observation similaires concernant la spécificité d'AAC3 pour ses substrats : cette protéine reconstituée dans du DPC n'était plus capable de distinguer entre l'ATP, son substrat naturel et le GTP, contrairement à ce qui a été observé dans des bicouches membranaires.

Nous avons par la suite essayé de comprendre quel type de perturbation structurale subissent ces transporteurs pour qu'ils soient encore capables de lier leurs substrats mais en ayant perdu la spécificité. Nous avons analysé les structures tertiaires des transporteurs. Des mutations de résidus dans des régions connues comme étant importantes pour la fonction et/ou compactes semblaient induire seulement des changements locaux (comme détecté par l'analyse des variations des déplacements chimiques), au lieu des propagations à travers une large partie de la structure. Ceci est inattendu puisque ces mutations devraient affecter la structure globale plus sévèrement. L'explication qui nous semble la plus plausible serait que ces transporteurs ne peuvent pas adopter une structure tridimensionnelle très compact dans les micelles de détergent. Nous avons alors examiné la stabilité des structures tertiaires des transporteurs AAC3 et ORC1 par des expériences de dénaturation thermique, observés par RMN. Les déplacements chimiques de deux protéines ont montré des changements linéaires en fonction de la température. Ceci ne correspond pas à ce qui est attendu pour la dénaturation d'une protéine bien repliée, qui est caractérisée par une transition sigmoïdale. Nous avons alors effectué des expériences supplémentaires sur GGC1 en caractérisant la compacité de sa structure tertiaire par des expériences d'échange hydrogène-deutérium par RMN. Comme une partie importante de la protéine est formée par des hélices transmembranaires supposées être localisées dans les micelles, on s'attend à ce que les protons amides des résidus respectifs soient protégés de l'échange avec le solvant aqueux. Les vitesses d'échange de ces protons amides devraient donc être extrêmement lentes. D'une manière complètement inattendue nous avons observé un échange hydrogène-deutérium très rapide pour toutes les régions de la protéine GGC1. Ceci suggère qu'au moins une fraction de la population dans notre échantillon est extrêmement flexible.

Nous avons également étudié les structures secondaires des trois transporteurs reconstitués dans le DPC. En analysant les déplacements chimiques ( $C\alpha$ ,  $C\beta$ , HN, N, C') nous avons obtenu de l'information sur les structures secondaires. Dans les trois transporteurs, les structures secondaires apparaissaient largement perturbées. Dans le cas d'ORC1, les déplacements chimiques observés pour tous les segments transmembranaires n'étaient pas en accord avec la formation d'hélices bien définies. Cette observation peut d'ailleurs expliquer, pourquoi cette protéine ne pouvait pas fixer son substrat. GGC1 et AAC3 montraient des perturbations moins sévère, mais néanmoins substantielles. Par exemple, en cas de GGC1, toute l'hélice transmembranaire 4 apparaît très flexible et désordonnée. Un tel désordre n'est pas attendu pour une région transmembranaire : dans toute structure de protéine membranaire, les régions transmembranaires sont formées soit d'hélices- $\alpha$ , soit des feuillets- $\beta$ . Ceci est certainement dû au fait que dans l'environnement hydrophobe de la membrane, des atomes polaires non-appariés, c'est à dire sans être impliqués dans une liaison hydrogène, sont énergétiquement très défavorables. Nous nous sommes alors demandé si les parties transmembranaires non-ordonnées étaient également exposées au solvant. Nous avons donc étudié l'exposition au solvant des résidus en ajoutant une sonde paramagnétique hydrosoluble, le gadodiamide. La présence de ce composé dans le solvant est supposée affecter fortement les régions qui sont accessibles au solvant, tandis que les hélices transmembranaires, protégées par le détergent, devraient être inaccessible et donc rester protégées. Cependant, aussi bien dans le cas de GGC1 que d'ORC1 nous avons vu que des parties importantes des segments transmembranaires ont montré une exposition au solvant substantielle. Encore plus intéressant, nous avons constaté un très bon accord entre les régions qui montraient les perturbations d'hélices- $\alpha$ les plus importantes avec les régions

avec une importante exposition au solvant. Il est également intéressant à noter que les résidus qui montraient de l'échange à l'échelle de la milliseconde se situent également dans ces mêmes régions. En conclusion, nos observations suggèrent que les transporteurs mitochondriaux possèdent des structures tertiaires perturbées lorsqu'ils sont reconstitués dans des micelles de DPC. Il apparaît qu'il n'y a pas un seul état mal replié, mais plutôt un ensemble de plusieurs conformations très flexibles. Les interactions non-spécifiques avec les substrats que nous avons observées pourraient être simplement dues à des interactions électrostatiques. En effet, ces protéines possèdent un large excès de résidus chargés dans leur cavité ce qui pourrait être suffisant pour attirer de molécules de charge opposée.



# Contents

<b>Acknowledgements</b>	<b>3</b>
<b>Abstract</b>	<b>5</b>
<b>Abstract in French</b>	<b>7</b>
<b>1 Membrane proteins and their environment</b>	<b>15</b>
1.1 Folding of membrane proteins vs soluble proteins . . . . .	16
1.2 Forces acting on proteins in phospholipid bilayers . . . . .	17
1.2.1 Electrostatic interactions and their modulation by bilayer . . . . .	17
1.2.2 Hydrogen bonds . . . . .	19
1.2.3 Salt bridges . . . . .	23
1.2.4 Aromatic interactions . . . . .	24
1.3 Non-electrostatic interactions . . . . .	24
1.3.1 Hydrophobic mismatch . . . . .	25
1.3.2 Effect of membrane elastic forces on proteins . . . . .	26
1.3.3 van der Waals interactions and packing . . . . .	27
1.3.4 Summary of membrane protein interactions . . . . .	30
1.4 Can membrane proteins fold without a complicated folding machinery? . . . . .	31
1.4.1 Single molecule studies demonstrate robust folding of some membrane proteins <i>in vitro</i> . . . . .	32
1.4.2 Refolding of membrane proteins <i>in vitro</i> . . . . .	34
1.4.3 Membrane proteins can be rendered soluble by replacement of their hydrophobic surfaces . . . . .	35
1.4.4 Membrane proteins appear to contain information about their tertiary structures in their sequence . . . . .	35
1.5 Differences between lipid bilayers and detergent micelles . . . . .	36
1.5.1 General features of detergent micelles . . . . .	36
1.5.2 Critical micelle concentration . . . . .	38
1.5.3 Detergents used for NMR and X-ray crystallography . . . . .	38
1.5.4 Effects of SDS on the structures of membrane proteins . . . . .	39
1.5.5 DPC in solution state studies of membrane proteins . . . . .	45
1.5.6 Are membrane proteins intrinsically very flexible, or is the flexibility caused by the detergent? . . . . .	51
<b>2 Mitochondrial carriers</b>	<b>54</b>
2.1 ADP/ATP carrier . . . . .	55
2.1.1 The function of AAC dictates its strong substrate specificity . . . . .	55
2.1.2 Two conformational states of AAC revealed by the studies with inhibitors . . . . .	56
2.1.3 Binding properties of AAC inhibitors . . . . .	57
2.1.4 AAC1 transport properties . . . . .	58
2.1.5 Lipid importance for AAC transport . . . . .	59
2.2 Structural details of mitochondrial carriers . . . . .	60
2.2.1 Sequence based structural information . . . . .	60
2.2.2 X-ray structural information . . . . .	61
2.2.3 Importance of charged residues in the TM domains of AAC . . . . .	62
2.3 Transport mechanism suggested for AAC. . . . .	65
2.4 Other mitochondrial carriers . . . . .	67
2.4.1 GTP/GDP carrier . . . . .	67
2.4.2 Ornithine/citrulline carrier . . . . .	68

2.5	NMR studies of mitochondrial carriers . . . . .	69
2.5.1	Human SCaMC1 - determination of specificity towards ATP·Mg <sup>2+</sup> . . . . .	69
2.5.2	Human SCaMC1 - the role of the N-terminal domain in the activation of the carrier . . . . .	70
2.5.3	GGC1 - conformational flexibility . . . . .	72
2.5.4	Analysis of UCP2 proton and fatty acid transport . . . . .	74
2.5.5	Substrate modulation of millisecond dynamics in AAC3 . . . . .	75
2.6	Motivation to perform NMR studies on mitochondrial carriers . . . . .	77
<b>3</b>	<b>Materials and methods</b> . . . . .	<b>79</b>
3.1	Sample preparation . . . . .	79
3.1.1	Reagents . . . . .	79
3.1.2	Mitochondrial carrier constructs . . . . .	79
3.1.3	Protein expression . . . . .	79
3.1.4	Protein purification ( <i>S. cerevisiae</i> ORC1) . . . . .	80
3.1.5	Protein purification ( <i>S. cerevisiae</i> GGC1) . . . . .	81
3.1.6	Protein purification ( <i>S. cerevisiae</i> AAC3mut) . . . . .	82
3.1.7	Titration of mitochondrial carrier samples with different molecules . . . . .	83
3.1.8	Reconstitution of mitochondrial carriers in liposomes (Performed by L. Capobianco) . . . . .	84
3.2	NMR experiments . . . . .	85
3.2.1	Acquisition and processing of NMR spectra . . . . .	85
3.2.2	Chemical shift perturbation analysis . . . . .	85
3.3	Assignment of ORC1 backbone resonances . . . . .	86
3.3.1	Assignment of methyl groups of GGC1 . . . . .	89
3.3.2	CPMG data collection . . . . .	90
3.3.3	CPMG data analysis . . . . .	91
3.3.4	Jackknife analysis . . . . .	92
<b>4</b>	<b>Results: Dynamics, Interactions and Structure of Mitochondrial Carriers in DPC detergent</b> . . . . .	<b>94</b>
4.1	Three different mitochondrial carriers undergo extensive millisecond dynamics . . . . .	94
4.1.1	GGC1 undergoes millisecond dynamics . . . . .	95
4.1.2	Millisecond dynamics are also present in ORC1 and AAC3 . . . . .	98
4.2	Substrates have very small impact on dynamics . . . . .	99
4.2.1	Non-functional mutants display essentially identical dynamics as WT proteins . . . . .	101
4.2.2	Millisecond dynamics in GGC1 <sup>2P→2R</sup> mutant . . . . .	102
4.2.3	Cardiolipin slows down the exchange in GGC1, but the effect is not biologically relevant . . . . .	104
4.3	Summary of millisecond dynamics in mitochondrial carriers . . . . .	105
4.4	DPC refolded carriers are not specific for their substrates . . . . .	106
4.5	ORC1 in DPC micelles is not able to bind its substrate L-ornithine . . . . .	107
4.6	GGC1 in DPC micelles does not show binding specificity . . . . .	108
4.6.1	GGC1 is able to bind GTP . . . . .	108
4.6.2	DPC reconstituted GGC1 bind ATP in an identical manner as GTP . . . . .	110
4.6.3	GGC1 shows interaction with carboxyatractyloside . . . . .	111
4.6.4	GGC1 <sup>2P→2R</sup> mutant binding to GTP . . . . .	114
4.6.5	Cardiolipin binds to GGC1 non-specifically . . . . .	115
4.6.6	GGC1 displays similar features refolded from the inclusion bodies or expressed by a cell-free system . . . . .	116
4.7	AAC3 <sup>c-salt-bridge</sup> mutant in DPC does not have binding specificity . . . . .	118
4.8	The lack of substrate binding/specificity to the carriers in DPC micelles suggest perturbed tertiary structure . . . . .	120

4.9	Mitochondrial carriers do not appear to have well defined structure in DPC micelles	122
4.9.1	GGC1 in DPC micelles lacks compact, well folded structure	122
4.9.2	AAC3 in DPC micelles lacks compact, well-folded structure	124
4.9.3	Summary of the observations on the tertiary structure of mitochondrial carriers in DPC	128
4.10	Investigation of mitochondrial carrier structure in DPC micelles	129
4.10.1	Secondary structures of mitochondrial carriers	130
4.10.2	Transverse relaxation rate measurements for GGC1	135
4.10.3	Probing solvent exposure by solvent PREs	136
4.11	Supplementary figures	138
<b>5</b>	<b>Discussion</b>	<b>164</b>
5.1	Structure of UCP2 in DPC micelles	164
5.2	Mechanism of proton transport by UCP2	167
5.3	Studies on ATP/Mg-P <sub>i</sub> transporter inhibition by the cytoplasmic Ca <sup>2+</sup> domain	169
5.4	Molecular basis of MgATP Selectivity of the Mitochondrial SCaMC carrier	171
5.5	Summary of the observations on the mitochondrial carriers in DPC	173
5.6	Studies of other membrane proteins in DPC	174
5.6.1	Perspectives	176
	<b>References</b>	<b>178</b>
<b>6</b>	<b>Manuscript: Visualizing the impact of the crystalline environment on slow protein dynamics</b>	<b>194</b>
6.1	Preface	194
6.2	Abstract	195
6.3	Introduction	196
6.4	Results	198
6.4.1	Type-I ↔ type-II β-turn exchange is slowed down by crystal contacts	198
6.4.2	The relative populations of βI/βII states differ in the different crystals	202
6.4.3	Molecules in cubic-PEG-ub undergo tens-of-μs overall rocking motion	204
6.4.4	Identifying intermolecular contacts that alter the conformational equilibrium	205
6.4.5	Coupling between rocking and βI ↔ βII exchange	207
6.4.6	Probing relaxation dispersion effects	210
6.5	Discussion	212
6.6	Methods	213
6.6.1	Sample preparation	213
6.6.2	Solid-state NMR and data analysis	213
6.6.3	MD simulations	214
6.7	Supporting information	217
6.8	References	222
<b>7</b>	<b>Appendix</b>	<b>226</b>
	<b>Cross-Correlated Relaxation of Dipolar Coupling and Chemical-Shift Anisotropy in Magic-Angle Spinning R<sub>1ρ</sub> NMR Measurements: Application to Protein Backbone Dynamics Measurements</b>	<b>226</b>
	<b>Sensitive proton-detected solid-state NMR spectroscopy of large proteins with selective CH<sub>3</sub> labelling: application to the 50S ribosome subunit</b>	<b>236</b>



## 1 Membrane proteins and their environment

Membrane bound compartments are very advantageous for living organisms, as charged and polar molecules have high energetic restrictions in diffusing across the lipid bilayer encapsulating them. Thus, it allows the soluble constituents of a cell to remain at high concentration inside a membrane-limited vesicles, allowing biochemical reactions to occur. Furthermore, different biochemical processes, which otherwise would be incompatible with each other can be segregated to different vesicles inside the cell, and concentration gradients can be formed across the membrane for energy storage. However, as lipid vesicles tightly isolate the cells from their environment, the by-products of chemical reactions could not diffuse away from the vesicle, as well as the precursors could not enter inside. In addition, such a cell would lose all of its communication with the environment and with any neighboring cells. Life has overcome all these limitations by the use of a special class of proteins, which are abundant in the lipid bilayers - the membrane proteins.

A typical lipid (see Figure 1.1), constituting membranes of eukaryotes and bacteria, is composed of a polar head-group, which can be negatively charged (i.e. phosphatidylglycerol (PG), phosphatidylserine (PS)) or neutral (i.e. phosphatidylcholine (PC), phosphatidylethanolamine (PE)) and aliphatic chains 14-18 carbon atoms in length, which can differ in their level of saturation. In aqueous environments present in living cells, aliphatic chains are drawn to each other by the hydrophobic effect to form the most energetically favorable phase of lipid bilayers. Lipid bilayers are composed of two leaflets of lipid molecules arranged in a manner that polar head-groups are always in contact with water molecules on both sides of the bilayer, while the hydrophobic chains in the two leaflets are touching by the ends. Aliphatic chains would still be exposed to the solvent at the edges of the planar bilayers, thus hydrophobic effect further makes the bilayers to close into spherical vesicles.

As mentioned above, lipid bilayers of living organisms contain an abundance of proteins specialized to perform functions such, as signal transduction or transport. In their natural environments membrane proteins are surrounded by hydrophobic lipid alkyl chains. However, it is often too difficult to perform detailed structural investigations on these proteins in their native membranes. Thus detergents, as in this study, are used either in the extraction steps or in the final sample which is characterized. Detergent environment is very different from a bilayer environment, and thus it can result in the structural and dynamic perturbations of the native state of a membrane protein. To understand potential perturbations it is important to understand the process of membrane protein folding and forces which define the native conformation. This chapter is thus dedicated to look into the state-of-the-art understanding of membrane protein interactions (primarily  $\alpha$ -helical membrane proteins are considered) in the native membranes and in detergents.

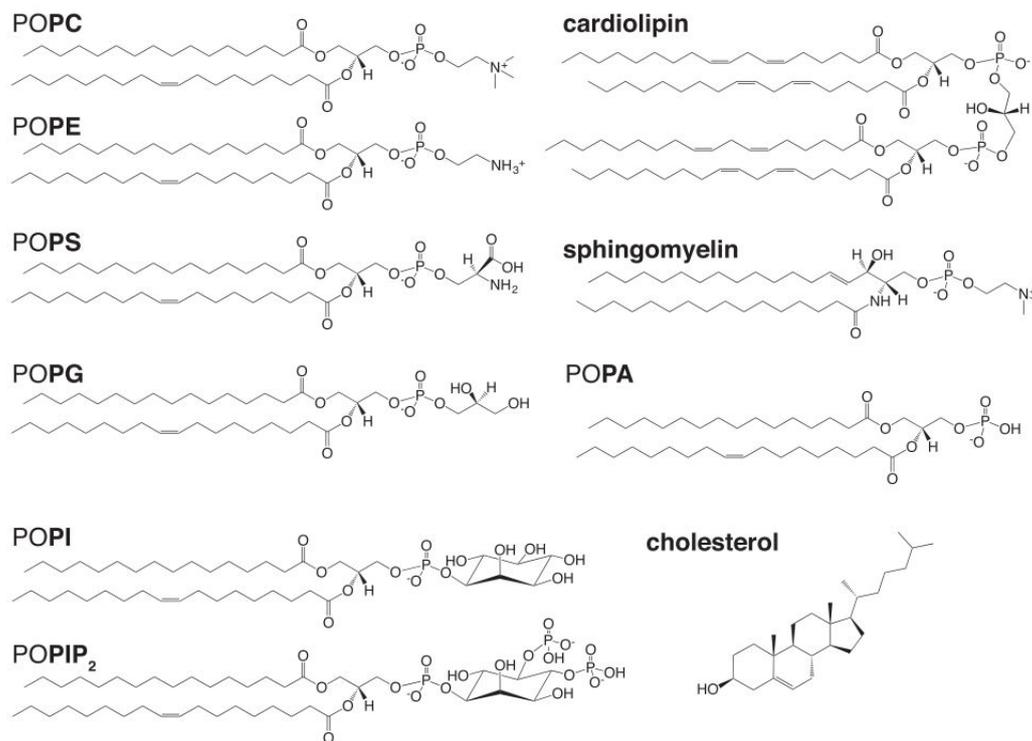


Figure 1.1: Most common cellular lipids. Figure taken from ref. (1)

## 1.1 Folding of membrane proteins vs soluble proteins

Membrane proteins differ fundamentally from the soluble proteins. The understanding of soluble protein folding and stabilizing interactions is substantially more advanced than that of membrane proteins. Hydrophobic effect and hydrogen bonds have been suggested to be the major players in conferring the stability to the folded state of soluble proteins, where they have to overcome large energy costs imposed by the reduction in conformational entropy. The rotational freedom of various bonds in an unfolded protein is much greater as compared to a folded protein, which is estimated to impose 1.5 – 4 kcal/mol per residue energy cost upon protein folding into a compact, native state. Such large increase in the free energy has to be overcome by favorable interactions in a folded state. A plethora of interactions among which the hydrogen bonds, providing -1.1 kcal/mol per bond of free energy on average and hydrophobic effect combined with van der Waals forces, resulting in similar gain of free energy per -CH<sub>2</sub>- group buried, are considered now to be the main driving forces for soluble protein folding. Other effects and interactions stabilizing the folded state of soluble proteins include disulfide bonds (do not exist in TM regions of membrane proteins)<sup>1</sup>, salt-bridges (especially important for thermophilic soluble proteins, as the contribution becomes more substantial at higher temperatures), various interactions with the aromatic groups ( $n \rightarrow \pi^*$ , cation- $\pi$ , aromatic-aromatic) and oligomerization

<sup>1</sup>In some proteins, for example mitochondrial carriers Cys residues may be present inside the solvent exposed cavities.

of proteins are known to contribute free energy to the folded state of a protein (2).

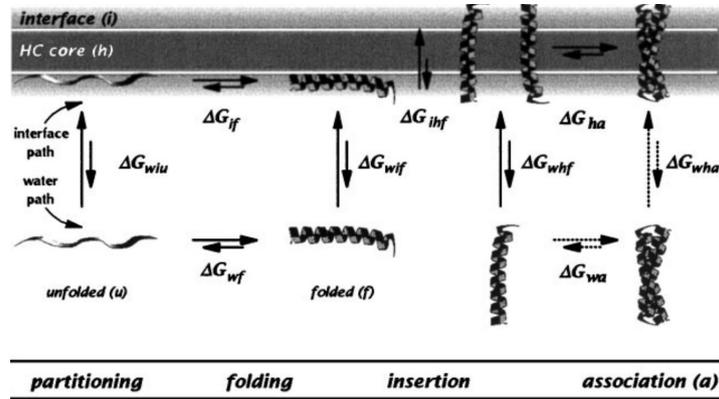


Figure 1.2: A schematic illustration of the complexity of membrane protein folding (aggregation pathway is not depicted). The scheme is designed to represent the association of a dimeric peptide, but remains correct for a polytopic transmembrane protein. Typical experiments, where a membrane protein tertiary structure is disrupted with a denaturant, such as SDS, provide information about the  $\Delta G_{ha}$ , as the transmembrane helices remain largely surrounded by the detergent. However, a polar substitution, for example, might have a beneficial effect on  $\Delta G_{ha}$ , but at the same time change the  $\Delta G_{ihf}$  and thus the preference of a single helix to be excluded from the bilayer, thereby inhibiting the dimerization (or folding) (3).

Membrane proteins exist in a largely anisotropic environment - lipid bilayer, as opposed to soluble proteins, so large differences in their folding and (de)stabilizing interactions must be expected. It is difficult, however, to extract the energetic contributions of different interactions in membrane proteins experimentally, as it is generally not possible to establish a system where membrane protein would be in equilibrium between a folded state in a bilayer and an unfolded state in solution (to understand the complexity of membrane protein folding, see Figure 1.2). This is related to the fact that common denaturants, such as urea or SDS (more detailed overview is provided in the section 1.5.4), are inefficient in unfolding membrane proteins (4). Therefore, most of the thermodynamic studies on interaction strengths analyze the equilibrium of folded and unfolded states in a hydrophobic environment (without considering that these interactions are influenced by the factors which membrane protein is experiencing in aqueous solvent). It is important to note, that in most cases the extent of the unfolding of a membrane protein in denaturant is also not known. The next section will discuss what is known about the stabilizing interactions in membrane proteins and provide a comparison of similar interactions in soluble proteins.

## 1.2 Forces acting on proteins in phospholipid bilayers

### 1.2.1 Electrostatic interactions and their modulation by bilayer

For soluble proteins, the effects coming from the environment are isotropic, as they are surrounded by water from all the sides. For the proteins residing inside the lipid

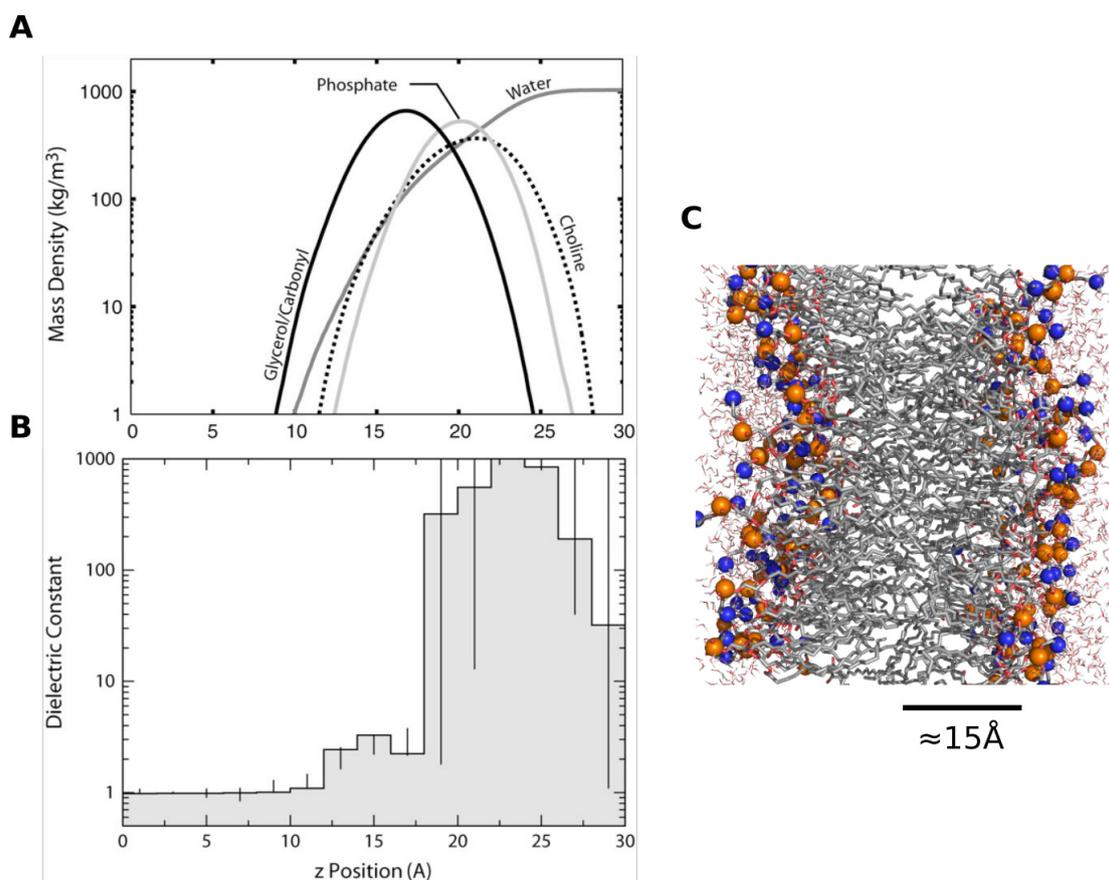


Figure 1.3: A, Distribution of moieties comprising lipids in a POPC bilayer along the bilayer normal (only one leaflet is illustrated), as obtained from MD simulations. 0 corresponds to the center of the bilayer. B, Dielectric profile along the bilayer normal. Vertical lines correspond to confidence limits. As can be seen, alkyl chains possess low dielectric constant, where it starts increasing at around 15 Å due to the presence of carbonyl groups (compare to the A). Large increase is observed at the phosphocholine head-groups which cannot be accurately estimated, however is assumed to be several times larger than that of bulk water. Figure adapted and more information can be found at ref. (5) C, graphical illustration of POPC bilayer. Alkyl chains are represented as gray sticks, phosphorus and nitrogen atoms as orange and blue spheres, respectively, and water molecules as lines. Coordinates of the atoms of MD simulated bilayer taken from the website of Prof. P. Tieleman's group.

bilayer this, however, is not the case. Transmembrane proteins are exposed to polar environment at the surfaces of the bilayer, while the core of the bilayer is highly hydrophobic. Some properties of the bilayer which have so far been identified and believed to be important for membrane protein folding and stability will be presented here.

The hydrophobic part of the bilayer, comprised of aliphatic chains of phospholipids, is around 2.5-3.5 nm in cross-section<sup>2</sup> (see Figure 1.3A, C), while the polar groups

<sup>2</sup>the lipid composition varies depending on the organism, the type of vesicle and even environmental conditions (6).

comprise at least 1 nm on each side of the bilayer (7, 8). Such a composition imposes a gradient of polarity (Figure 1.3B). Dielectric constant in the bulk water solution is  $\approx 60$ -80. In the case of POPC bilayer, however, the polar head-group region is poorly solvated (water concentration is less than 50% as compared to the bulk solution (7) and thus the dielectric constant of this region is much higher, approaching 200 (Figure 1.3B). On the contrary, interfacial and hydrophobic regions of the bilayer have very low dielectric constants of 2-5 (5, 9). It is very important to note that the polarity profile in living organisms is most often not symmetric as opposed to the case of an artificial POPC bilayer, as the lipid composition of the inner and the outer leaflet of lipid bilayer is usually different (10). For example, in the bacterial inner membrane positive residues of membrane proteins are more abundant on the cytoplasmic side of the protein (11). This very likely has influence on the directionality of membrane protein insertion (i.e. which side of the bilayer the top of the protein is facing).

A low dielectric constant in the aliphatic region implies that the polar interactions should be very strong inside the core of the bilayer. Indeed, it was initially believed that membrane proteins are inside-out soluble proteins, having strongly hydrophilic interiors, while their exteriors which are exposed to the lipids should be hydrophobic (12, 13). In such model, hydrophobic effect would drive the folding of a membrane protein in the same way it is affecting folding of soluble proteins. This idea, however, quickly appeared to be inaccurate (12), when bioinformatic analyses of known protein structures have shown that the interiors of transmembrane proteins are as hydrophobic as the ones of soluble proteins. However, their exteriors appear to be even more hydrophobic (8). The lack of highly polar cores of membrane proteins is maybe not overly surprising, as the partitioning between aqueous and membrane phases depends on hydrophobic effect (Figure 1.2), and considering the folding of membrane proteins, the presence of too many polar residues could disfavor membrane protein insertion into phospholipid bilayers (14).

## 1.2.2 Hydrogen bonds

### 1.2.2.1 Evidence that backbone hydrogen bonds are stronger in hydrophobic environment

A low dielectric constant is nevertheless important for the membrane protein folding. First, all of the known structures of membrane proteins display a limited variety of folds: they are either composed of  $\alpha$ -helices or  $\beta$ -barrels in the transmembrane sections. These types of architecture ensure the largest amount of hydrogen bonds formed between peptide groups, thereby decreasing the energy cost of transferring polar moieties to the hydrophobic environment. Furthermore, it has been experimentally measured that transmembrane helices are extremely resistant to heat or chemical denaturation. Computational studies of peptide to alkane transfer estimate that the transfer of an unfolded coil to a hydrophobic environment is  $\approx 4$  kcal/mol per residue less favorable than transferring a hydrogen bonded peptide (15). Second, the hydrogen bonds on the lipid exposed faces of  $\alpha$ -helices (0.198 nm) are shorter than inside protein interior (0.202 nm) or the surfacial hydrogen bonds of soluble proteins (0.208 nm). This implies that the hydrogen bonds on the lipid ex-

posed surface of membrane proteins are stronger. It is also commonly observed that the  $\alpha$ -helices of membrane proteins possess slightly different geometry than soluble proteins, so the exposure of polar peptide bonds to hydrophobic surroundings is reduced (16). This is in agreement with the fact that membrane protein surfaces are even more hydrophobic than their interiors (8, 16) - implying destabilizing effects of polar residue exposure towards the hydrophobic part of the membrane.

#### 1.2.2.2 Polar and charged residues are disfavored on surfaces exposed to lipid tails

Therefore, polar or charged residues are very rarely exposed to the lipidic environment (Figure 1.4) and their appearance on the transmembrane protein surface is often correlated with disease, hinting to perturbations in protein tertiary or quaternary structure (18, 19). Ser or Thr residues, however, are exposed more often, as they can participate in the hydrogen bonding with the amides of the peptide backbone (due to their short side chain) (20). Asn or Gln, on the other hand, are largely disfavored on the protein surface, and the presence of these residues may drive oligomerization of  $\alpha$ -helices (21) or reduce the stability as for bacteriorhodopsin (22). Furthermore, although charged residues such as Lys or Arg can be found at the periphery of hydrophobic regions of the membrane, their positively charged side chains are always directed towards the head-groups of the phospholipids, a phenomenon termed 'snorkeling' (17). In turn, if these residues were to face aliphatic interior of the bilayer, water molecules would penetrate to the very core of bilayer in order to solvate the charged side-chain, as revealed by MD simulations (18).

There are, however, some peculiar observations reported for growth factor receptors<sup>3</sup>. Disease causing mutations, which result in a ligand-independent signaling are related to neutral to charged residue substitutions in the transmembrane domain. It has been experimentally measured, that some of these mutations (Gly380Arg or Ala391Glu in FGFR3 (23) or G770E/F777E in VEGF2 (24)) are able to increase the strength of dimerization (23). The supposed mechanism of this mutant is a hydrogen bond formation between protonated Glu and neighbouring residue (25). It has to be considered, however, that such a protein has two large soluble domains on both sides of the TM domain, and thus the TM domain is not able to be expelled from the bilayer, as it would require the translocation of one of the soluble domains across the bilayer.

#### 1.2.2.3 Although hydrogen bond strength is modulated by the hydrophobicity of the environment, their stabilizing contribution does not appear to be larger for membrane proteins as compared to soluble proteins

Considering the aforementioned observations, it is natural to assume that the hydrogen bonds would play a major role in stabilizing the tertiary structures of  $\alpha$ -helical

---

<sup>3</sup>These receptors are composed of extracellular ligand binding domain, a single pass transmembrane domain and an intracellular kinase domain, and exist as inactive homodimers in native membranes. Ligand binding to the extracellular domain induces conformational changes which activate the intracellular kinase domain.

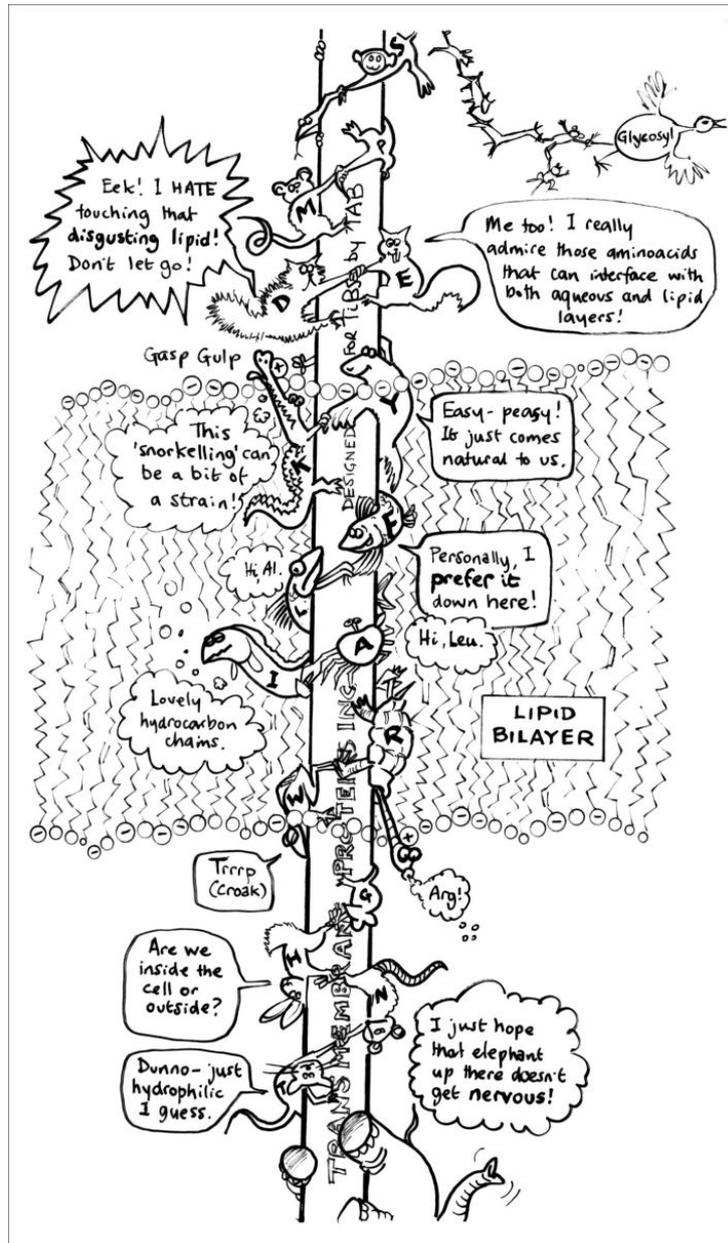


Figure 1.4: The preference of different amino acids in transmembrane proteins in respect to the lipid bilayer. Amino acids are represented as cartoons. The letter on a cartoon corresponds to the amino acid name in a single-letter notation. As can be seen, glycines and polar amino acids are preferentially distributed outside of the bilayer core, where hydrophobic amino acids are prevalent. R and K residues might be exceptional, as their long alkyl chain might partition in the membrane, where the charged group points towards the interface. Aromatic residues, such as Y or W tend to cluster at the membrane interface. (17)

membrane proteins by promoting helix association. This idea is supported by quantum mechanical calculations in vacuum, estimating the strength of these interactions to be  $\approx -6$  kcal/mol, while in aqueous environments, where water molecules can com-

pete with the hydrogen bonding partners, the strength of hydrogen bonds is reduced to  $\approx -1$  kcal/mol (26). Indeed, it was demonstrated by a careful study that in soluble proteins the hydrogen bond strength depends on the polarity of microenvironment. If a bulky, hydrophobic side-chain is introduced close to hydrogen bonding residues, the local dielectric constant and the ability of water molecules to access and compete for the hydrogen bond are reduced, thereby increasing the strength of the bond by up to  $-1.2$  kcal/mol (16). It is therefore surprising, that instead of expected  $-4$  to  $-5$  kcal/mol hydrogen bond energies, experiments indicated that most hydrogen bonds are only modestly stabilizing, providing  $-0.5$  to  $-1$  kcal/mol for the stability of a transmembrane domain, and being comparable to the stability enhancements granted to soluble proteins by hydrogen bonds (27). For example, in an  $\alpha$ -helical rhomboid protease GlpG, out of 26 residues involved in hydrogen bonding, only 3 appeared to have large stabilizing effects ranging from  $-1.8$  to  $-2.9$  kcal/mol (28)<sup>4</sup>. All of these residues appeared to be located in the periphery of GlpG, involved in hydrogen bonding with multiple partners, thereby linking several helices together. Furthermore, DgkA tolerates mutations of all the polar residues to non-polar ones (30). How such low stabilizing effects of hydrogen bonds could be explained then? First of all, the residues involved in hydrogen bonding exist in proteinaceous environments – they are not isolated chemical moieties as often simplified in quantum mechanical calculations. Thus these residues are surrounded by a multitude of polar partners, which can result in increasing dielectric constant and providing additional sites which compete for hydrogen bonding partners. Furthermore, many membrane proteins, such as membrane transporters, might have water accessible cavities which further reduce the strength of hydrogen bonds. Additionally, evolutionary optimization may be very important for the fine-tuning of hydrogen bond strength. In GlpG, none of the hydrogen bonds close to catalytic site appear to be important for protein stability (28). As the function of many membrane proteins require substantial conformational changes, weak hydrogen bonds result in lower transition state energies, allowing such changes to occur. Indeed, it has been shown that hydrogen bond networks are reshuffled in different functional states of membrane proteins (31). Considering the evidence available so far, it seems that the hydrogen bonds in transmembrane proteins are not stronger than in soluble proteins, and although they do provide an important energetical contribution, they are not sufficient to explain membrane protein folding in the absence of hydrophobic effect.

It has to be mentioned that unusual type of hydrogen bonds was proposed, namely a bond between  $\text{C}\alpha\text{H}$  and carbonyl group of peptide (33, 34) (see Figure 1.5). However, the experimental evidence of stabilizing effects of these bonds is conflicting. A single mutation of Thr, which was predicted to be involved in  $\text{C}\alpha\text{H}$  interaction with its side chain, to Ala in bacteriorhodopsin resulted in a more stable protein (35), suggesting that the interaction strength is negligible. However, single  $\text{C}\alpha\text{H} \cdots \text{O}$  interactions might not contribute substantially to the stability of the protein, but

---

<sup>4</sup>However, it is difficult to interpret these thermodynamic parameters, as the unfolding of the protein was followed as the decrease in enzymatic activity (by increasing SDS concentration). Activity of an enzyme may be very sensitive to small local perturbations and it may not report well on the degree of unfolding. The authors have claimed that there is a good correlation between decrease in activity and decrease in Trp fluorescence at 320 nm. However, this interval of fluorescence is sensitive to changes in environment, i.e. exchange of detergent (29)

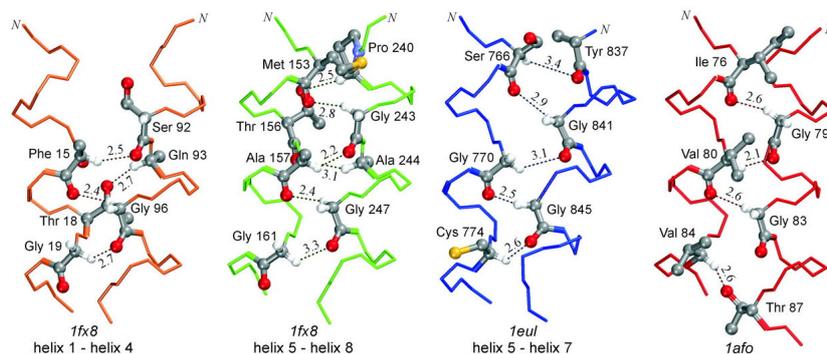


Figure 1.5:  $C\alpha H \cdots O$  hydrogen bonds in closely packed helices orange, green - glycerol facilitator, blue - calcium ATPase, red - GpA (32)

an extended network in appropriately packed helices may have an important contribution (36). Close packing of transmembrane helices could favor these interactions, and membrane protein structure analysis have revealed that the geometry of interacting helices of membrane proteins from experimentally determined structures exhibits high propensity for the formation of  $C\alpha H \cdots O$  hydrogen bonds (33).

### 1.2.3 Salt bridges

Besides hydrogen bonds, other electrostatic interactions, such as charge-charge interactions are expected to be affected by low dielectric environment of the bilayer core. Therefore, one could expect that salt-bridges contribute substantially to the stability of a membrane protein. However, large desolvation related penalties of a charged residue are expected (as mentioned above) when it is transferred to a hydrophobic environment, as well as destabilizing effects related to longer range Coloumbic interactions in apolar surroundings. However, to our knowledge there are no reported experimental investigations on the strength of salt-bridges in  $\alpha$ -helical transmembrane proteins.

Charged residues are uncommon in transmembrane parts of the helices, but substantial amount of experimental evidence hints to their importance in function related conformational transitions in membrane proteins. Alternating salt-bridge interactions that allow a membrane protein to change between distinct conformational states were suggested for many membrane proteins studied: mitochondrial carriers (37), proton-dependent oligonucleotide transporters (POTs) (38), lactose permease (LacY) (39), cystic fibrosis transmembrane conductance regulator (CFTR) (40), bacteriorhodopsin (41), glucose transporter (GLUT1-4) (42). This indicates that salt bridge interactions can be quite strong, as they are able to stabilize particular states of a membrane protein, but at the same time they are dynamic to allow functionally important conformational transitions. These energetic features of salt-bridges might be necessary to allow conformational transitions in membrane proteins, which are dominated by weak, short range interactions (43).

This conclusion can be further supported by the fact that charged ligands can make strong charge-charge interactions with the proteins, thereby greatly stabilizing their

structure. A few examples include ADP/ATP carrier, which is highly stabilized by its inhibitor carboxyatractyloside (44) or TSPO, which is in turn stabilized by a diagnostic ligand PK11195 (45).

#### 1.2.4 Aromatic interactions

##### 1.2.4.1 Aromatic residue partitioning to water-bilayer interface and 'aromatic belts'

Another interesting observation is the presence of large amount of aromatic Tyr and Trp residues at the bilayer-water interface, which form so-called 'aromatic belts' (see Figure 1.6). Aromatic belts are important for defining the orientation of membrane protein in respect to bilayer normal, and they could also provide additional energy for bilayer deformation in the context of hydrophobic mismatch (see below).

It is not well understood why it is energetically favorable for Trp and Tyr residues to localize at the bilayer interface. On one hand, dipole-dipole interactions or hydrogen bonding could play a role: aromatic groups are in contact with aliphatic hydrophobic part of the bilayer, while polar hydroxyl (Tyr) or amide (Trp) groups point toward the head groups of the lipids (17). However, this suggestion was contested by the observation that even without the indole amide Trp rings tend to partition at bilayer interface (46). However, it may be that the aromaticity by itself plays an important role. These aromatic belts are situated 0.45-0.6 nm below the choline head-groups and thus cation- $\pi$  interactions with the positive charge of choline head-group might have a significant role in determining their particular location (43). As the exact forces defining this particular partitioning of aromatic residues are not well understood, they will not be discussed further.

##### 1.2.4.2 Cation- $\pi$ and aromatic-aromatic interactions

Although not investigated systematically in membrane proteins, it is known that cation- $\pi$  interactions can provide moderate stabilizing effects to soluble proteins, which can range between +0.4 to -2.6 kcal/mol and it appears that some of the residues which could participate in such interactions are conserved. Thus, cation- $\pi$  interactions could provide additional stability to some membrane proteins.

Similarly aromatic-aromatic interactions display similar favorable contributions. These interactions were not investigated in detail in membrane proteins, but for soluble proteins they might contribute  $\approx$ -1 kcal/mol energy. It was demonstrated that Phe can increase the stability of GXXXG motif of GpA (see below) and it was suggested that aromatic residues could participate in aromatic-aromatic stacking interactions, if they are separated by less than 0.7 nm, providing additional stability (-1 to -1.4 kcal/mol) per interaction to the protein (43).

### 1.3 Non-electrostatic interactions

So far it did not become obvious what drives the folding of membrane proteins. Polar interactions were believed to be stronger in membrane proteins, thus compensating the lack of hydrophobic effect, but it turned out that their contribution is not larger

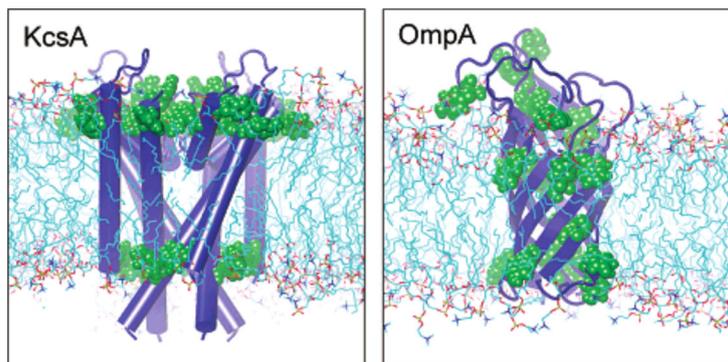


Figure 1.6: An example of the distribution of aromatic residues in an  $\alpha$ -helical (left) and a  $\beta$ -barrel protein. Aromatic residues are displayed in green. Figure taken from ref. (47).

than for soluble proteins.

The further section will delineate some forces acting inside the membrane and their interplay with weak van der Waals interactions, which are believed in fact to play a major contribution in overcoming the conformational entropy.

### 1.3.1 Hydrophobic mismatch

As mentioned before, the surface of a membrane protein traversing the aliphatic part of phospholipid bilayer is exclusively hydrophobic. Exposure of this part to the solvent would result in large energetic penalties. Indeed, there is a large body of experimental evidence on membrane embedded peptides, showing that such peptides adapt (by tilting, for example) to the bilayer width in such a way that the hydrophobic residues become embedded inside bilayer. This effect is called hydrophobic mismatch.

If the hydrophobic part of the peptide matches the thickness of the bilayer, these peptides are oriented in parallel to bilayer normal. However, if the hydrophobic surface of the peptide is longer than the thickness of the bilayer, the peptide starts to tilt along the bilayer normal, or in some cases, the  $\alpha$ -helix can start bending (48, 49). For proteins with several helices, depending on the amount of helical contacts, the helices can start tilting in respect to each other if the inter-helical interactions are sparse, or in opposite case the bilayer thickness adapts to the thickness of hydrophobic surface of the protein (see Figure 1.7). This may favour different conformational states of membrane proteins (50). In any case, the energy of the system is expected to increase (51) and this may effect the oligomerization state of the proteins (50). Therefore, proteins have evolved to minimize hydrophobic mismatch with respect to the membranes in which they are functioning (52).

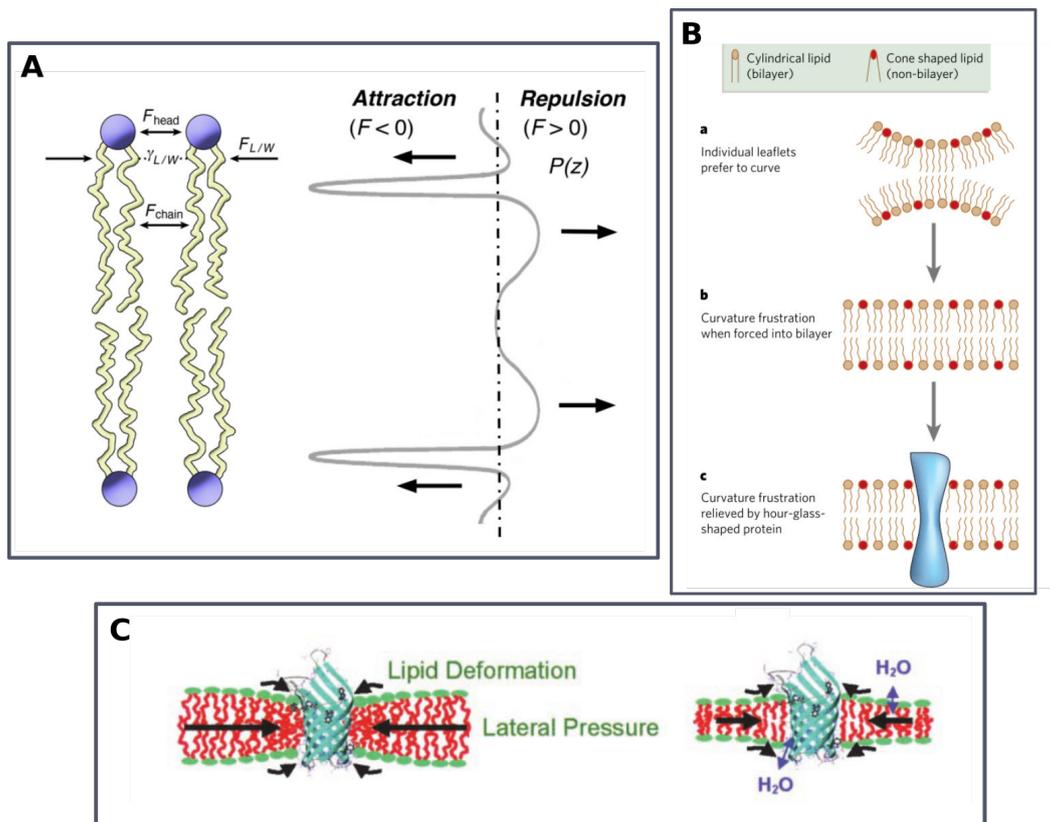


Figure 1.7: The forces determining lateral pressure profile. A, Coloumb force ( $F_{head}$ ) between the head-groups and steric force ( $F_{chain}$ ) between the alkyl chains of a lipid are of repulsive character, while the lipid-water interface region is drawn together by the hydrophobic effect ( $F_{L/W}$ ). These forces create a lateral pressure profile, which is depicted on the right. B, conical lipids (i.e. lipids possessing small head-groups) introduce negative bilayer curvature (53). C, Hydrophobic mismatch creates a force on the bilayer: protein containing shorter hydrophobic surface results in the compaction of the lipids around its cross-section, increasing the repulsive component of the lateral pressure (left); in the opposite case the lipid bilayer expands, reducing the pressure (48).

### 1.3.2 Effect of membrane elastic forces on proteins

#### 1.3.2.1 Lateral pressure

Polar head-groups at the surface of lipid bilayer repel each other because of the Coulomb force. Repulsive force is also present inside the hydrophobic core of the bilayer, because of short range steric interactions. The interface region which is exposed to water molecules, however, generates an attractive force, which is created because of the hydrophobic effect. All these forces in combination generate a lateral pressure profile along the cross-section of the bilayer (Figure 1.7A) (53).

Membrane proteins are believed to be evolved in such a way, that they are able to harness this lateral pressure profile to increase their stability, for example by being thicker at the region where attractive lateral forces are present. Additionally, molecular crowding effects were suggested as being important for the stability of some membrane proteins such as G-protein coupled adenosine A2a receptor (54),

and these effects would be expected to increase inside the hydrophobic part of the bilayer, due to repulsive lateral pressure forces.

### 1.3.2.2 Intrinsic membrane curvature

Another important effect related to the lateral pressure is intrinsic or spontaneous membrane curvature (Figure 1.7B). Phospholipids with small head groups, such as PE, or short chain phospholipids such as diC12PC have a conical shape. A bilayer formed by such lipids have a tendency to curve towards the water interface. However, this curvature is not possible because of the opposing hydrophobic effect, therefore negative spontaneous curvature is stored as an elastic force inside the bilayer. Such force increases the stability of helical proteins (measured on bacteriorhodopsin) (55) and  $\beta$ -barrel proteins (OmpA) (48), as well as the activity of small transporters EmrE and Tbsmr (56), presumably because of larger lateral pressure exerted inside the aliphatic part of the bilayer.

Furthermore, the association of M2 tetramer from single-pass transmembrane helices is favored in the presence of longer chain lipids – the dissociation constant is two orders of magnitude larger in DOPC (18C lipid) as compared to DLPC (12C lipid) (57). However, the negative spontaneous curvature disfavors the folding rate and insertion of proteins inside the bilayer (55). It could be related to the stabilization of some intermediates during protein folding.

### 1.3.2.3 Interplay between elastic bilayer forces

Lateral pressure, hydrophobic mismatch and spontaneous membrane curvature are somehow interlinked. Negative spontaneous curvature modifies the lateral pressure profile so that stronger repulsive forces are exerted inside the hydrophobic core of the bilayer. Hydrophobic mismatch may induce widening or narrowing of the bilayer cross-section, which again can change the lateral pressure profile.

Interestingly, dimerization energy of GpA measured in vesicles derived from mammalian membranes indicate that association energies are lower in the native membranes as compared to artificial bilayers, which are usually used. The decrease in dimerization energy has been ascribed to the presence of other membrane proteins, were non-interacting proteins compete with the interacting partners (58).

### 1.3.3 van der Waals interactions and packing

Lastly, the most important effect, as believed, for membrane protein folding and stability will be discussed. By using computational analysis, it can be shown than large hydrophobic residues (Val, Ile, Leu, Phe) tend to be exposed to the aliphatic chains of the lipids, while small and polar residues form the core of a membrane protein (19). Gly, for example, displays a strong conservation in the TM part of membrane proteins (56% of Glus are conserved) and they are largely buried (20% solvent accessibility) (59). This has two consequences: 1) the buried surface is increased (and maybe the packing density), as two helices can more tightly approach each other, 2)

the entropic cost of burying a residue with a large amount of rotamers is reduced (43).

Increase in buried surface and in the density of packing would favor more extensive van der Waals interactions. The strength of dispersive van der Waals interactions is strongly dependent on distance and is weakened by the competition with the solvent molecules (alkyl chains in membrane protein case) (60). These effects are illustrated by the work on GlpG protease (28). The terminal end of transmembrane helix 4 forms a tightly packed region of a protein, which is very sensitive to substitutions. Mutating larger Leu residues in this region to Ala results in substantial destabilization of the protein by 20 °C. Helices outside the core of the protein are affected less by single substitutions disrupting packing. Multiple mutations are required to reduce the stability substantially. It nevertheless indicates that even weak interactions, when present in multitude, imparts significant stabilization for a membrane protein. Closer packing, therefore, could in principle result in stronger interactions, as dispersion forces depend of the extent of two interacting surfaces. Furthermore, solvent (which in this case are the alkyl chains of a lipid or detergent) accessibility is reduced, allowing the interactions to remain stronger (60).

It is, however, not particularly clear whether membrane proteins are more densely packed than soluble proteins. Analyses of available structures gave mixed results. It is probably not overly surprising, as holes in proteins are important for functional roles, like substrate binding or transport (61). In addition to computational studies, experiments probing destabilization introduced by creating holes in bacteriorhodopsin indicated that the energetic contribution of packing in bacteriorhodopsin and T4 lysozyme are very similar (62). However, it should be noted, that the interior of soluble proteins is already very tightly packed, so the differences between soluble and membrane proteins might be very small and hard to measure.

High pressure stability studies could allow to investigate the packing density of membrane proteins – holes inside the protein structures are largely destabilizing at high pressures. Unfortunately, this type of investigations are very sparse. It has been shown however, that photosynthetic reaction center is extremely resistant to pressure, substantially more than soluble proteins (63).

#### 1.3.3.1 Experimental evidence supporting the importance of packing interactions

The importance of packing interactions for membrane protein stability can be well illustrated by some experimental observations. The transmembrane helix of glycophorin A (GpA) is able to form dimers (in fact these dimers are so strong, that 100% of GpA is dimeric in artificial bilayers (58)) although it does not contain any polar residues, which could participate in, e.g. hydrogen bonding (Figure 1.8). Instead, the peptide contains GXXXG motif, which is now recognized as promoting dimerization of single-pass TM helices and being very important for the stability of polytopic membrane proteins. More than 50% of transmembrane helices in membrane proteins contain (small-XXX-small) sequences (small= Gly, Ala, Ser) (64), otherwise known as GAS motifs and Gly residues are found in 65% of closely interacting helices in crystal structures (59). Similarly, a study on artificial peptides

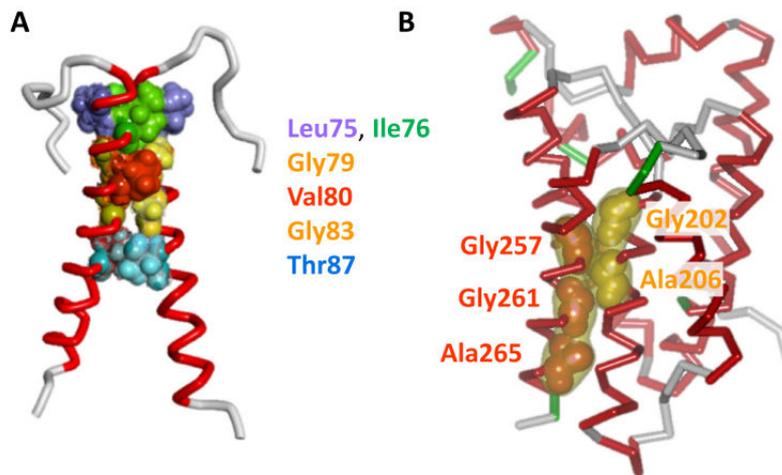


Figure 1.8: Packing interactions in GpA (left) and GlpG (right)

has demonstrated that  $\alpha$ -helix association depends on the size of the side chain. For two residues facing each other in the helix, the association is strongest when the side chain is smallest, i.e. Gly, and decreases progressively by increasing the bulkiness of the side-chain, being the weakest for Ile. Equilibrium constant for a dimer formation reaction being 5 orders of magnitude smaller for the former as compared to Gly (65). Small side-chains of these residues allow helices to pack more closely in respect to each other, increasing the surface of association and thereby allowing the formation of more van der Waals interactions.

In addition, co-variation analysis of trimers of interacting helices shows a narrow range of helix packing arrangements, with 56% of all known helical membrane protein structures falling into 6 well defined helix packing geometries, and the majority of packing interactions are usually occurring at the least flexible regions of the protein (66). This observation further supports that helix packing is an extremely important factor in membrane protein stability.

Experimental observations have revealed that strategically placed GAS motifs result in strong stabilization of polytopic helical membrane proteins as well. Scanning mutagenesis in GlpG revealed two transmembrane glycine residues, which were most detrimental for protein stability (among 151 mutations). One of these glycines (Gly261) involved in 257-GXXXGXXXA-265 motif which links C-terminal helix to the transmembrane helix 4, belonging to the core of the protein, reduces the stability by  $28.1 \pm 0.08$  °C when mutated to valine (28), and destabilization was also observed in force-unfolding studies, when another Gly in the same region was mutated to Ala (67). Similarly, the other important Gly probably is involved in packing interactions which stabilize the N-terminal helix (28).

It seems therefore, that GAS motifs allow the formation of extensive van der Waals interactions, which could further be braced by a multitude of  $\text{C}\alpha\text{H}-\text{O} \cdots \text{C}$  hydrogen bonds (33).

GAS motif, although abundant, is not the only way to ensure close packing and

stability of a membrane protein. Another example of packing interactions was revealed in  $\text{Cl}^-/\text{H}^+$  exchanger from *E. coli*, belonging to the ClC family of transporters. It is a homodimer composed of two monomers interacting via non-polar surface of  $12 \text{ nm}^2$ . There are no polar residues on the surface of the protein, which could drive the dimerization of this protein. However, there is a large shape complementarity, resulting from specific distribution of Leu and Ile residues on the surface. Even a single Ile/Leu to Trp substitution substantially interferes with the dimerization of the protein, with some single residue substitutions effectively converting the protein into monomer. The monomeric protein possesses essentially the same structure as a wild-type dimer (68).

### 1.3.3.2 The interplay between packing interactions and the elastic forces of the bilayer

It could be anticipated that there is some interplay between tight packing in membrane proteins and elastic membrane forces exerted on them. Indeed, studies on GpA transmembrane helix (GpA-TM) revealed that bilayer environment, such as POPC, can stabilize the dimer by 4-5 kcal/mol as compared to DM micelles. Furthermore, packing defects introduced resulted in much stronger destabilization of the GpA-TM dimers in bilayers as compared to DM or C8E5 micelles (69).

### 1.3.3.3 Packing-improving mutations appear to be favoured in thermophilic membrane proteins

An informative way in understanding the most important contributions for stabilizing membrane proteins could be a comparison of thermophilic and mesophilic membrane proteins. However, the amount of available structures of membrane proteins is relatively low and there is little information about their thermostability (for example more stable mesophilic proteins might dominate the membrane protein structure base). Nevertheless, slight differences were observed: thermophilic proteins have on average 10% less polar residues in their transmembrane domains (Asp, Glu, Asn, Gln, Arg, Tyr; which is interesting, as more charged residues are observed in thermostable soluble proteins), while 10% increase is observed for small residues Gly and Ala, and 5% increase for Ile, Pro, Ser, Val. Furthermore, although the packing density of the analyzed proteins did not seem to increase, it appears that 5% more surface area is buried by thermophilic membrane proteins as compared to mesophilic ones. This strengthens the idea that van der Waals interactions play a major role in membrane protein stability. The effect of polar interactions cannot be neglected, however, as decrease in dielectric constant at larger temperatures could favor stronger inter-chain hydrogen bonds (70).

### 1.3.4 Summary of membrane protein interactions

Analysis of different interactions and forces acting on membrane proteins allow us to understand the stability of membrane proteins better. As mentioned in the beginning of this chapter, folding of a soluble protein is disfavoured by a large 4-5 kcal/mol per

residue conformational entropy. We know now that hydrophobic effect results in a tremendous favorable contribution, being one of the main factors stabilizing soluble proteins besides the hydrogen bonds.

For membrane proteins, however, this effect did not appear to be of major importance, at least as long as the protein is inserted into a bilayer. It may, however, help to compensate a part of conformational entropy - membrane insertion of a membrane protein is driven by the hydrophobic effect, and a protein inside a two-dimensional bilayer posses fewer degrees of freedom.

In addition, the hydrogen bonds were believed to be stronger in membrane proteins, thus compensating the lack of hydrophobic effect, but it turned out that their contribution is not larger as for soluble proteins. Many other interactions so far discussed also does not appear to be much stronger than in soluble proteins either.

Why then some membrane proteins (such as diacylglycerol kinase or bacteriorhodopsin) are as stable or even more stable than soluble proteins? It may well be that extensive van der Waals interactions involving small residues, in combination with the lateral forces of the bilayer provide these proteins the edge over the substantial decrease in the conformational entropy upon folding.

#### 1.4 Can membrane proteins fold without a complicated folding machinery?

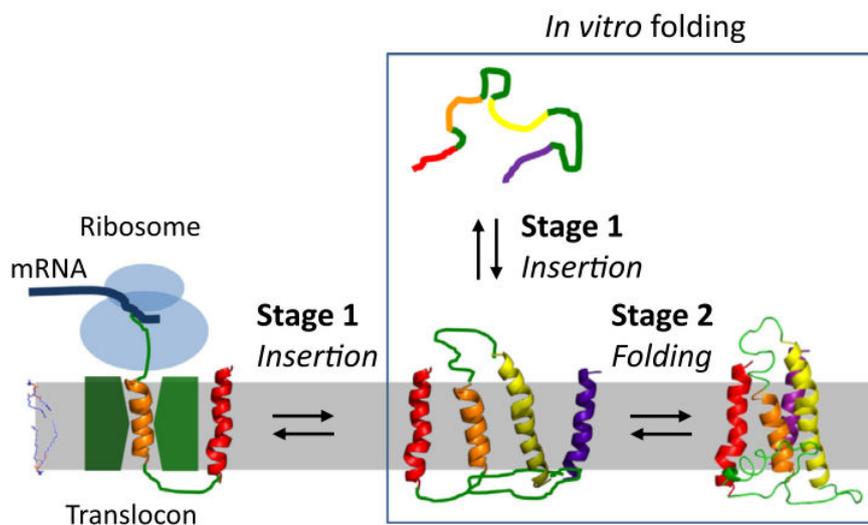


Figure 1.9: Two stages of membrane protein folding. A helical membrane protein might be inserted into the bilayer by a translocon, as occur *in vivo*, or from a detergent solubilized state *in vitro*. Figure taken from ref. (43).

The most simplistic model describes membrane protein folding as two stage process: 1) the hydrophobic chain partitions into a membrane bilayer (secondary structures are believed to be largely formed before this stage) and 2) inserted helices arrange (in the case of  $\alpha$ -helical proteins) into a native state of a membrane protein (Figure 1.9) (71).

The first stage occurs co-translationally *in vivo*, where ribosome docks onto a translocon (ER) or Sec complex (plasma membrane of bacteria) and the nascent helices are inserted in a controlled manner. In the case of membrane proteins residing in outer membrane of gram negative cells, or the inner membrane of mitochondrion or chloroplast, membrane proteins have to cross an aqueous spacing between the two membranes. This problem is solved by soluble membrane protein chaperones present in the space between the two membranes (Skp in case of *E. coli* and Tim and Tic chaperones for mitochondria and chloroplasts respectively). These proteins bind membrane proteins translocating through the first membrane barrier and transfer them past the aqueous environment to systems in the opposing membrane, where transmembrane proteins are inserted in a controlled manner.

The insertion is generally energetically favorable, as membrane protein secondary structures possess strongly hydrophobic segments. Therefore it is believed that the insertion is dictated by transmembrane helix hydrophobicity, and no external energy is used (72). However, less hydrophobic transmembrane helices, such as occurring in membrane transporters (where polar residues line the inner surface of a protein), might require a pre-formed hydrophobic core to be partitioned into lipid bilayer efficiently.

In the second stage, the interactions defining the tertiary structure of membrane proteins are established. In this step protein-protein long distance interactions must overcome protein-solvent (in this case aliphatic chains of lipids) interactions, so that solvent is expelled from the protein interior and a compact structure can be formed. Both inter-helical interactions and forces present in a bilayer define the energy of a native state of a membrane protein.

The previous chapter, hopefully, illustrated that membrane proteins are stabilized by a multitude of weak interactions which are in interplay with the forces exerted by the lipid bilayer. It is known that the bilayer forces has a large contribution to the stability of membrane proteins. However, can these proteins fold in the absence of complicated folding machinery – the translocon, and more importantly, without the lipid bilayer?

#### 1.4.1 Single molecule studies demonstrate robust folding of some membrane proteins *in vitro*

Several studies of membrane proteins by forced unfolding demonstrated that at least some membrane proteins can robustly refold in lipid bilayers without the need of insertion machinery. These single-molecule force microscopy studies revealed interesting details about the folding of  $\alpha$ -helical proteins bacteriorhodopsin (bR) and GlpG as well a  $\beta$ -barrel protein OmpA from *Klebsiella pneumoniae*. As this work is mainly focused on  $\alpha$ -helical proteins, a very interesting study on OmpA performed by Bosshart et al (73) will not be discussed in detail.

In the first study, the C-terminal end of bR embedded inside its native membranes, immobilized on a support, was attached to the AFM cantilever. The cantilever can be slowly moved away from the surface of the membrane, which exerts a force on the protein and results in extraction of the helices from lipid bilayer and their unfolding.

When the cantilever is moved closer to the membrane, the helices were observed to reinsert into the bilayer against a substantial force (tens of pN), while being able to translocate hydrophilic loops across the bilayer.

This study has indicated the formation of intermediate states in some cases, where the helices are inserted into membrane, but they do not form tertiary contacts to the rest of the protein. Nevertheless, in many cases the helices reinserted correctly and the native contacts of bacteriorhodopsin were established, even when N-terminal helices were not inserted into the bilayer (74).

Both stages of membrane protein folding were considered in this study (insertion

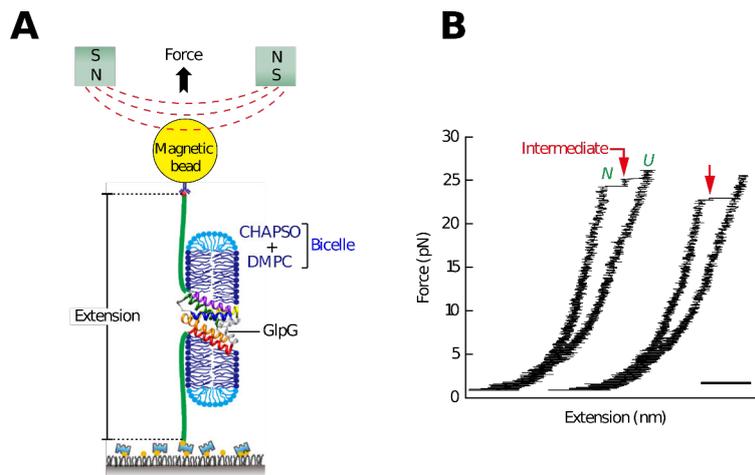


Figure 1.10: An illustration of the force unfolding study done by Min et al., 2015 (67) on GlpG unfolding without extraction of the membrane protein from the bilayer. A, bicelle embedded protein is attached by its termini to DNA linkers. One linker is then immobilized on a support, while the other linker is attached to a magnetic bead. Modulation of the magnetic field allows to exert the force on the protein and unfold it. B, an illustration of unfolding curves observed. Sharp transition around 20-25 pN is observed which was interpreted as corresponding to a cooperative, single stage unfolding of GlpG. In contrast to bR, refolding of GlpG requires reduction of the external force down to 5 pN for the refolding to occur. However, several repeated unfolding cycles appear to give very reproducible outcome.

of secondary structures into the membrane and subsequent formation of a tertiary structure), and due to limited resolution, deconvolution of the contribution from the second stage of folding was not possible.

On the other hand, attempts on GlpG dealt only with the second stage of folding – the protein was embedded inside a bicelle (DMPC/CHAPSO) and was unfolded by applying force to its termini without extraction of the helices from the bilayer (see Figure 1.10A). Interestingly, the unfolding process appeared to be very cooperative, occurring at 25 pN of pulling force and resulting in abrupt 40 nm extension, which corresponds to the length of extended GlpG. The refolding procedure appeared to be very robust, albeit it only occurred at much lower force of 5 pN. Strikingly, even when the lipids were washed away from the protein and subsequently re-introduced, GlpG displayed the same unfolding and refolding behavior (see Figure 1.10B) (67). Interestingly, in neither of the experiments off-pathway metastable states with im-

proper tertiary contacts were observed to be populated to a substantial amount – suggesting that membrane proteins have evolved in a way to avoid these intermediates (67).

#### 1.4.2 Refolding of membrane proteins *in vitro*

Empirical evidence coming from X-ray crystallography suggests that membrane proteins do not crystallize if refolded *in vitro* and even when over-expressed they may obtain non-native structure in lipid bilayers (75). Therefore, the goal is to express protein in an environment which mimics its functional environment most intimately. Biophysical studies however, demonstrated that at least a subset of membrane proteins can be refolded in a variety of membrane mimicking systems, some of them having only a remote resemblance to a lipid bilayer. In 2014 Jean-Luc Popot compiled a list of successfully refolded membrane proteins. The list includes 89 membrane proteins, among which there are 52  $\alpha$ -helical proteins, which include well-studied bR, and GpIG, but also contains a multitude of GPCRs, which are considered to be less stable membrane proteins (4).

The most common procedure for  $\alpha$ -helical protein refolding starts from SDS solubilized or cell-free expressed proteins, which are later inserted into detergent micelles, nanodisks, amphipols or liposomes. In most of the cases, a membrane protein is confronted with a very distinct environment as compared to native bilayers (see discussion about bilayers above). Even the bilayer environment, in which membrane protein is refolded *in vitro* have substantial differences from natural bilayers: the absence of electrochemical gradients, the difference in lipid composition (both head-groups and aliphatic chains), the two leaflets of a bilayer are no longer asymmetric and the thickness of the bilayer might be different.

Detergent micelles introduce further differences, such as heavily distorted lateral pressure profile, absence of membrane elastic forces and maybe different hydration of head-groups. It is noteworthy that GPCRs, which are considered to be sensitive proteins, were successfully refolded in detergent micelles and amphipols – surfactant molecules whose structure is very different from that of lipids or detergents (71).

The yield of refolding is usually not reported in the papers, and it might well be that for some less stable proteins it is not substantial. Although not included in the report of Popot, a multitude of mitochondrial carriers were refolded and reconstituted in lipid bilayers, and this procedure has been a principle way for mitochondrial carrier characterization. However, it can be estimated from these studies, that the amount of refolded functional carriers is very low ((76–80) et cetera.), except in the case of SCaMC1, which reportedly was refolded into liposomes with 100% efficiency (81).

Amphipols do not have well-defined head-groups, as hydrophilic and hydrophobic parts of these polymers are intertwined, thus, besides lacking the properties of both detergent micelles and lipid bilayers, they cannot provide polar interactions of the head-groups. An NMR study of bR in different environments (DDM micelles, DMPC nanodisks and amphipols) indicated a splitting of peaks corresponding to the F loop of bacteriorhodopsin in DDM micelles and nanodisks (which is suggestive of different conformations of the loop). This peak splitting was not present in amphipols,

and more surprisingly it was absent from the solid-state NMR spectrum of bacteriorhodopsin in its native membranes. Nevertheless, other loop regions showed larger perturbations in amphipols, indicating the presence of non-specific interactions with the polar groups of amphipol (82).

#### **1.4.3 Membrane proteins can be rendered soluble by replacement of their hydrophobic surfaces**

Most amazingly, hydrophobic surfaces of a few membrane proteins have been redesigned into polar surfaces, therefore rendering these proteins soluble. Such redesign yielded KcsA which could be expressed in *E. coli* in high amounts, forming tetramers as a wild-type protein. Importantly, the structure of a soluble KcsA variant remained very similar to the wild-type protein and it retained high selectivity for potassium ions (83).

Similar redesign strategies were applied to transmembrane domain of monomerized nicotinic acetylcholine receptor, which yielded biologically feasible NMR structure in solution and was able to bind anesthetics to some degree (84). Human  $\mu$ -opioid receptor was redesigned in a similar way, displaying near wild-type affinity for its antagonist (85). An attempt to solubilize phospholamban proved less successful – although the protein was still able to form pentamers, an NMR analysis displayed badly resolved spectra, which is indicative of molten-globule structure (86).

#### **1.4.4 Membrane proteins appear to contain information about their tertiary structures in their sequence**

These findings support the idea that a large fraction of membrane proteins can be folded *in vitro*, without the need of translocons and appropriate lipid environment – that is, the information determining the tertiary structure of a membrane protein is encoded inside its sequence, as for soluble proteins. It is perhaps not overly surprising. Evolutionary, membranes experience larger variation as compared to proteins embedded in them. While membranes of archaea and bacteria or eukaryotes might have vastly different compositions, the three dimensional structures of proteins from these organisms display remarkable similarity. Furthermore, plants and unicellular organisms are able to radically change membrane composition in response to the environmental changes, and *E. coli* can survive knock-out of the majority of its lipids (87) – although these bacterial strains are not as robust as wild-type *E. coli*. Lastly, in eukaryotic cells membrane proteins are usually expressed and fold in endoplasmic reticulum before they are transferred to target membranes. Thus it seems that many membrane proteins are adapted to some extent to fold and preserve their structure in varying membranes (6).

## 1.5 Differences between lipid bilayers and detergent micelles

### 1.5.1 General features of detergent micelles

The natural environment of membrane proteins is very complex: the native membranes are composed of a large variety of lipids, usually distributed asymmetrically across the bilayer and each of the leaflets, with many different types of membrane proteins co-existing. Although the study of membrane proteins in their native environments would provide the most accurate description of their structure and function, it is not yet experimentally feasible (although some promising results have been achieved by in-cell NMR (88, 89)).

Very few membrane proteins are sufficiently abundant in the membranes to be studied without over-expression and the techniques for structure characterization cannot generally deal with mixtures of different types of proteins. Furthermore, membrane systems in living organisms are generally out of equilibrium and the established methods for functional characterization are usually performed in well-controlled systems under equilibrium.

Therefore, for most practical purposes membrane proteins have to be isolated by purification. Although some recently established methods, such as the utilization of styrene-maleic acid co-polymers (90) allow the isolation of membrane proteins without perturbing the composition of the surrounding lipids, purification almost invariably implies nearly complete removal of the native lipids and neighboring proteins by the use of detergents. Detergent solubilized membrane proteins then can be transferred to artificial bilayers for characterization by solid state NMR or electron microscopy, or monoolein cubic phases for study by X-ray crystallography. Nevertheless, the majority of structural studies are performed in detergent micelles.

Detergents are amphiphilic molecules, somehow related by their structure to lipids

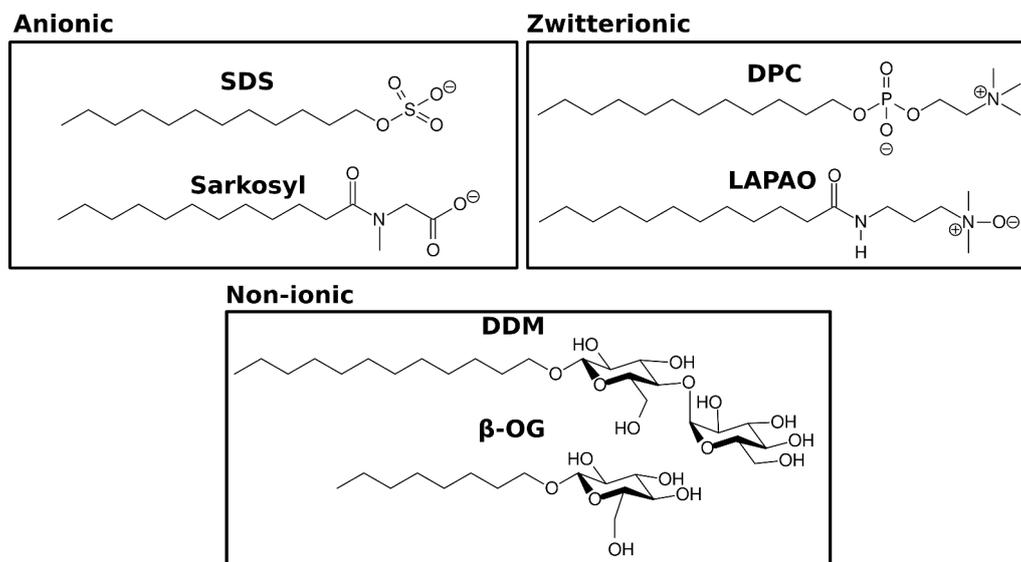


Figure 1.11: Some commonly used detergents

(Figure 1.11). Most commonly used detergents for membrane protein solubilization

have a polar head-group and usually only a single aliphatic chain, although some detergents can have two short chains. Aliphatic chains of detergents are usually shorter than those of lipids. The geometry of detergent molecules (short, single aliphatic chain) dictates that they do not form bilayers as lipid molecules, but rather assemble into spherical or oblate structures called micelles.

Detergent micelles are often depicted in cartoon representations as spheres, where

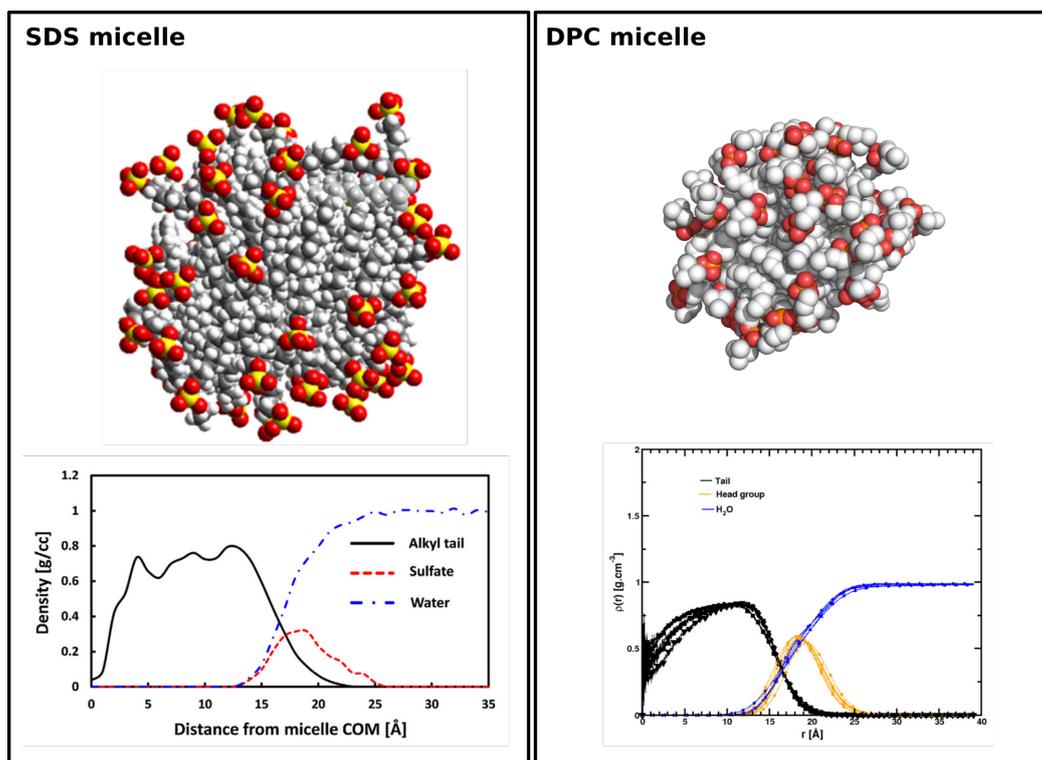


Figure 1.12: SDS micelle (60 molecules, left) and DPC micelle (54 molecules, right) and the corresponding distribution of water molecules, alkyl tails and head-groups (below) as obtained from MD simulations. Alkyl tails are displayed in gray. Note that the size of the micelles in the pictorial representations is incorrect, as two micelles are actually similar in size. SDS figures were adapted from (91). The distribution of moieties in DPC micelle were taken from (92). The atomic coordinates for DPC micelle were obtained from the website of Prof. P. Tieleman.

the aliphatic chains radiate linearly from the center of the micelle, while the micelle-water interface is completely covered with polar head-groups. This type of representation is highly misleading (Figure 1.12). MD simulations revealed, that the aliphatic chains in the SDS micelles are significantly bent, while the polar head-groups cover less than 50% of the surface of the micelle, and DPC micelles acquire similar shape. Thus a large proportion of the micellar solvent-exposed surface is hydrophobic. Furthermore, these micelles are not perfectly spherical and can be slightly elongated. Nevertheless, it appears that water cannot penetrate into the hydrophobic core of the micelle and up to 1-1.2 nm away from the center of a micelle there are no water molecules present (Figure 1.12). The diameter of the hydrophobic core of the SDS micelle is thus comparable to the lipid bilayer, which is 2.5-3.5 nm (91). At large

Table 1: An example of several common detergents used for membrane protein study and their critical micelle concentration (in water)

Detergent type	CMC
<b>Anionic</b>	
SDS (Sodium dodecylsulfate)	8.2 mM
Sarkosyl	13.7 mM
<b>Zwitterionic</b>	
DPC (n-dodecylphosphocholine)	1.5 mM
LAPAO (3-laurylamido-N,N'-Dimethylpropyl amine oxide)	1.56 mM
<b>Non-ionic</b>	
DDM (n-Dodecyl $\beta$ -D-maltoside)	0.15 mM
Brij-35	0.09 mM

concentration ( $\approx 400$  mM DPC), some detergents such as DPC or SDS transit into worm-like, elongated micelles – this is directly related to reduction of water exposed area of hydrophobic chains (93).

### 1.5.2 Critical micelle concentration

The tendency of a detergent molecule to form micelles is described by its critical micelle concentration (cmc) – below this concentration of detergent micelles are not formed, and detergent exist in solution as separate molecules. When this concentration is reached, association into micelles quickly occur. The cmc depends both on the type of head-group and alkyl chain and external factors, such as temperature and salt concentration. Considering detergents with the same length of alkyl chains, the cmc values are determined by the charge of a head group. Anionic detergents typically have high cmc values as compared to non-ionic detergents, while zwitterionic detergents are somewhere in between the two (as measured in water at room temperature; Table 1). This is most likely determined by the destabilization of the micelle due to electrostatic repulsion of the head-groups.

On the other hand, detergents with longer alkyl chains will have lower cmc values: DPC ( $\approx 1.5$  mM), n-tetradecylphosphocholine (FC-14) ( $\approx 0.12$  mM) and n-hexadecylphosphocholine ( $\approx$ FC-16) (0.013 mM), which is dictated by the hydrophobic effect.

Studies of GpA dimers revealed a large variation of dimerization energies in different detergents, but no direct relationship between cmc and dimerization strength was observed (94).

### 1.5.3 Detergents used for NMR and X-ray crystallography

Preferable detergents for both NMR spectroscopy and X-ray crystallography generally have short alkyl chains, ranging between 8 and 12 carbon atoms. Detergents with longer aliphatic chains tend to form larger micelles, which might prevent proteins from crystallizing by limiting protein-protein contacts. Therefore, the most

successful detergents for membrane protein crystallization appeared to be DDM ( $\alpha$ -helical proteins) and  $\beta$ OG ( $\beta$ -barrel proteins) (95).

Solution-state NMR has even more stringent requirements for micelle size, as the detergent molecules add additional mass to the protein-detergent complex, which in turn decreases the tumbling rate of this complex and results in peak broadening and disappearance – which highly complicate NMR data analysis. An average DDM micelle, for example has a size of 70 kDa, while the empty micelles of DPC and SDS are only 20-25 kDa in size. For this, and other reasons, which will be discussed below, DPC is the most commonly used detergent in NMR spectroscopy to solve membrane protein structures. Quite surprisingly, it is extremely rarely used for X-ray crystallography.

#### 1.5.4 Effects of SDS on the structures of membrane proteins

##### 1.5.4.1 Effects of SDS on soluble proteins

SDS is the most studied detergent regarding its effects on proteins. These effects are fairly well understood for soluble proteins. In fact, they are so reproducible, that SDS became a common technique in determining the approximate molecular weight of a protein and it is used in every molecular biology laboratory. SDS, combined with heat, very efficiently unravels soluble protein structures, covering the hydrophobic parts and inducing helical structures in them, while more hydrophilic regions form solvent exposed random-coil linkers. Thereby protein adopts “necklace and bead structure”, which is stable in solution. Furthermore, the amount of SDS bound per protein molecule is very well defined, being 1.5-2g SDS/g protein. As all of the proteins adopt similar structures and bound SDS becomes the main species contributing the charge of the complex, soluble protein electrophoretic motility on acrylamide gels is strictly determined by their molecular size (96).

##### 1.5.4.2 Effects of SDS on membrane proteins

SDS affects membrane proteins in a much more complicated manner (see Figure 1.13). Although SDS is able to unfold membrane proteins and a well-established reverse unfolding system of membrane proteins is employing SDS (29) to disrupt tertiary structure, this detergent is generally not able to unfold  $\alpha$ -helices or  $\beta$ -sheets of membrane proteins. Furthermore, there is evidence that compact unfolded states might persist in the presence of SDS. A compact denatured state was observed for DsbB upon SDS unfolding (97), as well as for bacteriorhodopsin (98–100). Some single-pass helical dimers, such as GpA (101) or PsbF (102) retain their dimeric structure on SDS-PAGE gels. In addition it is well known that membrane protein-SDS complexes may aggregate upon heating (which is probably related to increasing hydrophobic effect at elevated temperatures), indicating the lack of SDS ability to efficiently unfold and solubilize membrane proteins. Furthermore, many outer membrane proteins ( $\beta$ -barrel fold) retain their tertiary structures in SDS below room temperatures (103). Thus even exposed to  $\approx 2\%$ <sup>5</sup> ( $\approx 8$  cmc) of SDS or more, membrane proteins retain

<sup>5</sup>A typical SDS concentration in the sample buffer used for polyacrylamide gel electrophoresis

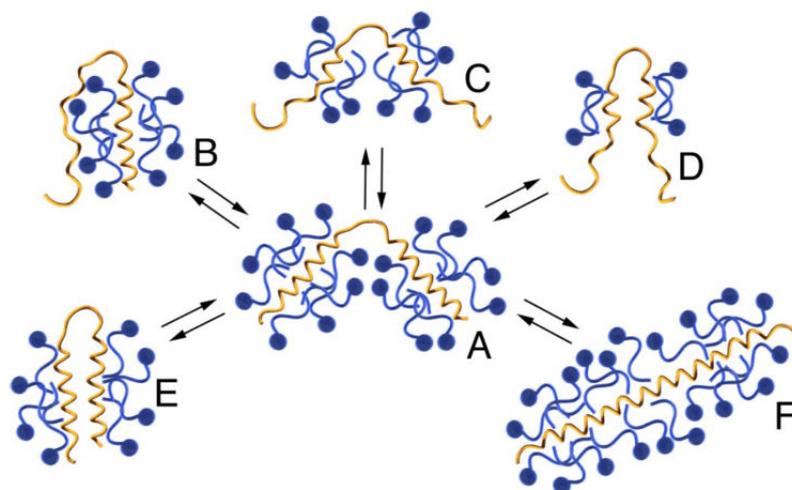


Figure 1.13: A speculative representation of different modes of interaction between the SDS and transmembrane helices of a protein. A typical "necklace and bead" model is represented in A. Helices with lower hydrophobicity might be fully or partially uncoated by SDS, like in B, C or D. More hydrophobic helices might retain their compact structure (E) or result in increased helicity (F) (96).

some degree of their tertiary contacts. This conclusion is corroborated by the different motility of different membrane proteins of the same size on acrylamide gels.

#### 1.5.4.3 Detailed study of SDS effects on a fragment of cystic fibrosis transmembrane conductance regulator (CFTR) fragment

A comprehensive study of the  $\alpha$ -helical protein CFTR fragment differing in hydrophobicity revealed large differences in SDS binding for different mutants, ranging from 3.4 to 10 g SDS/g protein. The amount of bound detergent appeared to be correlated with hydrophobicity and helicity. Even solubilized at large concentrations (74 mM, 10 cmc) of LDS<sup>6</sup> wild-type CFTR fragment migrates at a similar rate as a soluble protein of the same size with less SDS bound, indicating preserved tertiary interactions in CFTR fragment (96). These CFTR fragments were further investigated in their ability to preserve the tertiary structure in the presence of either SDS or DDM. By attaching pyrene fluorophores on the neighboring N- and C- termini of CFTR hairpins, which change their emission wavelength depending of the distance, it was shown that DDM is able to support a compact structure of different CFTR mutants (except a variant where Asp was placed inside the transmembrane region), while in SDS all of the tested variants displayed >1 nm distances between the ends of the fragment, despite preserving relatively large amounts of helicity. The larger distance between the ends of the helices in the presence of SDS is difficult to interpret due to the low resolution provided by the study. It may either suggest that a hairpin forms a continuous helical rod or indicate that the termini

<sup>6</sup>lithium dodecylsulfate, DS salt more soluble at lower temperatures

where pyrene moieties are attached become more flexible. However, some variation in helicity and amount of tertiary structure was observed even in DDM. The variations were dependent on the hydrophobicity of different fragments (104). Much less hydrophobicity-dependent variability on the secondary structures of CFTR dimers was observed when fluorinated detergent, sodium perfluorooctanoate (SPFO) was used, indicating the importance of detergent-protein interactions (most likely van der Waals dispersion interactions were affected) in denaturing membrane proteins (PhD thesis of Vincent G. Nadeau).

#### 1.5.4.4 Evidence that transmembrane hairpins might retain compact structures in SDS

There are several structures supporting the idea that membrane proteins may retain compact states in SDS. A structural model of a helical hairpin comprised of the first two helices of bR have been proposed by combining solution state NMR in 90 mM SDS and a low resolution electron cryo-microscopy map. A compact structure in SDS have been suggested based on correlation time measurements, however, no inter-helical NOE cross-peaks were observed (105). The disappearance of NOE cross-peaks could be caused by the large variation in distance between two interacting nuclei during the NOE transfer delay - or in other words, due to high flexibility.

Analysis of bacterial mercury transporter MerF in SDS micelles (500 mM, 70 cmc) by solution state NMR indicated that some proteins are able to form compact structures even in the presence of large amounts of SDS. In these samples, MerF contains two transmembrane helices (17 and 16 residues in length) arranged in an anti-parallel manner and resistant to hydrogen-deuterium exchange (which in protein-detergent complex indicates that hydrophobic parts are well covered by detergent). N- and C-termini appear to be disordered, as can be judged by NMR relaxation parameters, secondary structure predictions from chemical shifts and residual dipolar couplings (106).

However, the structure solved in DMPC liposomes revealed longer TM helices (23 and 21 residues, respectively), with one of the helices (which was truncated in the solution study) further extending 13 residues into aqueous environment (107) (see Figure 1.14). In addition, the long-range NOE information was not used in the structure calculation of the MerF structure determined in SDS. The absence of long-range NOE might indicate that the helices are moving in respect to each other on a faster time-scale than NOE transfer delay. It could be that the two helices are diffusing in respect to each other while maintaining parallel orientation - RDC data (which was used due to the lack of long-range information from NOEs) does not give information about the faces of interaction between two helices.

#### 1.5.4.5 Effects of SDS on polytopic membrane proteins

As already discussed above for denaturation of helical hairpins of CFTR, bR and MerF, SDS can also denature native states of polytopical membrane proteins. Addi-

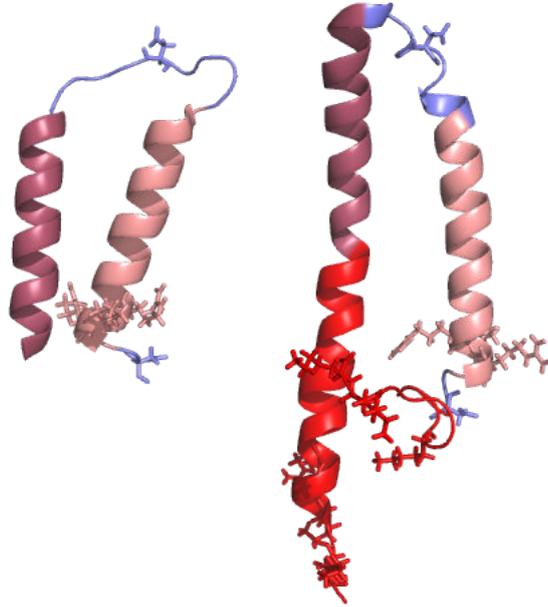


Figure 1.14: NMR structures of MerF in SDS 'micelles' (500 mM or 15% SDS) (left) and DMPC liposomes (right). The construct used for the solution NMR is truncated at the termini (red color, right), as the relaxation properties indicated that they are unfolded. Dark and light pink colors represent residues involved in  $\alpha$ -helical structures in the structure solved in SDS 'micelles'. Blue color indicates residues which were disordered. All charged residues are displayed as sticks.

tion of SDS results in partial unfolding of bR. bR denaturation can be observed by the loss of its retinal cofactor, which is covalently bound to native state of bR via a Schiff-base. In addition, rhomboid protease GlpG completely loses its activity at  $\approx 2\%$  (8 cmc) of SDS. This denaturation, however, is also not complete, as experimental data shows that, as in the case of helical hairpins, bR retains large amounts of secondary and tertiary structure (100).

#### 1.5.4.6 Transmembrane helices containing charged residues tend to partition to the micelle-water interface in SDS

There are no high resolution studies investigating the unfolding of membrane proteins by detergent to date. Electron paramagnetic resonance illuminated the process at the resolution of individual helices. To observe SDS denaturation of bR in purple membranes, each helix was labeled with paramagnetic labels on its C- and N- termini, and thus using DEER experiments the distances between helix ends could be measured in a denatured protein and compared to bR reconstituted in DMPC/CHAPS bicelles. Denatured bR displays a large variation in the lengths of presumed helices indicating large amount of flexibility (rather than one misfolded structure), but the average distance remains similar to folded bR, indicating that the helical nature of the secondary structures is well preserved (108). In an independent study by NMR

it was demonstrated, that the motion of the termini of (eukaryotic) rhodopsin helices shift from millisecond to nanosecond timescale even at  $\approx 30$  mM SDS ( $\approx 5$  cmc) (109).

Larger length distributions of bR helices in an SDS denature state in bR observed by DEER experiments were investigated with MD simulations on isolated helices. It was revealed that the amount of observed helicity is somehow related to the coating of helical fragments by SDS molecules, which in turn was dictated by the hydrophobicity of the helices - very similar observation as made for CTFR fragments experimentally (Figure 1.15). In the MD simulations, helices composed of solely hydrophobic residues were well surrounded by alkyl chains of SDS and retained their helical structure. However, the ones containing positive charges in their transmembrane regions tended to partition towards micelle-water interface, allowing charged residues to make multiple interactions with water molecules or polar groups of the detergent. This type of helix interactions with water may result in kinking of the helices or general unfolding of helical structure (110).

Experimental study of SDS binding to an artificial peptide containing Asn substi-

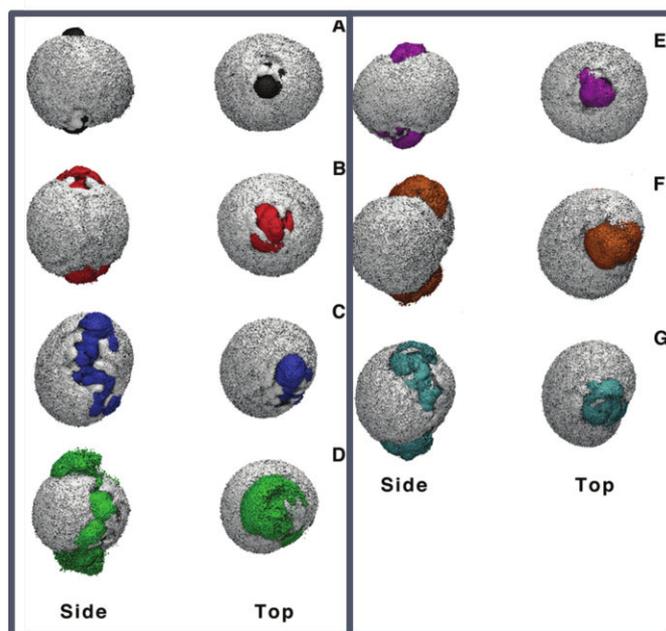


Figure 1.15: Behaviour of individual bR helices in SDS, as observed by MD simulations. Each figure represents populations of a single helix of bR averaged over the whole simulation. White spheres represent SDS micelle. As can be seen, some helices, as in A, B and E are well embedded inside the detergent and are not dynamic. Other helices, especially the one in D are partitioned at the micelle-water interface and display much larger degree of flexibility. The partitioning is likely caused by the presence of charged residues inside the transmembrane helix (110).

tutions at different positions have revealed that peptide partitioning depends on the hydrophobic momentum<sup>7</sup> of the helix – i.e.  $\alpha$ -helices with strong momentum tend to bind less detergent at the more hydrophilic side of the helix. This observation is in

<sup>7</sup>Hydrophobic momentum describes the distribution of polarity along the helix. Helices having more polar residues only on one face possess larger hydrophobic momentum

a good agreement with the MD data (111).

Furthermore, a complementary study of artificial AI10 and AI5<sup>8</sup> peptides (differing in the number of Ile residues, and thus hydrophobicity) indicated that more hydrophobic AI10 peptides are able to form dimers in SDS, while AI5 exist in an equilibrium of dimer and monomer. However, while the addition of Asn in AI10 peptides result in tetramerization, AI5 peptide equilibrium is strongly shifted to monomeric species. This could be explained by the tendency of helices with low hydrophobicity and polar residues to partition towards micelle-water interface, while strongly hydrophobic helices are still able to remain embedded inside hydrophobic environment of SDS (112).

A clear correlation between the hydrophobicity of DsbB lipid-exposed surface and the resistance to SDS unfolding was also observed in DsbB (97).

Interestingly, the property of more hydrophilic regions of TM helices to partition at the micelle-water interface is not restricted to SDS. KdpD and ArcB solubilized in LMPG micelles display similar behaviour. Hydrogen-deuterium exchange data indicate that the helices which contain polar residues (N,Q,Y) or residues of low hydrophobicity (A, G) in KdpD and ArcB display faster H/D exchange rate. The regions exchanging with water are not limited to the ends of the helices, but can be located in the middle of a helix (113).

In addition to secondary structure perturbations in SDS denatured bR, interhelical distances were also determined using DEER experiments (Figure 1.16), displaying large heterogeneity in distance distributions (again, indicating high degree of flexibility) as well as increased average distances by 1.6-2.4 nm, indicating extensive loss of tertiary contacts (108) (in agreement with NMR study on rhodopsin) and FRET study of bR (99)). The increase of inter-helical distance, although large, is too small to account for complete dissociation of the secondary structures and formation of “necklace and bead structure” as in soluble proteins (108).

#### 1.5.4.7 Can SDS (or other detergent) denatured states be very flexible, instead of representing an ensemble of misfolded, rigid structures?

The presence of secondary structures and relatively compact shapes, at least in some membrane proteins in SDS is reminiscent of a situation sometimes observed in soluble proteins. Non-aggregating, soluble proteins with well-preserved secondary structures, as observed, for example, by CD, can sometimes display highly broadened NMR spectra. These proteins are believed to exist in “molten-globule” states, where inter-secondary structure interactions are formed and broken on millisecond time-scale.

SDS denatured proteins generally display very sharp peaks, which would indicate that tertiary interactions are formed on a faster (ns) time-scale. The rate of hydrogen breaking in water, as determined by molecular dynamics simulations, is at least one order of magnitude slower than in a lipid bilayer, which can be explained by the need of concerted breakage of hydrogen bonds among clustered water molecules

---

<sup>8</sup>AI10 corresponds to KKKKKFAIAIAIIAWAIAIIAIAIAI-KKKKK and AI5 to KK-YAAAIAAIAWAIAAIAAAIAA-KKK

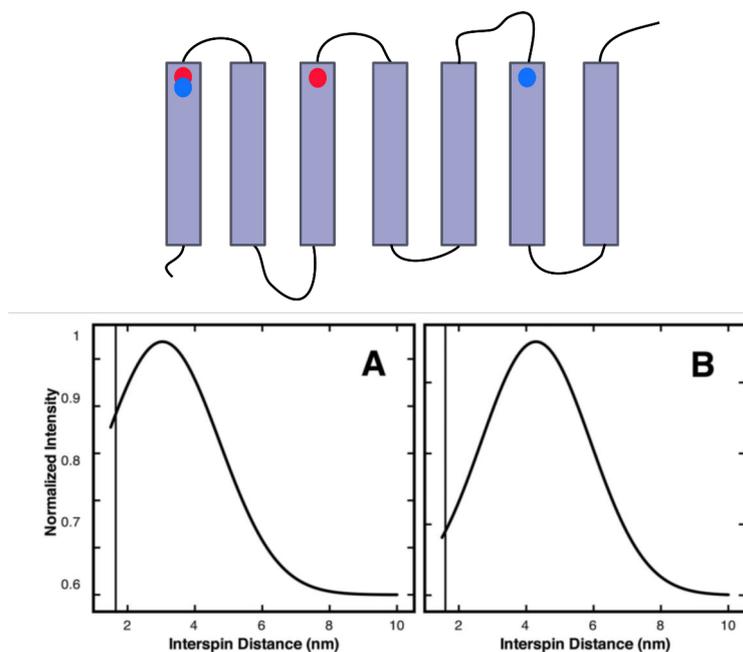


Figure 1.16: An illustration representing the DEER data on inter-helical distances in SDS denatured bR. Purple squares represent the helices of bR. The paramagnetic labels were placed either on helices 1 and 3 (red dots, correspond the the data in A) or 1 and 6 (blue dots, data in B). As it can be seen, the distance distributions are quite broad and the average distance is larger as compared to  $C\beta$ - $C\beta$  distances in the crystal structure (vertical lines). However, the distance distribution increases only by  $\approx 1$  nm in A and B case, suggesting that the helices in the protein remain rather compact (110).

(114). The rate of breaking van der Waals interactions should also be fast, considering the situation where SDS molecules are competing for protein-protein interactions. Static NMR and X-ray structures have made us to believe any kind of denaturation as an introduction of non-native interactions. This is partially true, as discussed above. However, it may be speculated that it is not impossible that membrane protein denaturation in strong detergents results from loss of the important packing interactions, rendering these macromolecules very mobile. A lack of persistent interactions in SDS are supported by some evidence as discussed above, and it seems that refolding of proteins from SDS is feasible, and the most efficient mechanism seems to be a rapid dilution of SDS into a refolding buffer (4). Thus such "molten-globules" are not inconceivable and more experimental data is needed to disprove them (99).

### 1.5.5 DPC in solution state studies of membrane proteins

DPC is by far the most widely applied detergent in the study of membrane proteins by NMR, as can be illustrated by the amount of membrane protein structures solved in different detergents (Table 2). This is obviously related to the high solubilizing capacity of DPC, as it is able to prevent membrane protein aggregation for extended periods of time at elevated temperatures, required to record multidimensional NMR spectra (115). In addition, as compared to other detergents, DPC more often results

Table 2: Membrane protein structures solved by solution state NMR<sup>9</sup>

Membrane mimicking system	Structures solved
C7-DHPC/LMPG	YgaP
C7-DHPC	KvAP-VSD
DHPC	OprG, OprH, PR, OmpA, OmpX, SRII, A-M2, B-M2, AM2-BM2
DPC(FC-12)	Opa60, APLNR-TM, GpA-TM, IR-TM, ErbB3-TM, FGFR3-TM, VEGFR2-TM, <b>STX1A, A-M2 (S31N), p7, TatA, UCP2, PLN, CcdA, OmpG, OmpA, OmpX, PagP, MCP-M13, ErbB2, ZZ, DAP12, VAMP-2, DsbB, TSPO, TSPO (A147T), DgkA, IFITM3 p75-TM, KcsA, mEpoR-TMD, hEpoR-TMD</b>
FC-14	MCU
LDAO	Mistic, $\alpha$ 7-nAChR-TM, VDAC-1
LPPG	hGlyR- $\alpha$ 1-TM
SDS	Nicastrin-a, FXYD1, MerF
DMPC/DHPC bicelles	Env-TM, ErbB4-TM, ErbB2-TM, ErbB1/ErbB2-TM, EphA1-TM

in the best resolved NMR spectra (116).

Many of these structures concern single transmembrane helices or  $\beta$ -barrel proteins. There are, however, several structures solved of polytopic or oligomeric membrane proteins (Table 2, bold). TSPO structure was solved in the presence of strong

<sup>9</sup>YgaP – Inner membrane protein YgaP, *E. coli* (117) | KvAP-VSD – Voltage-sensing domain of voltage-gated potassium channel, *A. pernix*, (118) | OprG – Outer membrane protein G, *P. aeruginosa*, (119) (DHPC) | OprH – outer membrane protein H, *P. aeruginosa*, (120) (DHPC) | PR – proteorhodopsin, marine gamma proteobacterium, (121) | OmpA – outer membrane protein A, *E. coli*, (122) (DHPC); (123) | OmpX – outer membrane protein X, *E. coli*, (124) (DHPC); (125) (DPC) | SRII – sensory rhodopsin II, *N. pharaonis*, (126) | A-M2 – M2 proton channel from influenza A, (127) | B-M2 – M2 proton channel from influenza B, (128) | AM2-BM2 – chimeric A-M2-BM2 channel from influenza, (129) | Opa60 – Opacity protein opa60, *N. gonorrhoeae*, (130) | APLNR-TM – transmembrane domain of Apelin receptor, (131) | GpA-TM – transmembrane domain of glycophorin A, (132); (133) | IR-TM – transmembrane domain of insulin receptor, *H. sapiens*, (134) | ErbB3-TM – transmembrane domain of receptor tyrosine-protein kinase ErbB-3, *H. sapiens*, (135) | FGFR3-TM – transmembrane domain of fibroblast growth factor receptor 3, *H. sapiens*, (136) (DPC:SDS mixture 9:1) | VEGFR2-TM – transmembrane domain of transmembrane domain of vascular endothelial growth factor receptor 2, *H. sapiens*, (24) | STX1A – Syntaxin 1A, *R. norvegicus*, (137) | A-M2(S31N) – M2 proton channel from influenza A with Ser to Asn substitution, (138) | p7 – p7 hexamer channel, Hepatitis C virus, (139) | tatA – twin-arginine translocase, *E. coli*, (140) | UCP2 – uncoupling protein 2, *M. musculus*, (141) | PLN – phospholamban pentamer, *H. sapiens*, (142) (DPC); *H. sapiens*, (143) (DPC); (144) (DPC) | CcdA – cytochrome c biogenesis protein, (145) | OmpG – outer membrane protein G, *E. coli*, (146) | PagP – lipid a palmitoyltransferase PagP, *E. coli*, (147) | MCP-M13 – major coat protein of M13 phage, (148) | ErbB2 – transmembrane domain of receptor tyrosine-protein kinase erbB-2, (149) (DPC); (150) (DMPC/DHPC) | ZZ – transmembrane zeta-zeta dimer of the TCR-CD3 complex, *H. sapiens*, (151) | DAP12 – TYRO protein kinase-binding protein, *H. sapiens*, (152) | VAMP-2 – Synaptobrevin/vesicle-associated membrane protein 2, *R. norvegicus*, (153) | DsbB – disulfide bond formation protein B, *E. coli*, (154) | TSPO – translocator protein, *M. musculus*, (45) | DgkA – diacylglycerol kinase alpha, *E. coli*, (155) | IFITM3 – interferon inducible transmembrane protein, *H. sapiens*, (156) | p75 – BDNF/NT-3 growth factors receptor transmembrane domain, *R. norvegicus*, (157) | MCU – mitochondrial calcium uniporter – pore forming region, *C. elegans*, (158) | Mistic – protein mistic, *B. subtilis*, (159) |  $\alpha$ 7-nAChR-TM – Neuronal acetylcholine receptor subunit alpha-7 transmembrane domain, *H. sapiens*, (160) | VDAC-1 – voltage-dependent anion-selective channel protein 1, *H. sapiens*, (161) | Nicastrin-a –  $\gamma$ -secretase nicastrin-a transmembrane domain, *H. sapiens*, (162) | FXYD1 – Phospholemman, *H. sapiens*, T(163) | Env-TM – envelope glycoprotein gp160 transmembrane domain, HIV1, (164) | ErbB4-TM – receptor tyrosine-protein kinase erbB-4 transmembrane segment, *H. sapiens*, (165) | ErbB1/ErbB2-TM – ErbB1/ErbB2 transmembrane segment heterodimer, *H. sapiens*, (166) | EphA1-TM – ephrin type-A receptor 1 transmembrane domain, (167) | KcsA – pH-gate potassium channel KcsA (168)

inhibitor, PK11195. Without this inhibitor poor quality of the spectra indicated extensive millisecond dynamics (see Figure 5.7A). DsbB would only display a good quality spectrum when a disulfide bond linking the TM helices 2 and 4 were formed, otherwise the spectrum displayed very poor resolution (see Figure 5.7B). The quality of DsbB spectrum would further be improved by addition of co-factor ubiquinone, which presumably plays an important role in the stabilization of the periplasmic loops. M2 structure was stabilized with inhibitors, rimantidine analogs. The KcsA structure was solved in the presence of a strong inhibitor charybdotoxin. In all these cases inter-helical NOEs were observed (methyl NOEs, as for such big proteins extensive deuteration is mandatory, which limits side-chain NOE measurements), indicating compact and stable conformations. DgkA and UCP2 do not report using methyl NOE data. Phospholamban pentamer showed inter-helical NOEs although it was not stabilized by inhibitors, while p7 did not display any long range NOEs (139).

DgkA spectra in DPC (155), for example, displays peak dispersion barely larger than 2 ppm. However, a spectrum of sensory rhodopsin II in LHPC (and bicelles) (126) has a much larger peak dispersion, in agreement with a good packing of TM helices. Observation of loosely packed TM helices were also reported for KdpD structure solved in lyso-PG, where the average inter-helical distance was  $\approx 0.94$  nm (113).

Several of these structures are associated with some controversy, as they differ substantially from structures determined using other methods. The errors observed could arise either because the detergent directly perturbs the native structure of a protein, or calculation errors are introduced because of limited amount of experimental restraints, improper force fields or bias (by introducing hydrogen-bond restraints in the calculations, for example).

### **Phospholamban**

Comparison of phospholamban solution structure in DPC with a hybrid solid-solution state NMR structure indicated two major differences: amphiphatic extramembraneous helices in solution state structure were pointing almost parallel to the transmembrane helices, which was most likely related to issues related to structure calculation. Additionally the transmembrane helices were considerably bent towards exterior of the bundle, as opposed to hybrid structure. This may result from the averaging of residual dipolar couplings because of dynamics (see Figure 1.17A) (169).

### **Diacylglycerol kinase**

In the case of DgkA, domain swapped architecture of a trimer was observed, while X-ray structures indicated a regular trimer (see Figure 1.17B and inserts). In addition, transmembrane helices are bent outwards (this feature is only observed in detergent micelles in general), resulting in a poor packing of transmembrane helices and exposure of polar backbone atoms to hydrophobic environment. Outward bending of the helices might indicate the partitioning of transmembrane helix to the interface of micelle: polar residues present in the TM helix would favor such partitioning, and the hydrogen bonds would lengthen due to increased dielectric constant (as discussed in the section about hydrogen bonds and SDS denaturation of proteins). Lengthening

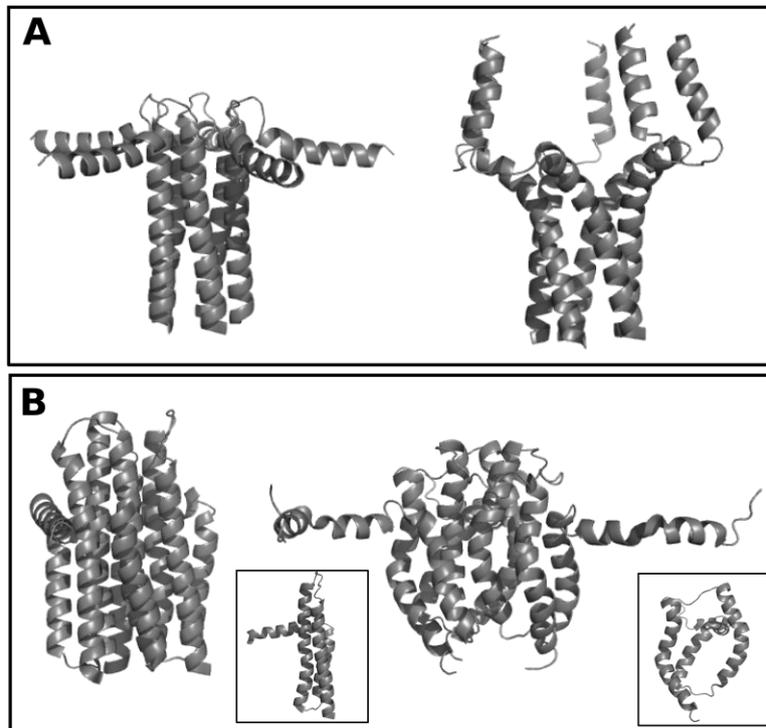


Figure 1.17: Structures of membrane protein structures solved by different methods. A, structure of phospholamban pentamer solved by hybrid solid-solution state NMR approach (left, PDB:2KYV) and solution state NMR in DPC (right, PDB:1ZLL). B, structure of DgkA solved in DPC micelles (right, PDB:2KDC) and in monoolein LCP (left, PDB:4UP6). Besides obvious perturbations in the NMR structure, DgkA displays domain-swapped architecture. Inserts display DgkA monomers in both cases to show the helical structures and the packing interactions.

of hydrogen bonds in the middle of the helix in turn would result in the outward curvature.

Furthermore, some terminal polar residues are exposed to the hydrophobic part of bilayer (9). As discussed in the introduction, such features are highly incompatible with a correctly folded membrane protein structure. Although it seems appealing to interpret the data as direct evidence of structural perturbations induced by the micelle, the lack of restraints used in structure calculation could have resulted in some of these non-native features. Structural perturbations are, however, evident in DAGK in DPC micelles, as the  $K_M$  is increased (9).

### p7

Interesting feature observed in micelles, the domain swapping was also observed in p7 channel (139). As 18 Arg side chains are exposed on TM surface of the protein, the structure was suggested to be non-native (170).

Some peculiar observations related to fos-choline detergents were observed for the soluble cytoplasmic domain of GlpG - namely their ability to catalyze domain swapping at micromolar concentrations (171). Further investigations are necessary to understand whether domain-swapped structures in the aforementioned examples were induced by DPC.

### **Uncoupling protein 2**

UCP2 also display some features that are not compatible with properly folded membrane protein - loose packing interactions and exposure of polar residues on TM surface. Investigation of the structure by MD simulations revealed, that the cavity of UCP2 in DPC is not closed on one side as in AAC structures and instead traverses the whole transmembrane section of the protein. This is inconsistent with the function of the protein. When the structure is allowed to relax in a lipid bilayer, it quickly collapses removing the packing defects(172). As in the case of DAGK, it is difficult to deconvolute effects arising from improper structure calculation from those related to detergent induced structure perturbation.

### **Translocator protein**

Some differences in TSPO structure in DPC by NMR and X-ray structures in LCP and DDM micelles indicate possible perturbations in NMR structure. However, it has to be mentioned that the structures of TSPO solved by two different crystallography groups display substantial differences: the dimerization interface observed is different (which might reflect a native ability of TSPO to form wider range of oligomers) and one of the helices is tilted along the axis perpendicular to bilayer normal. This might be an artifact, or it might indicate some native flexibility of TSPO. However, in the NMR structure some rotation of the helices along the axis parallel to the bilayer normal results in a partial exposure of Asn residue. As the crystal structures concerns bacterial homologues and the NMR structure is of a mammalian protein, there might be some functional importance in the exposure of Asn, as the mammalian TSPO seems to be able to form higher order oligomers. However, the same rotation of the helices result in the deposition of conserved residues outside of the binding cavity, while in the X-ray structures they are clustered inside (45, 173–175).

#### **1.5.5.1 DPC in X-ray crystallography**

In contrast to NMR structures, there is a scarce amount of X-ray structures reported in DPC: a structures of a helical  $Zn^{2+}/H^+$  exchanger YiiP (176), a  $\beta$ -barrel protein OmpF, which crystallized as a contaminant in a trial of histidine kinase receptor KdpD (177) (the crystal structure of KdpD was not reported up to this date, however NMR backbone structure was solved, albeit in lysoPG (113)), and a mitochondrial outer membrane anchored monoamine oxidase (178). Furthermore, there is information in the literature that diffracting crystals were obtained of a thermostabilized GPCR  $\beta$ 1-adrenergic receptor inhibited with cyanopindolol (179) in a related detergent fos-choline 10. It is not clear, however, whether the low amount of structures represent high rate of failure, or the DPC and other fos-choline detergents

are very rarely used in X-ray crystallography.

#### 1.5.5.2 Experimental evidence of the denaturing ability of DPC

Although SDS is widely accepted as a strong detergent able to denature protein structures (as discussed already), DPC is regarded as a much milder detergent. The mechanism of protein solubilization of these two detergents might be different, as one of them is anionic, while the other - zwitterionic. However, there are clear indications that at least in some cases DPC can behave in similar way as SDS. For example a study of solubilization of different GPCRs indicated, that both SDS and DPC are able to extract much larger amount of these proteins from eukaryotic membranes as compared to DDM or digitonin. The conclusion was made that both DPC and SDS are able to extract misfolded protein from the membrane, which is not extracted by weaker detergents (75). The authors did not discuss the denaturing capacity of these detergents. However, SDS purified GPCRs appeared to lose pre-bound ligands, and in many cases the amount of remaining ligand was non-existent or much lower in DPC samples as compared to DDM or digitonin samples (which is surprising given larger quantity of total protein extracted). Thus is also likely that both DPC and SDS are able to further denature even the properly folded fraction of GPCRs.

There are other examples illustrating the denaturing capacity of DPC. Studies on water-soluble Dengue fever protease NS2BFL-NS3pro construct have revealed that fos-choline detergents stood out among 80 detergents tested by their ability to disrupt the activity of this enzyme (180), although these detergents were among the ones showing the largest yields of solubilized protein during purification (181), as observed in the aforementioned GPCR study.

Extensive detergent screening of over 100 detergents for purification of multidrug resistance protein 3 (MDR3) and bile salt export pump (BSEP) revealed, that only fos-choline detergents were able to extract both proteins from plasma membrane in a quantitative manner. BSEP displayed no ATPase activity in fos-choline detergents, however some binding to ATP could be recovered after the exchange of detergent to DDM. MDR3 purified in fos-16 displayed no ATP hydrolysis activity, however, it could be induced by addition of 2-5 mM DPPC or DOPC (182).

Mouse erythropoietin receptor (mEpoR) have been shown to exist as a dimer in the native membranes (183), however, it is not able to dimerize in DPC micelles (184). Human variant of EpoR possessing high sequence identity to mouse receptor, however, is able to form dimers and trimers in DPC micelles, as have been shown by cross-linking studies (184).

Interestingly some reports could be found where DPC solubilized proteins are more active than the DDM solubilized ones. For example  $\beta$ -barrel protein OmpLA refolds better in DPC (20% refolded) than in DDM (5% refolded) and displays the specific activity larger than any other non-ionic detergent tested ( $\approx 0.05 \mu\text{mol}/(\text{min} \cdot \text{mg})$  vs  $10 \mu\text{mol}/(\text{min} \cdot \text{mg})$ ) (185).

It has indeed been shown that DPC is able to perturb even such robust proteins as bR. A comparison of different fos-choline detergents in their ability to retain the native structure of bacteriorhodopsin revealed that these detergents are much milder

than SDS, as the observed absorbance coincided with the absorbance of retinal bound to bacterioopsin (whereas in the SDS samples free retinal was dominant). However, increasing the DPC concentration above 80 mM resulted in a rapid loss of retinal, indicating denaturation (186). Even larger amounts of DPC are used for recording the data for NMR structure calculations of most membrane proteins, some of which are clearly less stable than bR.

There is also some evidence hinting that DPC might be able to induce helical structure in membrane proteins, at least to some extent. Comparison of predicted TMD regions to those observed in the NMR structures of single-pass helices indicate, that on average the helices are 20% longer in DPC micelles as compared to the predictions. In some cases these helices extend up to 16 residues C-terminally (VEGFR2-TMD) beyond predicted TMD regions. It could be argued, of course, that TMD predictions are not accurate.

The best way to identify detergent-related perturbations is to compare the protein structures obtained from several different environments. However, there are not enough structures of the same TMD solved in DPC micelles and bicelles to allow direct comparison. Out of several available examples, in two instances helices are longer in DPC (by 1 residue), but in one case they are shorter (by 7 residues). However, in the case of ErbB3-TMD the helical segment propagates into C-terminal His-tag (64), indicating the ability of DPC to induce non-native helicity in proteins. Peculiar example of a single-pass TMD of VEGF2 have been reported. Introduction of two Glu residues on one face of a transmembrane region resulted in a different dimerization interface, where each helix turns by  $\approx 180^\circ$  (as compared to the wild-type structure) and brings four Glu residues in close proximity. These NMR structures have been solved at pH of 4.5 ( $pK_a$  of Glu side chain 4.25) and the authors argued that the structure is physiologically relevant, as they did not observe changes in C $\gamma$ Glu residues in the pH range 4.5-7.0. Thereby it appears that TM regions with high hydrophobic momentum are not transferred to the surface of DPC micelle, as it would have been expected in SDS micelle (discussed above), at least in this case (24). However, it might also indicate large hydrophobicity of the TM fragment, as in the case of artificial AI10 peptides (discussed above).

#### **1.5.6 Are membrane proteins intrinsically very flexible, or is the flexibility caused by the detergent?**

Low dispersion (2 ppm on proton dimension) (see Figure 5.7) of many detergent reconstituted membrane protein spectra and missing inter-helical NOE cross-peaks prompted the idea that these proteins are inherently flexible or even disordered in the membrane (see ref. (187) as an example). On the other hand, it is known that these proteins need close packing of the helices to maintain their native structures. It is thus tempting to speculate that DPC might also attack protein-protein interactions, affect the packing and render the proteins less compact and more dynamic, as might be in the case for SDS. More investigations are necessary to answer this question.

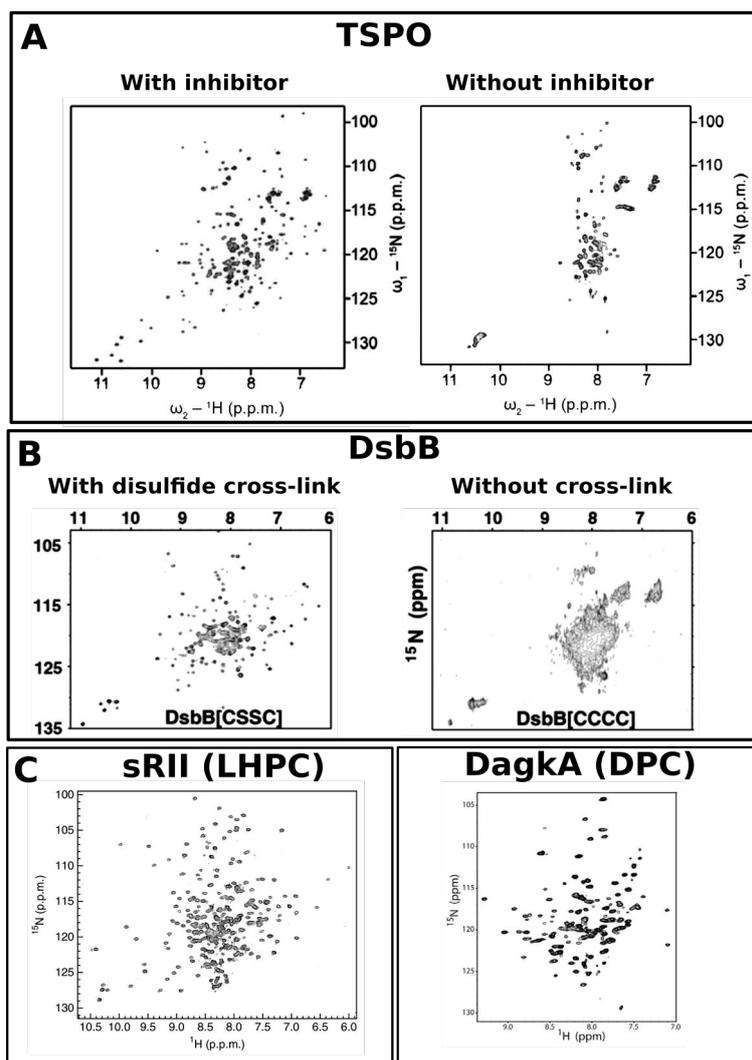


Figure 1.18: Helical transmembrane proteins not necessarily show a narrow peak dispersion. A, TSPO with (left) and without (right) PK11195 in DPC micelles; middle - DsbB with (left) and without (right) a stabilizing disulfide bridge in DPC micelles; C, spectrum of sensory rhodopsin in LHPC micelles (left) and a spectrum of DgkA in DPC micelles (right). Both proteins are of similar size.



## 2 Mitochondrial carriers

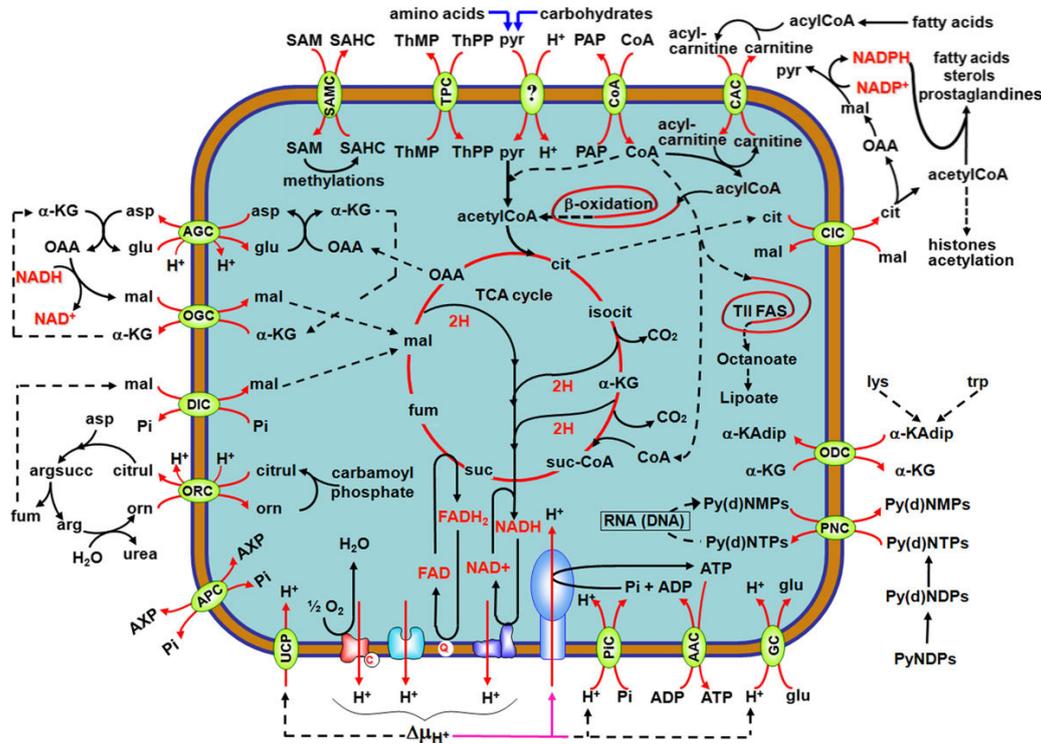


Figure 2.1: An illustration of various metabolic processes occurring inside of the mitochondrion and the role of mitochondrial carrier in connecting these processes to the cytosol. Figure taken from ref. (188)

Many important metabolic processes in eukaryotic cells are dependent on biochemical reactions occurring both in the cytosol and mitochondrial matrix. It is therefore necessary to shuttle the intermediates of these reactions efficiently between the two compartments. Although the outer mitochondrial membrane is easily permeable to solutes smaller than a few kDa due to the presence of porins, such as VDAC, the diffusion of these metabolites is hampered by the tight inner membrane. Therefore, eukaryotic organisms are dependent on a large family of membrane proteins, called mitochondrial carriers (MCs), which, by transporting solutes across the membrane, provide a linkage between biochemical reactions comprising the same metabolic pathway and occurring in different compartments (see Figure 2.1). Although MCs appear to be structurally similar, the metabolites transported by different MCs can be of various chemical characters: keto acids, amino acids, nucleotides, inorganic ions and co-factors, such as NADH. Decades of research on MCs resulted in many insights regarding their function: the substrate specificities of many carriers have been elucidated, and the general understanding of their functional mechanism have been achieved. Out of the large MC family, ADP/ATP transporter (AAC) is understood the best. As it is believed that other transporters share a lot of similarities with AAC regarding their structure and function, the current understanding of AAC will

be overviewed in the following section.

## 2.1 ADP/ATP carrier

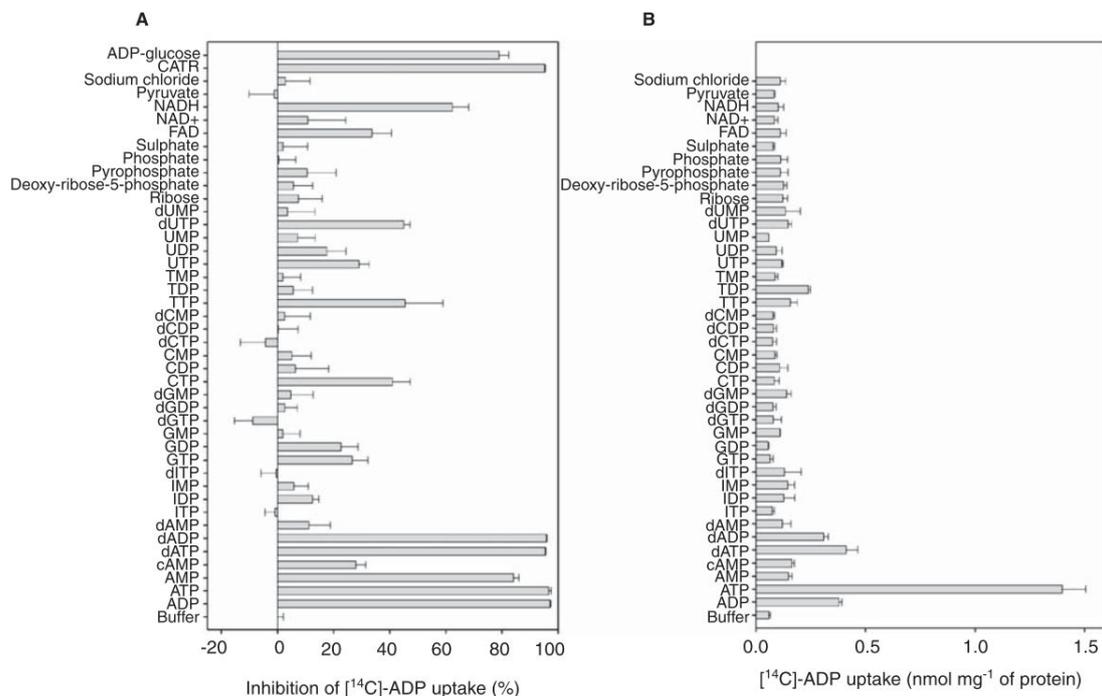


Figure 2.2: Selectivity of human AAC1 reconstituted in proteoliposomes, without membrane potential. A, ability to bind compounds, which have some similarity to the transported substrate, determined by their ability to inhibit transport at 1000-fold larger concentrations. B, transport efficiency of these compounds, determined as efflux of an indicated compound out of a liposome (189)

### 2.1.1 The function of AAC dictates its strong substrate specificity

ADP/ATP transporter (AAC) is the best characterized transporter in functional and structural terms among all of the mitochondrial carriers. Being the most abundant protein in mitochondrial inner membranes, AAC allowed the characterization of the transport properties of its different homologues both in the native and artificial membranes, well before the dawn of molecular cloning. In contrast to AAC, the majority of other mitochondrial carriers are expressed in small amounts and many of them are tissue specific (188). Furthermore, AAC is the only mitochondrial transporter for which strong inhibitors are available. This fact facilitated the resolution of three structures of different homologues (bovine heart AAC1 (btAAC1) (44) and yeast (*S. cerevisiae*) AAC2 (yAAC2) and AAC3 (yAAC3) (190)), as well as helped in understanding the transport mechanism. Therefore, AAC stands as a model mitochondrial carrier, according to which the properties of other MCs are understood. A high abundance of AAC in mitochondrial membranes is dictated by its crucial role in aerobic respiration. ATP synthase produces ATP inside the mitochondrial matrix,

but the majority of it is consumed to fuel the processes occurring in the cytosol or in other compartments. AAC therefore functions to export large amounts of ATP out of mitochondrial matrix, while importing ADP inside, to be further recycled back into ATP by ATP synthase. Non-specific binding of other solutes to AAC could result in a competitive inhibition thereby diminishing ATP turnover. Even more importantly, a large transport capacity of AAC could easily perturb metabolic pathways by upsetting the balance of various metabolites present in the mitochondrial matrix. Indeed, AAC is extremely specific as determined for bhAAC1 in mitochondria: while it can readily transport both ATP and ADP, a very similar molecule, AMP is not transported and does not display substantial binding capacity. Other similar molecules, such as GTP or ITP are also not transported by AAC and do not bind or bind very weakly to the protein. Similar observations are reported for human AAC1 regarding its transport properties in both bacterial membranes and proteoliposomes (189) (Figure 2.2).

It is also interesting, that although ATP and ADP generally exist as  $Mg^{2+}$  complexes in mitochondria, AAC can only transport free nucleotides and addition of  $Mg^{2+}$  inhibits the function. The  $K_D$  for both ADP and ATP on the cytoplasmic side of the carrier is similar, being in the range of 5-12  $\mu M$  (191).

### 2.1.2 Two conformational states of AAC revealed by the studies with inhibitors

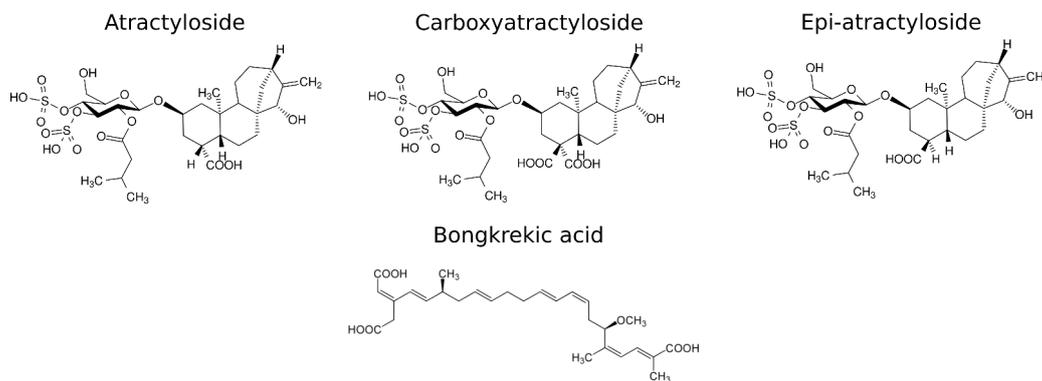


Figure 2.3: Two classes of strong AAC inhibitors. Atractylosides are only able to inhibit AAC in the *c*-state, while bongkreikic acid and its analogues inhibits the carrier in the *m*-state.

Biochemical experiments revealed a wealth of information about the carrier mechanism even before the first X-ray structure was solved. Two classes of inhibitors, atractylosides (ATRs) and bongkreikic acid (BKA) were instrumental for these achievements (Figure 2.3). Atractyloside and a related compound - carboxyatractyloside (CATR), both of which are present in Mediterranean gum thistle, display a very strong inhibition of AAC when added to samples of mitochondria. It is interesting that other known nucleotide carriers of MCF family (GTP/GDP, dNDP/dNTP, Mg-ATP/ $P_i$  and peroxisomal AMP/ADP/ATP carriers) are not inhibited by atractylosides (192).

Later, the crystal structure of bovine heart AAC in complex with CATR at 2.2Å res-

olution enlightened the mechanism of this inhibitor. CATR was found to be bound inside the cavity of AAC situated in the state that is open towards cytoplasm<sup>10</sup> (c-state), blocking the access for ADP or ATP, and at the same time stabilizing this conformation of the transporter (44).

However, if the mitochondrial inner membrane is inverted, for example by sonication, thus placing the inner leaflet outside and exposing the matrix side of AAC, CATR loses the ability to inhibit AAC (193).

Another powerful inhibitor, BK comes into play. It can only bind to AAC which was not prior stabilized by ATR inhibitors, and the full binding requires catalytic amounts of AdNs. This has two very important implications:

- First, it suggests that besides a state which is open to the cytoplasmic side of the membrane there is another state of mitochondrial carriers, which is open to the matrix side. This conclusion is further strengthened by the fact that in all AAC structures of the c-state the cavity is tightly closed by several salt-bridges towards the matrix side, and the bound substrates could not reach mitochondrial matrix without a major change in the conformation of AAC (44, 190).
- Second, the two conformations - one facing to the cytosol (c-state) and the one facing towards the matrix (m-state) are stable and no interconversion occurs in the absence of the substrates.

Furthermore, these two conformations appear to be also stable in mild detergents, such as DDM (194). However, it appears that different detergents are able to stabilize different conformations of AAC: in LAPAO and DDM (194), for example, c-state is preferred, while in CHAPS more AAC is in m-state (195).

### 2.1.3 Binding properties of AAC inhibitors

$K_D$  determined in mitochondria is 1-10 nM for CATR for the c-state of btAAC, while ATR affinity is slightly lower (196, 197). CATR displays fully uncompetitive inhibition - it binds to AAC so strongly, that even large excess of ADP does not result in the displacement of CATR from the binding site of AAC in mitochondrial membranes.

ATR is a weaker inhibitor and can be removed from the active site by excess of ADP. Different AAC homologues, however, respond differently to CATR. For example the binding constant is slightly lower for yeast AACs (198).

The interactions defining the binding strength of these inhibitors are not completely understood. CATR has an additional negatively charged carboxylate group as compared to ATR, and considering highly positively charged binding cavity of AAC, the charge interaction might be invoked to explain the binding. However, epi-ATR, which has differently positioned carboxylate group (see Figure 2.3) binds AAC almost as strongly as CATR, despite lacking an additional charge (199). Furthermore, some ATR analogues display strong binding affinity, even when their sulfate groups are replaced with hydrophobic moieties (198).

---

<sup>10</sup>inter mitochondrial space, to be precise

BK has vastly different structure from ATR inhibitors - while the later are generally hydrophilic, BK is composed of long, polyunsaturated aliphatic chain, rendering it quite hydrophobic. Nevertheless, it is a very strong inhibitor of the m-state of AAC - with a  $K_D$  of  $\approx 20-60$  nM. It is however, weaker than CATR, as it can dissociate from AAC, thus resulting in quasi-competitive inhibition in the presence of ADP.

The structural details of BKA binding to AAC are not understood, as there is no available structure of the m-state. Considering the ability of ADP to compete for the binding with BKA, it seems that at least part of the inhibitor occupies the binding cavity. A part of BKA containing two carboxylate groups is a likely candidate to bind inside the cavity upon being deprotonated. However, isomerization of the two carboxyl groups from trans to cis orientation results in only a modest decrease ( $K_D \approx 80-600$  nM) of binding affinity (193), implying the importance of aliphatic chain.

Detergent-solubilized AAC displays large decrease in CATR affinity - it was observed that bovine heart AAC in LAPAO has a  $K_D$  of  $\approx 300$  nM (44), while the affinity of yeast AAC3 in DPC is reduced to an extreme value of  $150 \mu\text{M}$  (200). Disolution of AAC in SDS abolishes CATR binding completely (201). AAC reconstituted in proteoliposomes display similar binding affinity for inhibitors as in the mitochondrial membranes (202), thereby indicating that the lipid environment is necessary for proper stabilization of the native state.

BKA binding is also reduced in detergent-reconstituted AAC, decreasing to  $1.5 \mu\text{M}$  in LAPAO (202). ADP and ATP binding is perturbed as well in the presence of this detergent, being estimated as  $>100 \mu\text{M}$  (202).

Nevertheless, these observations show that even detergent-reconstituted AAC has access to both of its postulated conformations, as both of the inhibitors are still able to bind. However, it has been observed that even in mild detergents such as DDM, an AAC from a thermostable fungus is associated with larger degree of "floppiness", as CATR fully binds to this AAC in the absence of ADP - which is not the case for membrane-embedded carriers<sup>11</sup> (194).

#### 2.1.4 AAC1 transport properties

AAC is thought to have a common binding site, which is located approximately mid-way along the bilayer, and is accessible from both m- and c-states of the carrier (203). Therefore, both ADP and ATP can bind and be transported from both matrix and cytosolic sides of AAC<sup>12</sup>. However, one transport half-cycle yields a stable conformation which is now open to the other side of the bilayer, and in the absence of a substantial amount of substrates on both sides of the bilayer, AAC cannot interconvert between different conformations and thus the transport does not proceed

---

<sup>11</sup>CATR is only able to bind to the c-state of the carrier. Detergent-solubilized carriers might exist in either c- or m-states. In the study of King et al., for example, BKA cannot bind to a fraction of AAC, indicating that some of it is not in the m-state. However, CATR in the same study was shown to bind to all AAC molecules. This can only be explained if we consider that CATR itself might induce conformational change in the carrier (which is not observed in AAC in mitochondrial membranes).

<sup>12</sup>In proteoliposome transport assays, for example, it does not matter if the liposomes are prefilled with ADP or ATP, and which substrate is outside.

(204). Depending on the partial concentrations of both ADP and ATP on both sides of the bilayer, different transport modes can be populated to different extents, determining the net amount of ADP and ATP transported. These transport modes include  $\text{ADP}_{\text{internal}}\text{-ADP}_{\text{external}}$ ,  $\text{ADP}_{\text{internal}}\text{-ATP}_{\text{external}}$ ,  $\text{ATP}_{\text{internal}}\text{-ADP}_{\text{external}}$  and  $\text{ATP}_{\text{internal}}\text{-ATP}_{\text{external}}$  exchange (198).

In addition to the differences in substrate concentration, these transport modes, and thus net transport, is strongly affected by the electric potential of the membrane - up to 30 times more ATP can be exported in the case of equal partial concentrations of substrates on both sides of the membrane and a membrane potential of +180 mV in AAC proteoliposomes (205). This influence is not dictated by the fact that the binding specificity for ADP or ATP is affected by the membrane potential, but rather because of the charge differences in transported substrates: ADP is transported as  $\text{ADP}^{2-}$  and ATP as  $\text{ATP}^{3-}$  moieties<sup>13</sup>. Therefore, under large negative electrostatic potential the  $\text{ATP}^{3-}$  has a higher free energy inside the mitochondrion, where negative charge density is larger than outside, and export of  $\text{ATP}^{3-}$  is preferred (over  $\text{ADP}^{2-}$ ). Although AAC appears to be a strict antiporter driven by the electrostatic potential of the membrane, some other MCs may utilize different modes of transport, such as uniport (carnitine/acylcarnitine carrier) or symport of the substrate with a proton (glutamate carrier). However, the majority of carriers are antiporters - although instead of electrogenic mechanism they may operate by electroneutral mechanism, utilizing co-transport of protons as the energy source (188) or just the concentration gradient.

ATP appears to be translocated more slowly ( $160 \text{ s}^{-1}$ ) than ADP ( $400 \text{ s}^{-1}$ ) (204) in AAC reconstituted in lipid vesicles (where no directional insertion of AAC is retained (206)). This difference might be caused by either a stronger binding of ATP due to its more negative charge and stabilization of a ground state conformation, or increase in the energy of the transition state because of the larger size of ATP.

### 2.1.5 Lipid importance for AAC transport

Studies of reconstituted AAC revealed the importance of lipids, especially cardiolipin, on its transport activity. Cardiolipin molecules were revealed to be bound to AAC in all the crystal structures solved. Furthermore, in NMR studies of Triton X-100 solubilized AAC, cardiolipin signal could only be detected after addition of SDS, indicating both the ability of SDS to denature AAC and strong binding of cardiolipin<sup>14</sup>(207). Measurement of kinetics could not detect the dissociation of a bound CL on an hour time-scale (208). The binding to CL appears to be mediated mainly by its head-group, as the modifications of the aliphatic chain (such as hydrogenation of its double bond) does not appear to interfere with the binding (208).

CL appears to be indispensable for the transport activity in refolded AAC<sup>15</sup> (210) and is also acting as an activator of the transport. The activating function might

---

<sup>13</sup>At physiological pH.

<sup>14</sup>AAC bound cardiolipin would not be visible as it would possess the same correlation time as AAC, thereby broadening its signal below detection limit

<sup>15</sup>However in the complete absence of cardiolipin, yeast are still able to grow on non-fermentable carbon source; a structurally similar lipid PG, which has not been depleted in these yeast, could have partially substituted CL (209).

not necessarily be ascribed to the specific binding of CL, because some other lipids, which are uncommon in the inner membrane, also activate transport function. Replacement of PC in proteoliposomes by PE, for example, resulted in a 7-fold increase in the transport activity of AAC (206). Another study has shown  $\approx 50$  fold higher transport activity of proteoliposomes composed of 50/50 PC/PE as compared to pure PC proteoliposomes (211). Both CL and PE are conically shaped lipids, which are known to confer positive intrinsic pressure to the membrane. The lateral forces thus could be important for the transport activity of AAC. However the influence of anionic head-group of cardiolipin could not be discarded, as partial replacement of CL with another anionic lipid PS would retain high activity (however, large amounts of PE are still necessary) (211).

## 2.2 Structural details of mitochondrial carriers

### 2.2.1 Sequence based structural information

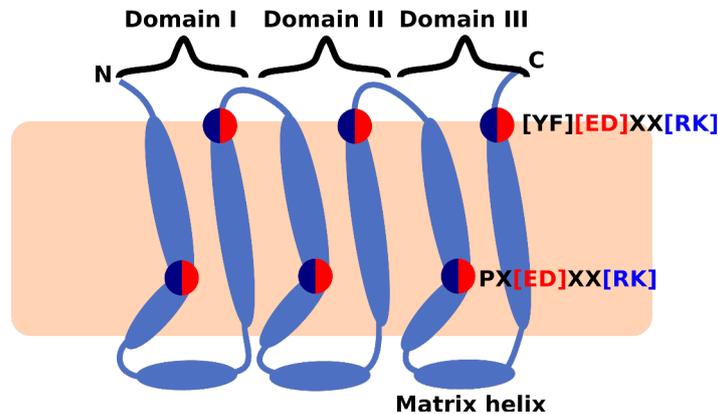


Figure 2.4: A scheme illustrating the general architecture of mitochondrial carriers. The carriers are composed of three domains, as illustrated. In each domain odd TM helices contain a PX[ED]XX[RK] motif with a Pro residue introducing a kink in the helix and charged residues contributing to the matrix salt bridge network. Even helices contain [YF][ED]XX[RK] motifs, which are suggested to form cytoplasmic salt bridge network.

Bovine heart AAC was the first carrier which shed some light of the structural intricacies of MCF members by having its sequence revealed by Edman's degradation. Sequence information of other family members allowed to discover some common features. The majority of the carriers are exclusively TM proteins<sup>16</sup>, composed of  $\approx 300$  residues, which are further subdivided into 3 pseudo-repeat domains. Each domain is composed of two transmembrane helices. Odd helices contain a strongly conserved mitochondrial carrier signature motif (MCS motif): PX[DE]XX[KR]X[KR]-X<sub>(20-30)</sub>-[DE]GXXXX[WYF][KR]G (212), which extends to the matrix helices and reaches the N-terminal ends of even helices. A less conserved [YF][DE]XX[RK] motif was suggested to be present in even helices (37) (see Figure 2.4). The carriers also contain an unexpectedly large amount of charged residues ( $\approx 20\%$ ) for an exclusively

<sup>16</sup>Aspartate/glutamate and ATP-Mg/P<sub>i</sub> carriers are an exception, as they contain  $\approx 150$  residue soluble N-terminal regulatory domains, which are implicated in Ca<sup>2+</sup> binding

Table 3: X-ray structures of AAC

Protein	Detergent	PDB accession code	Space group
btAAC1 (44)	LAPAO	1OKC	P <sub>2</sub> <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
btAAC1 (213)	LAPAO	2C3E	C222 <sub>1</sub>
yAAC2 (190)	Cymal-5	4C9H	P <sub>2</sub> <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
yAAC2 (190)	Cymal-5	4C9G	C222 <sub>1</sub>
yAAC3 (190)	DM	4C9Q	P2 <sub>1</sub>
yAAC3 (190)	DM	4C9J	P <sub>2</sub> <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>

transmembrane protein. Due to the substantial presence of charged residues in the transmembrane helices, hydrophobicity based predictions might underestimate the length of TM helices. The importance of all these features was partially uncovered by X-ray structures.

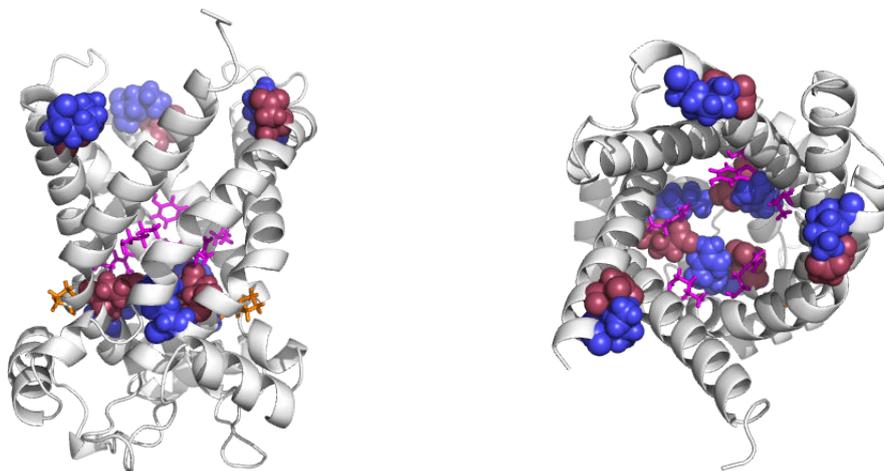


Figure 2.5: An X-ray structure of yeast AAC3 transporter (PDB:4C9Q, invisible loops modeled with I-TASSER). Residues involved in cytosolic and matrix salt bridge networks are depicted in spheres. Residues involved in substrate binding are colored in magenta. Prolines involved in kink regions are colored in orange.

### 2.2.2 X-ray structural information

Structures of three AAC carriers inhibited by CATR were solved by X-ray crystallography (Table 1). Although different detergents were used during protein extraction procedure, the structures of all three carriers show striking similarity, with RMSD of  $<1\text{\AA}$  (190). These structures indicated a unique fold of mitochondrial carriers among all membrane transport proteins (Figure 2.5). Six TM helices, which are highly tilted with respect to the bilayer normal, form a basket-like structure, with a deep, funnel-shaped cavity. The cavity penetrates  $>2$  nm into the would-be bilayer ( $\approx 2/3$  of the bilayer width), being  $\approx 2$  nm wide at the top and narrowing to

$\approx 0.8$  nm deeper inside the bilayer (192). Due to such a large cavity, the TM helices can only form packing interactions with two neighboring helices for most of their span. Inter-domain interactions are likely promoted by small residues on the odd-numbered helices, such as in the GXXXG motif. They allow a close approach of the  $\alpha$ -helices, as the side-chains of residues on the opposite face of even-numbered helices do not protrude into the inter-helical interface. Only few polar interactions are observed between the helices (214). This type of architecture stands in a stark contrast to many other membrane protein structures, where bundles of helices are tightly packed. Besides the transmembrane helices, AAC contains three short matrix helices lying perpendicular to the bilayer normal (see Figure 2.4. These helices link two TM helices of the same domain on the matrix facing side.

TM helices surrounding the cavity reflect the three-fold pseudo-symmetry observed in the primary structure. The proline residue, starting the MCS motif is involved in sharp kink ( $20\text{-}30^\circ$ ) of the odd helices. Pro in this position can be substituted by a Ser (as observed in yeast AACs, for example). The hydroxyl group of Ser adopts an unusual geometry and bends back towards its backbone nitrogen, thereby disrupting the peptide hydrogen bond arrangement in the helical structure, mimicking the effect of Pro (190). The kinked structure is expected to be stabilized by the tertiary structure of AAC<sup>17</sup>.

### 2.2.3 Importance of charged residues in the TM domains of AAC

#### 2.2.3.1 Residues involved in the substrate binding

The cavity is lined with polar residues and possesses a charge opposite to the transported substrates. In the case of AAC, the overall charge of the cavity is dictated by an excess of positively charged residues, which are expected to stabilize the phosphate groups of  $\text{ADP}^{2-}$  and  $\text{ATP}^{3-}$  molecules (see Figure 2.6). As CATR is not a structural analogue of any of the substrates of AAC, several independent approaches were used to identify the residues involved in ATP or ADP binding, which have shown a good convergence. Based on this, a defined binding geometry was proposed for ADP/ATP binding, that is a true association site (192).

Although AAC possesses a three-fold structural symmetry, its substrates are not symmetric, nor the residues involved in the substrate binding. The substrate binding residues are thought to be distributed on even helices. For btAAC1, residues K22, R79 and R279 were inferred to participate in phosphate binding, while the adenine base was found to be stabilized by G182, I183 which allows hydrophobic interactions, Y186 stacking with the aromatic purine ring and S227 involved in the binding of amide group of adenine. In yAAC3, the same residues correspond to K27, R85, R283 (see Figure 2.6, magenta) and G187, I188, Y192 and S231. Analysis of the proposed binding site of AAC suggested the mechanism determining its specificity towards ATP over GT(D)P - the N6 amino group, which acts as a hydrogen bond donor is replaced with a carbonyl group in GT(D)P and C2 carbonyl group of gua-

---

<sup>17</sup>A similar observation was made on bacteriorhodopsin, where Pro50 residue, which kinks a helix by  $20^\circ$ , can be mutated to Ala, preserving the stability and function of the protein (215).

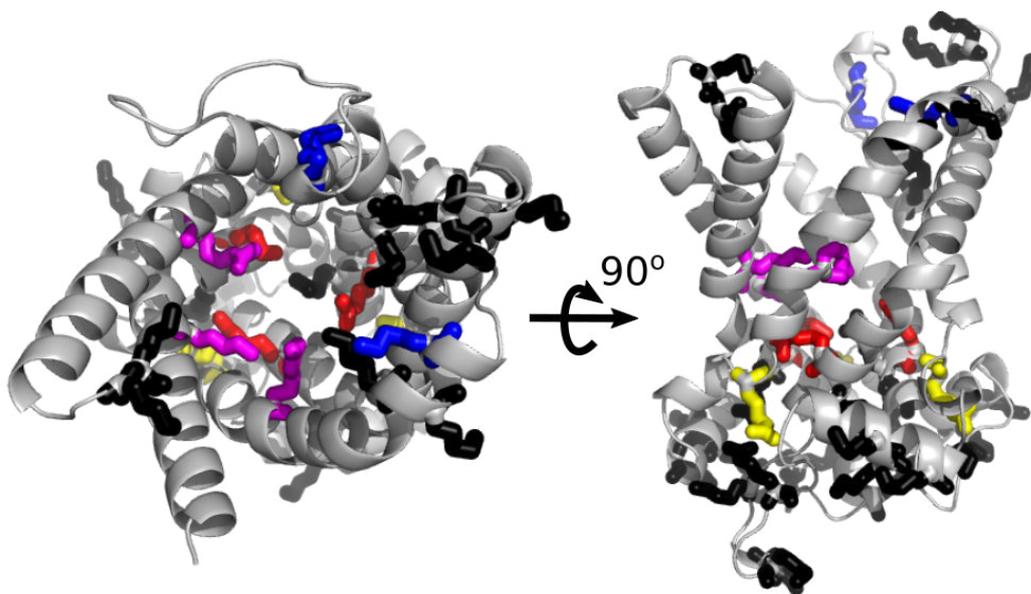


Figure 2.6: Distribution of positively charged residues (Arg and Lys) on the structure of yeast AAC3. Blue indicates residues involved in the cytoplasmic salt-bridge formation, magenta - residues involved in the binding of phosphate groups of the substrate (AdP), red - matrix salt-bridge network and yellow - residues present in cardiolipin binding site. Black residues do not have an assigned function. Top and side views are shown

nine would be facing a hydrophobic pocket in the binding site (216).

### 2.2.3.2 Residues involved in salt-bridge networks

Besides the three Arg or Lys residues, situated approximately midway in the TM region, additional positively charged residues line the cavity of AAC. Three [RK] residues, present in the MCS motif, are forming a salt-bridge network in combination with the [DE] residues from the same motif. These residues tightly close the cavity towards the matrix side and are implicated in stabilizing the c-state of the AAC (and other mitochondrial carriers).

The structurally poorly understood m-state is assumed to be stabilized by the [FY][DE]XX[RK] motif, which was suggested to form a cytosolic salt bridge network (Figure 2.5). The presence of this salt-bridge network was investigated by the thermostability analysis of BKA inhibited AAC<sup>18</sup> and activity assays. The results indicated that the transport is inhibited upon mutation of residues comprising the proposed cytoplasmic salt-bridge (194).

<sup>18</sup>A correlation of predicted strength of the salt-bridge network and measured stability of the protein was observed. However the assay was based on fluorescent molecule reaction with cysteine located in a close proximity to the modified region. Cysteine reactivity is strongly dependant on the surrounding charge: positive charge decreases the pK<sub>a</sub> of Cys, thereby increasing -S<sup>-</sup> population, which participate in the reaction, and thus the reaction rate. On the other hand, if the negative residues surround Cys, the population of -SH is increased and the reaction rate is reduced. See, for example (217).

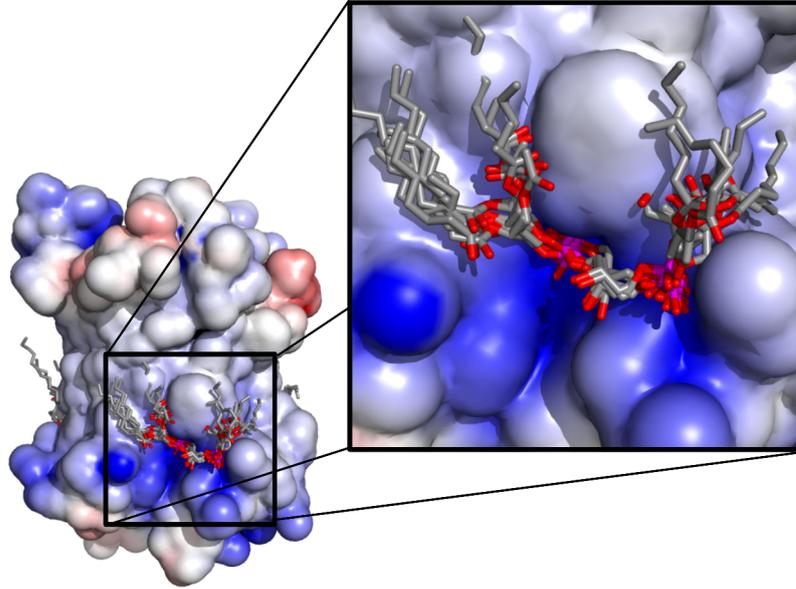


Figure 2.7: One of the three similar binding sites of cardiolipin in AAC. Cardiolipin molecules from all the AAC structures solved are displayed over the binding site of bhAAC. Electrostatic contours? are from -5 (red) to +5 (blue)

### 2.2.3.3 Residues involved in cardiolipin binding

In all but one crystal structures<sup>19</sup> of AAC, three CL molecules are bound per protein molecule. Whereas the aliphatic chains of CL are largely not visible in the structures, presumably due to the lack of a defined binding mode, the head-groups of cardiolipin are visible and situated around the second part of the MCS motif ([DE]GXXXX[WYF][KR]G) and another conserved [YF]XG motif. This implies the function of MCS motif to bind and stabilize cardiolipin molecules.

Mapping the electrostatic potential over the structures of AAC reveals small, strongly positively charged invaginations on the outer surface of the protein at the positions of the cardiolipin binding sites (see Figure 2.7). Such large electrostatic potential is likely contributed by the N-terminal helix macrodipole (214) in combination with the capping by positive Lys/Arg residues. Charged surfaces are strongly disfavored in TM regions (discussed in section 1.2.2), therefore CL binding and subsequent charge compensation is expected to be strongly favourable. Cardiolipin also appears to make interactions with the N-terminal end of the matrix helix from the adjacent domain, with the [YF]XG motif. The [DE]G part of the MCS motif is situated further away from the head-groups of cardiolipin molecules, on the C-termini of the matrix helices<sup>20</sup>.

Considering limited amount of packing interactions, which are assumed to be one of the main stabilizing factors for membrane proteins according to the current paradigm, CL binding to MCs can be a very important stabilizing factor. Alternatively, CL

<sup>19</sup>In the PDB:2C3E structure only two cardiolipin molecules are visible

<sup>20</sup>As the C-terminal termini of helices have a partial negative charge, a positioning of a negatively charged residue is expected to be destabilizing. Therefore, its function is not clear.

may protect MCs from interaction with other molecules, which could otherwise limit the flexibility of the dynamic regions (190).

#### 2.2.3.4 Positively charged residues without a clearly defined function

Lastly, several additional positively charged residues are present in the cavity without a known function assigned to them. As it was suggested that AAC's cavity might act as an electrostatic funnel (218) to facilitate substrate binding and it is inhibited by high concentration of chloride ions (219), it is likely that these residues play a specific role in directing ADP or ATP towards the binding site.

Furthermore, a lot of Lys and Arg residues are also present in the matrix helices of mitochondrial carriers. These residues might be important for the function in the m-state.

### 2.3 Transport mechanism suggested for AAC.

All the experimental results on AAC are best explained by an "induced transition fit" mechanism suggested by Klingenberg (198). This model is similar to the enzyme catalysis model, where the transition state of the substrate would have the strongest affinity to the enzyme. In the case of AAC and other MCs, however, the transition state of the carrier has the strongest affinity to the transported solute. As the carriers need to switch between two stable states, which involves large conformational changes (i.e. redistribution of many non-covalent interactions, including salt-bridges), the transition barrier for such a conformational change is large. The main contributors defining the stability of the two end-states (c- and m-state) of the carriers are believed to be the strongly conserved salt-bridge networks on both sides of the carrier(37). Thus the strong binding of the substrate to the transition state of the carrier provides necessary energy for salt-bridge network breaking. At the same time strong specificity is ensured, as molecules which are not able to provide chemical groups to satisfy all the binding interactions might not provide enough energy to break the salt bridges (198, 214)). Although AAC does not inter-convert between its two stable conformations in the absence of the substrates and therefore acts as exchanger, other MCs have been shown to promote uniport. This quality is likely dictated by the weaker salt-bridge network in the m-state, thereby reducing the stability of the m-state and allowing spontaneous interconversion to the c-state due to the thermal energy (37).

Although the residues involved in the substrate binding are not symmetrically distributed inside the cavity, the overall conformational changes leading to the transitions between the two states are believed to be symmetrical. This assumption is based on the high degree of overall tripartite symmetry of the carriers. If the symmetry was not relevant for the function of the carriers, it would have likely been lost due to neutral mutations. Besides the symmetric charged residues in the strongly conserved PX[DE]XX[KR] motif in odd-numbered helices, which are involved in salt-bridge formation, the Pro residues are thought to be of major importance, as they were suggested to act as hinges, allowing the opening of the protein towards the ma-

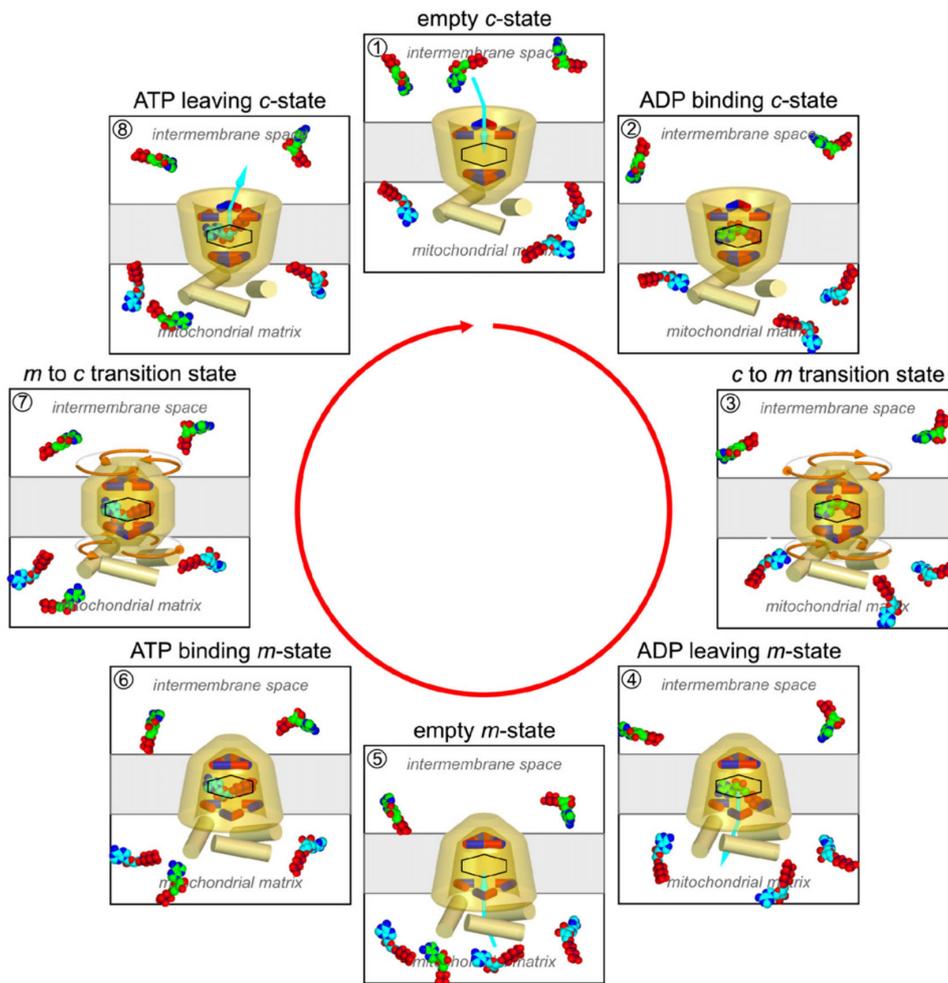


Figure 2.8: Transport mechanism suggested for AAC. An empty AAC molecule exists in a defined conformation (in this case, c-state stabilized by the matrix salt bridge network (1). Binding of a substrate (ADP or ATP) (2) provides the energy to break the matrix salt bridge network, which leads to a transition state (3). The transition state, in principle, should be able to further evolve back into c-state, or into m-state (4). After the substrate dissociates, AAC remains in an m-state, stabilized by the cytosolic salt bridge network (5). The second half of the cycle (6-7-8) occurs in a similar fashion. Figure taken from ref (220).

trix side. Similar proline residues are present above the proposed substrate binding site on even helices. As intra-helical hydrogen bonds appear to be very strong in TM proteins, rendering the helices very stiff, Pro (and probably Gly) residues, which have very low helical propensity, might allow kinking of the helices during conformational transitions (188, 221).

Another proposed mechanism suggests that opening could also be achieved by rigid-body movements of the three domains, where the domains slide along each other, thus closing the cytosolic salt-bridge network and opening the matrix gate (190, 214) (see Figure 2.8). As the ligand binding residues are distributed between different domains asymmetrically, this mode of conformational change would rearrange

their orientation and suggest that other residues might participate in the substrate binding in the m-state of the carriers.

## 2.4 Other mitochondrial carriers

Besides AAC, MCF consists of a large number of other members (see Figure 2.1). In total 35 MCs are present in *Saccharomyces cerevisiae*, 50 in humans and can reach more than 100 in some plants (212). These members can be subdivided into several subfamilies, based on their transported substrates: 1) nucleotide and dinucleotide carriers, 2) di-/tri-carboxylate and keto acid carriers, 3) amino acid carriers and 4) carriers for other substrates.

Due to the conservation of specific motifs, which were already discussed above, it is believed that most of the carriers share a lot of structural similarities with AAC and function in a similar mechanistic fashion. However, some marked differences are present in other members of MCF, especially regarding their sensitivity to inhibitors, utilization of different (or none) energy sources to enhance transport, their substrate specificity and promiscuity to transport of unrelated molecules. As it is out of the scope of this work to discuss the behaviour of different mitochondrial carriers (188), only GTP/GDP carrier (GGC) and ornithine carrier (ORC), which this work directly relates to, will be presented in regards to their most pronounced differences from AAC.

### 2.4.1 GTP/GDP carrier

GGC shares many similarities with AAC. It is also a nucleotide transporter, and like AAC it is not able to transport GMP. It is also functioning as a strict exchanger, as no transport was observed in proteoliposome assays when the vesicles were not pre-filled with one of the transported substrates (78). Although GTP/GDP carrier (GGC1) transports nucleotides, similarly to AAC, there are marked differences between these two carriers.

First of all, while AAC is ubiquitously expressed in all the organisms due to its important role in oxydative phosphorylation, GGC1 is only present in fungi<sup>21</sup>(78). Second, as opposed to AAC, the primary function of GGC is the import of triphosphonucleotides. Therefore, electrogenic transport mechanism cannot be employed by this transporter (as the nucleotides would be transported against the electrostatic gradient). GGC thus imports GTP complexed with H<sup>+</sup> in an overall electroneutral manner of transport.

Third, GGC is somehow less specific than AAC, as is able to transport 2'-deoxy variants of GTP/GDP and ITP and IDP (although the observed  $K_M$  values were one and two orders of magnitude higher for these compounds, respectively (212).

Fourth - although GGC is transporting nucleotides, it seems to form a separate class of mitochondrial carriers due to the limited similarity of its conserved repeats to

---

<sup>21</sup>This is dictated by the absence of GTP production in these organisms. GTP is used for mitochondrial nucleic acid synthesis and as a co-factor for the mitochondrial translation machinery. In most metozoans, two enzymes are responsible for the production of GTP in mitochondrial matrix: nucleotide diphosphate kinase and succinyl-CoA ligase. In fungi, however, succinyl-CoA ligase produces ATP instead of GTP, and NDK is absent in the matrix.

the other members; strikingly it appears to contain some conserved regions which resemble carboxylate carriers (212).

Fifth, GGC do not respond to the inhibitors CATR or BKA, compounds which are inhibiting AAC strongly (78).

GGC also differs from AAC by the composition of salt-bridge networks. The matrix network in GGC is not complete, as Ala or Ser replaces one of the positively charged residues. The cytoplasmic salt-bridge network appears to be very weak, as several charged residues are replaced with polar or even non-polar ones. Nevertheless, GGC is not able to catalyse uniport, as measured in proteoliposomes (37).

The proposed residues for substrate binding in GGC1 are K82, K89 and R279 for the binding of phosphate groups, and E133, R180 and N181 for the guanine base.

#### 2.4.2 Ornithine/citrulline carrier

ORC, as opposed to AAC and GGC is an amino acid transporter, and its substrates are positively charged, in contrast to nucleotide transporters. ORC transport mode is somehow different in unicellular organisms, such as yeast (77), and metazoans (222). This is dictated by the physiological differences between the organisms.

The urea cycle is important for multicellular organisms in the detoxification of ammonia by its conversion to a less toxic compound, urea. In metazoans, ornithine carrier mediates the metabolic steps of the urea cycle, which are composed of enzymes residing both in cytoplasm and in mitochondrial matrix. Ornithine is imported into mitochondrial matrix by ORC, where it is subsequently converted to citrulline, and the latter is exported by the same carrier in a 1:1 stoichiometry, which is a similar transport mode to previously discussed nucleotide carriers (222).

Citrulline is co-transported with an  $H^+$  ion, which moves against the electrochemical proton potential of mitochondria, while the overall charge is not transferred (as ornithine is positively charged and citrulline is neutral). The residue suggested for proton binding is E31 (37). The prevention of the counter-productive movement or non-productive movement of ornithine would then necessitate high intra-mitochondrial concentrations of citrulline (76).

Experimental studies on reconstituted rat ornithine carriers have shown somehow broader specificity to positively charged amino acids, such as ornithine, arginine, lysine and citrulline. The transporter inhibition by cations was observed, similar to AAC inhibition by chloride ions (223). The inhibition of ORC, as in the case of AAC, is suggested to be caused mainly by the competition of positively charged ions with the substrates.

The absence of ORC1 activity in humans causes hyperammonaemia-hyper-ornithinaemia-homocitrullinuria (HHH) syndrome.

Organisms such as yeast, however, are not faced with the problem of ammonia accumulation, and instead use ornithine for the production of arginine. Thus in yeast ORC works in an opposite direction, exporting ornithine from the matrix at the expense of mitochondrial proton gradient (77, 224). The matrix salt-bridge network is complete in these carriers, however the cytoplasmic network was observed to be

weaker, allowing for some uniport activity (37).

As opposed to nucleotide carriers, the positive nature of the transported substrates is reflected in a negatively charged cavity. The residues suggested by asymmetric residue analysis for substrate binding in ornithine carriers include E83 and E179 (scORC1), which are likely involved in stabilization of two amine groups of the substrates. R178 and R179 were suggested to bind the carboxylate group.

## 2.5 NMR studies of mitochondrial carriers

Several carriers (SCaMC1, GGC1, UCP2 and AAC3) were investigated by solution state NMR in DPC micelles in order to get structural and mechanistic insights into their function. Although considered as a relatively harsh detergent by some, due to its solubilizing ability (see section 1.5.5), DPC is preferable in NMR studies of mitochondrial carriers as it is able to prevent protein aggregation for prolonged times, required for NMR measurements. In addition, the carrier spectra in DPC usually display a single set of well resolved peaks, which may be interpreted as a structurally homogeneous ensemble of protein particles. Furthermore, the functional relevance of DPC solubilized mitochondrial carrier samples is supported by their ability to bind substrates and inhibitors (albeit the affinities are decreased). The following section will overview the current state of mitochondrial carrier research by solution state NMR and the insights provided by these experiments.

### 2.5.1 Human SCaMC1 - determination of specificity towards $\text{ATP} \cdot \text{Mg}^{2+}$

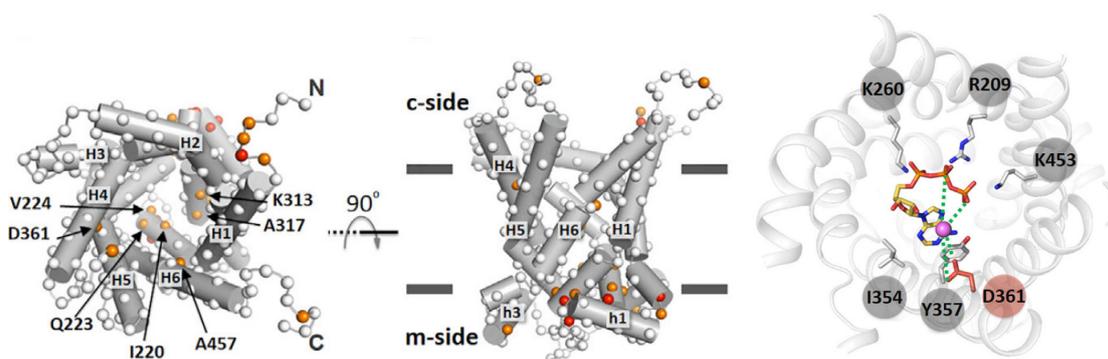


Figure 2.9: Left: the regions which show chemical shift perturbations upon titration from 0 to 16.7 mM  $\text{ATP} \cdot \text{Mg}^{2+}$ . Red spheres indicate chemical shift perturbations larger than 0.03 ppm, while the orange spheres perturbations in the range 0.015-0.03 ppm. Right: the proposed model for  $\text{ATP} \cdot \text{Mg}^{2+}$  binding to the SCaMC1 and the importance of D361 in substrate selectivity (225)

Calcium-binding mitochondrial carrier protein 1 (SCaMC1 or APC1) is an  $\text{ATP} \cdot \text{Mg}/\text{P}_i$  exchanger, importing external  $\text{ATP}$  and exporting  $\text{P}_i$  from the mitochondrial matrix. Unlike AAC, this transporter is able to transport  $\text{ATP} \cdot \text{Mg}^{2+}$  complexes in electroneutral manner. Moreover, these complexes are the preferential substrates for

SCaMC1 ( $K_M$  200  $\mu$ M for ATP  $\cdot$  Mg<sup>2+</sup> as compared to 330  $\mu$ M for ATP (226) in proteoliposomes). The underlying difference in substrate specificity between AAC and SCaMC carriers is not understood in structural terms. It was therefore investigated by NMR spectroscopy in residue-wise resolution, where transmembrane domain of SCaMC1 was refolded in 60 mM DPC micelles and functionally characterized.

The ligand binding was investigated by titrating the ligands to the protein sample and observing HN chemical shift perturbations. These experiments revealed a multitude of chemical shift perturbations along the structure of SCaMC (see Figure 2.9), which was interpreted as the ligand is being loosely bound inside the cavity (i.e. without a defined orientation) and a possibility of related conformational changes of the protein. As the chemical shift perturbations report on the changes of the local electronic environment of an atom, which can occur both due to ligand binding and conformational changes, further paramagnetic relaxation enhancement experiments were performed. These experiments measured the ability of ATP  $\cdot$  Mg<sup>2+</sup> to expel chemically similar ATP  $\cdot$  Mn<sup>2+</sup> ligand bound to the carrier. As Mn<sup>2+</sup> is a paramagnetic ion, the peaks in its vicinity have increased transversal relaxation rates and therefore are broadened. Replacement of Mn<sup>2+</sup> ions with Mg<sup>2+</sup> results in the recovery of peak intensity. Therefore, the sites which have specific ATP  $\cdot$  Mg<sup>2+</sup> binding are expected to display stronger peak intensities with increasing ATP  $\cdot$  Mg<sup>2+</sup> concentration. The results reportedly indicated two binding sites in SCaMC1: one located on helix 4 (in a good agreement to the binding site identified by the analysis of symmetry), with a relatively high apparent affinity ( $K'_D=830$   $\mu$ M) and the second one in the loop between helix 4 and helix 5 ( $K'_D=1$  mM).

Among the identified residues involved in substrate binding, D361 was suggested to be responsible in determining the selectivity of SCaMC to ATP  $\cdot$  Mg<sup>2+</sup>. Indeed, mutating this residue to Tyr resulted in a protein which was able to transport both ATP and ATP  $\cdot$  Mg<sup>2+</sup> at the same efficiency (albeit several-fold lower than wild-type protein) in proteoliposome assays. Analysis of D361A mutant by NMR revealed that Mg<sup>2+</sup> was no longer able to replace Mn<sup>2+</sup> ions. Based on the aforementioned experiments and the conservation of D361 among SCaMC carriers it was suggested that this residue is the foremost determinant of the particular specificity of SCaMC carriers (225) (Figure 2.9).

### 2.5.2 Human SCaMC1 - the role of the N-terminal domain in the activation of the carrier

Unlike many MCs, SCaMC contains a soluble N-terminal Ca<sup>2+</sup> binding domain (NTD; Figure 2.10). The NTD domain is composed of two lobes, each containing a pair of Ca<sup>2+</sup> binding EF hand motifs. The two lobes are connected by a long helix (helix H4-H5), with a short amphiphatic C-terminal helix (H9) intercalated between EF hands (81, 227). The domain regulates the carrier's activity by activating it in the presence of Ca<sup>2+</sup> ions. However, the mechanism of regulation in this type of carriers is not completely understood.

To understand the regulation mechanism on structural level, Ca<sup>2+</sup> induced changes in conformation and dynamics of NTD were investigated. Analysis of the relaxation

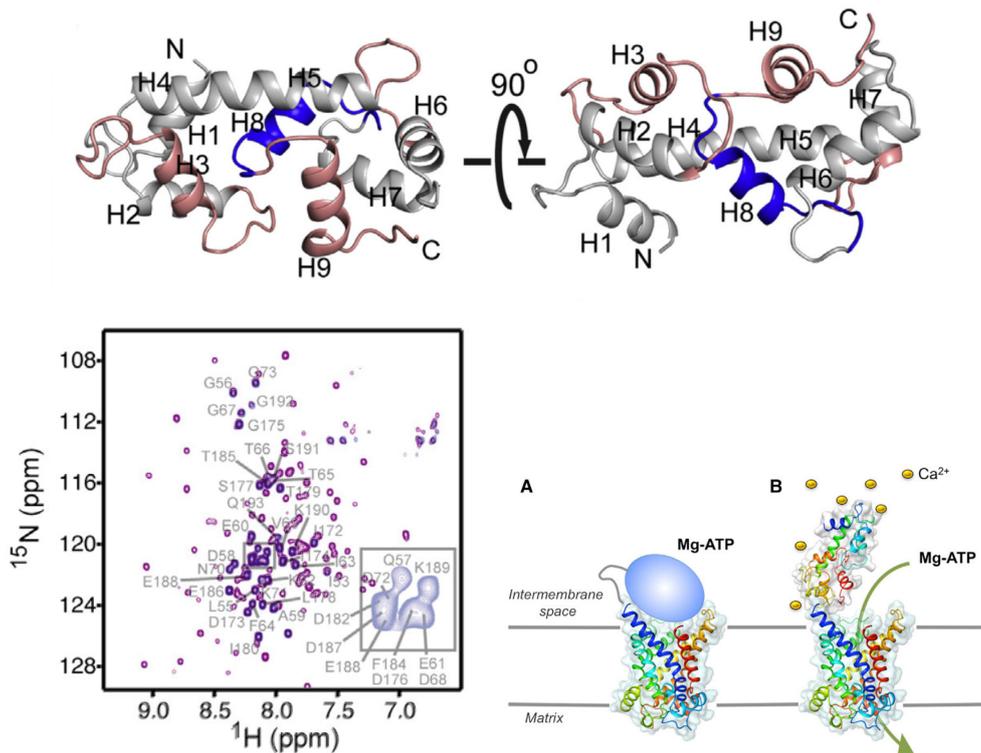


Figure 2.10: Top: Structure of the N-terminal domain of human SCaMC1 ATP-Mg/P<sub>i</sub> exchanger. The order parameter of pink regions is lower than 0.75 in the absence of Ca<sup>2+</sup>, indicating increased flexibility. The blue regions are broadened due to millisecond dynamics in the absence of calcium. Bottom left: overlay of spectra of SCaMC1 NTD domain in the presence of liposome-embedded TMD domain, and the presence (magenta) or absence (blue) of Ca<sup>2+</sup>. Bottom right: The mechanism proposed based on the observed peak disappearance of the complex of NTD-TM domains of SCaMC1 (81)

rates and heteronuclear NOEs indicated an increased flexibility in the domain upon removal of Ca<sup>2+</sup> ions, which points to a complete unfolding of a helix in the vicinity of calcium binding sites (H3; Figure 2.10) in one of the EF hands, and increased flexibility and partial unfolding of the amphiphatic helix (H9). To investigate if the observed conformational changes of NTD may result in the modulation of binding properties between NTD and TM domain, the analysis was further extended to the interaction measurements of the soluble NTD and DPC solubilized and refolded TM domain of SCaMC1 reconstituted in liposomes.

Addition of soluble 100 μM of NTD to 100 μM of liposome-reconstituted TM domain in the presence of Ca<sup>2+</sup> resulted in an identical spectrum as NTD in the absence of TM domain. It indicates that no binding to the TM occurs in the presence of Ca<sup>2+</sup>. However, in the absence of Ca<sup>2+</sup>, the spectrum of NTD mixed with TMD in an equimolar ratio had a very narrow peak dispersion and no peaks from the structured regions were visible (see Figure 2.10). These results were interpreted as the Ca<sup>2+</sup>-devoid NTD binding to the TM domain of the carrier, thereby blocking the cavity of the carrier and inhibiting the transport (see Figure 2.10). As the NTD forms a

stable complex with TM domain, which is embedded in large liposomes, the overall tumbling rate would broaden non-flexible parts of the protein beyond detection, such that only the loops remain visible.

NMR binding data was further supported by the SPR analysis, which revealed a binding constant of 250  $\mu\text{M}$  for the interaction of apo NTD domain with the TM domain (81).

The persistence of disordered regions in NTD upon complex formation prompted an independent study, where comparison of multiple EF hand domains from different calcium binding proteins suggested the closure of EF hand domains upon the release of  $\text{Ca}^{2+}$  and the concerted expulsion of the amphiphatic C-terminal helix from the interface of two EF domains. This helix was suggested to play the major role in blocking the cavity of the carrier and inhibiting its activity, in contrast to the aforementioned model, where the whole NTD blocks the cavity (227).

These results are in good agreement with the observed flexibility of C-terminal helix. However, they do not agree with the proteoliposome binding assays done by NMR. The peaks corresponding to the C-terminal extension, which would link the NTD to the TM domain are unperturbed both in their shifts or intensities. Moreover, considering that the H9 is excluded from the rest of the NTD, it should be flexible in respect to the rest of the NTD domain. Therefore, if H9 was bound to the TM domain, the peaks of the remaining NTD should be clearly visible. In contrast, only loop regions are visible (81), indicating that they remain disordered and do not participate in the binding to TM domain.

### 2.5.3 GGC1 - conformational flexibility

To understand the flexibility of uninhibited carriers, GGC1 was studied in micelle reconstituted state in 80-120 mM of DPC by measuring residual dipolar couplings. As it is not possible to determine the biological relevance of transporters in solution-state NMR compatible conditions by activity measurements, ligand binding is often used instead. Ligands (GTP and GDP) were shown to bind GGC1 in a similar manner as reported for SCaMC1, inducing backbone chemical shift perturbations along most of the structure. More detailed investigation revealed vastly different  $K_D$  values for different residues, which were interpreted to represent multiple binding sites along the cavity of GGC1 and a population-averaged picture of both of m- and c-state binding to the ligands. Multiple binding sites were suggested to facilitate ligand recruitment to the active site of the protein, in agreement to previously noted electrostatic funnel of AAC (218). Overall  $K_D$  for GTP was estimated to be 6.6 mM, while for GDP 23 mM. These numbers are substantially larger than reported  $K_M$  values ( $K_M^{GTP}=1.2 \mu\text{M}$  and  $K_M^{GDP}=4.5 \mu\text{M}$ , nonetheless the observed ratio of GTP/GDP affinity was preserved. In an agreement with SCaMC1, additional binding site was revealed, which could be mapped to the top of the cavity and involved helix 1 and helix 6.

Attempts to fit the residual dipolar coupling data indicated that even adjacent 7 residue segments in the same helix do not share the same alignment tensor (Figure 2.11). This indicates the movement of different segments in the same helix in respect to each other and large overall conformational heterogeneity. This conformational

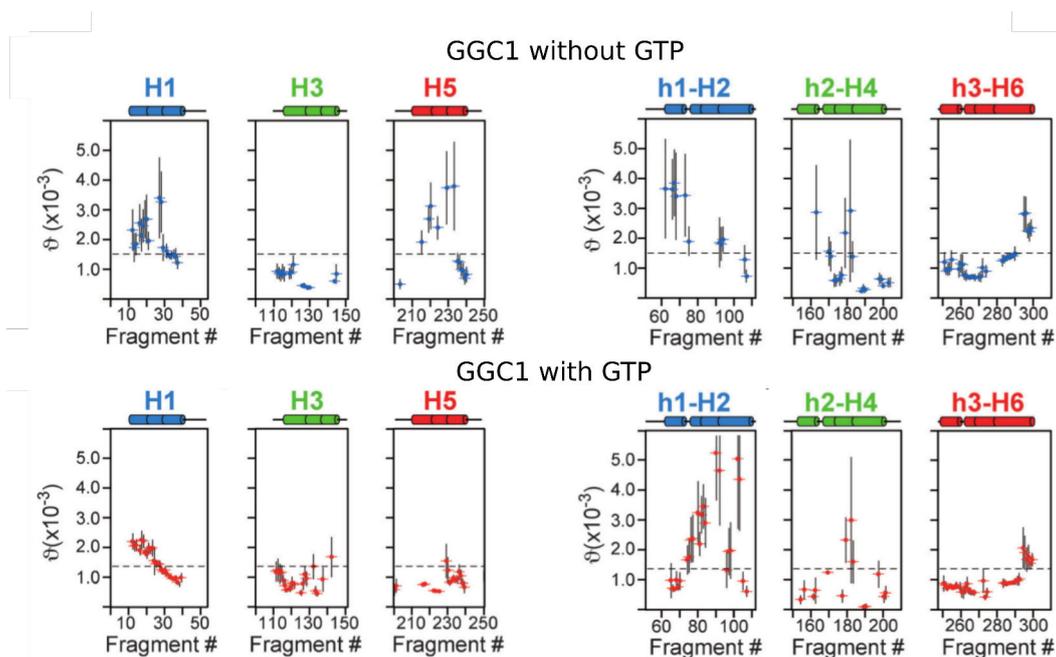


Figure 2.11: Conformational heterogeneity in GGC1 in the presence and absence of 30 mM GTP.  $\vartheta$  indicates the degree of conformational heterogeneity (228).

heterogeneity can be interpreted as flexibility of the protein on relatively fast time-scale ( $>10$  ms), as otherwise multiple sets of peaks would be visible in the NMR spectra, which is not the case.

The observed conformational heterogeneity was not in agreement with the three-fold symmetry of mitochondrial carriers. For example, among the odd-helices, helix 1 and 5 displayed relatively large amount of conformational heterogeneity, while helix 3 was quite uniform. Similarly, helix 6 was less heterogeneous (except the C-terminus), while helices 2 and 4 were sampling more conformations. Addition of GTP decreased heterogeneity in the odd helices, reportedly around the Pro-king regions. However, it has little effect on the even helices.

In general, the matrix helices were reported to be more heterogeneous than the TM helices. This behaviour was explained by assuming an equilibrium of c- and m-states in the absence of ligand, where matrix helices would be expected to sample larger conformational space.

These observations present a very interesting aspect of membrane proteins - although TM helices are generally assumed to be very stable, which is supported by a large amount of evidence, it appears that in some proteins, such as mitochondrial carriers, they can be very flexible and sample many different conformations. The mechanism leading to such large flexibility of helices was suggested to be hydrogen bond network rearrangements between  $\alpha$ -helix peptide bonds (228).

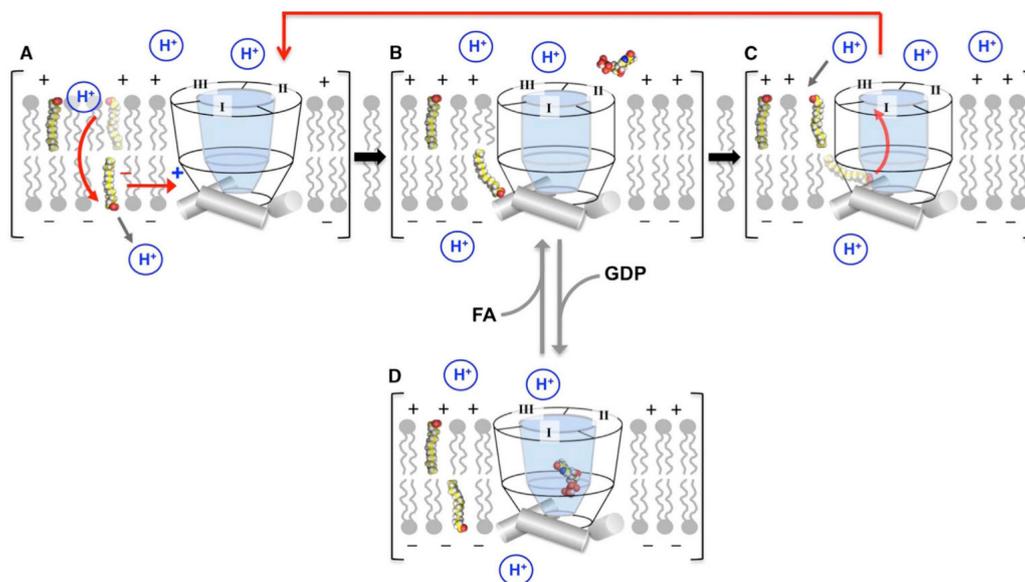


Figure 2.12: Suggested mechanism of fatty acid and proton transport by UCP2. A, an excess of protons on the outer side of the inner mitochondrial membrane results in a higher population of protonated fatty acids. This in turn facilitates the flipping of a protonated and therefore uncharged fatty acids across the bilayer. The fatty acids, however, become deprotonated on the inner side of the membrane. B-C, deprotonated fatty acid binds to the UCP2 in between helices 1 and 6, and the carrier facilitates the flipping of the charged FA to the outer side of the membrane. D, in the presence of GDP the FA is no longer able to bind to the carrier and transport is inhibited. Figure taken from (229).

#### 2.5.4 Analysis of UCP2 proton and fatty acid transport

The backbone structure of mouse UCP2 was solved using residual dipolar couplings and paramagnetic relaxation enhancement constraints by solution state NMR and its discussion was presented in the previous chapter. In addition to the structural information, the mechanism of transport by this carrier was also studied by a combination of NMR and biochemical studies. UCPS can transport  $H^+$  ions, but they are also reported to transport fatty acids. To understand if there was a link between proton and fatty acid transport, the interactions on UCP2 and fatty acids were revealed by solution state NMR. The binding of fatty acids from 50  $\mu$ M to 20 mM was measured as chemical shift perturbations, by titrating fatty acid to 1 mM UCP2 sample in DPC .

Reportedly, the most pronounced chemical shift perturbations occurred between helices 1 and 6. It has been observed that the inhibitor of UCP2, GDP, is able to displace the bound fatty acids by an allosteric effect - RDC measurements indicated wide-spread differences between UCP2-fatty acid (UCP2-FA) and UCP2-GDP sample, with largest differences reportedly concentrated around the fatty acid binding site observed by CSPs.

This feature was used to scrutinized the binding site by applying a similar strategy as was used SCaMC - namely replacement of a paramagnetic fatty acid by increasing

concentrations of GDP, and thus the recovery of signals in NMR spectrum. This method allowed to determine the fatty acid binding site more precisely, with the residues reporting displacement of fatty acids by GDP located in the midway of a TM region, both on the surface of the carrier and inside the polar cavity. The surfacial residues were of hydrophobic character, but two nearby charged residues, Arg60 and Lys271 were suggested to be involved in the binding of the head-group of a fatty acid. Indeed, mutation of these residues to Ser resulted in substantial reduction of fatty acid transport in proteoliposome assays. Furthermore, the same mutations correlated with a strong reduction of H<sup>+</sup> transport in UCP2. This which was interpreted as a proof of a strong link between H<sup>+</sup> and fatty acid transport. Based on this finding it was proposed that a larger population of fatty acids are protonated at the outer leaflet of the inner membrane, due to lower pH, therefore they can easily flip to the inner leaflet of the bilayer. However, at the inner leaflet they are deprotonated and UCP2 serves as a catalyst in flipping back the fatty acid to the outer leaflet, by allowing the head-group of fatty acid to enter its polar cavity. It was further suggested that the proton transport is essentially independent of the carrier activity, as the main substrate for the carrier is a deprotonated fatty acid (229) (see Figure 2.12 for suggested transport mechanism).

### 2.5.5 Substrate modulation of millisecond dynamics in AAC3

As discussed previously in the context of AAC, the currently available biochemical and structural information does not provide a detailed understanding of the transport mechanism of mitochondrial carriers. Although it is now known that these transporters interconvert between two states in a process where the transported substrates lower the free energy of the transition state, the structural changes involved in this interconversion are hypothetical, based on indirect observations. An attempt was made to apply solution state NMR techniques in order to elucidate the details of these conformational transitions during the transport cycle of AAC.

The AAC3 sample was refolded and reconstituted in 120 mM of DPC micelles. ITC measurements determined that the  $K_D$  of CATR binding to AAC3 sample was 15  $\mu\text{M}$ , which was interpreted as showing functional relevance of the sample. The  $K_D$  measured by titrating CATR to NMR samples was 150  $\mu\text{M}$ . This decrease in affinity was attributed to increased concentration of detergent in NMR samples. Correspondingly, the ADP binding constant was determined by NMR, which indicated a binding affinity of 500  $\mu\text{M}$ .

As observed in all the mitochondrial carriers studied by NMR, AAC displayed a well resolved NMR spectra, where the number of NMR cross-peaks coming from amide groups of the protein was approximately the same as the number of amino acids in a single molecule. Therefore, the spectra were interpreted to reflect a structurally homogeneous population of AAC molecules in the sample, reflecting a more stable c-state. However, CPMG relaxation dispersion (RD) measurements indicated millisecond dynamics on a subset of residues. CPMG RD measurements are a very sensitive way to detect millisecond exchange phenomena in proteins, where even low populated additional state(s) of a protein are in an equilibrium with the ground state.

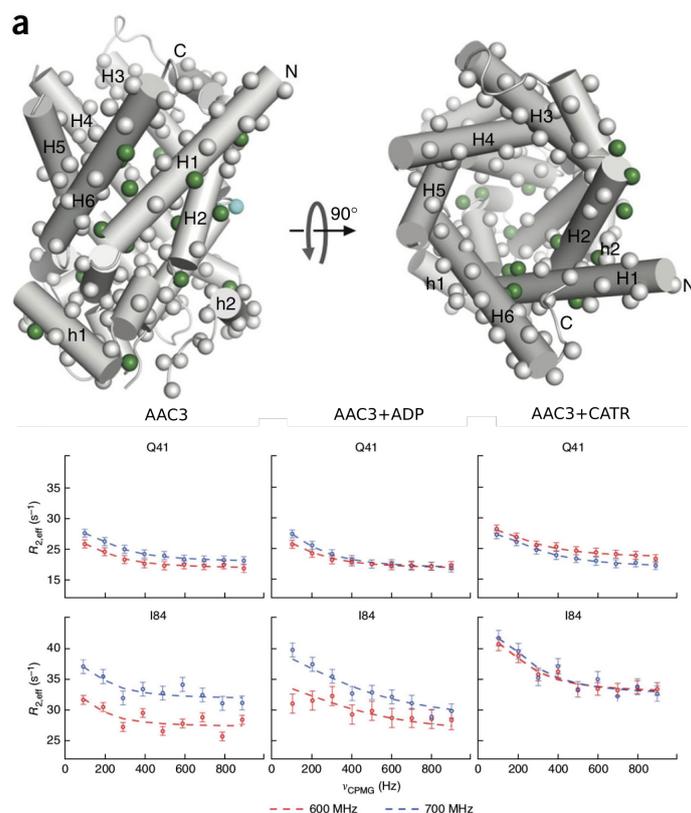


Figure 2.13: Top: distribution of observed CPMG dynamics on AAC structure. Bottom: The differences in relaxation dispersion profiles depending on the substrates and inhibitors (200)

Strikingly, the dynamics observed in the apo AAC3 were distributed in an asymmetric manner along the protein structure (see Figure 2.13, top), indicating a disagreement with a proposed three-fold symmetrical model of AAC transport mechanism. Strongest chemical exchange was reported for a Pro-kink region in helix 1, while the same region of helices 3 and 5 did not suggest any exchange. The difference of chemical shifts between the ground and excited states (i.e. structural differences), populations of the two states and exchange rates from CPMG profiles were extracted by fitting the data to analytical solution of Bloch-McConnell equation for two-state exchange. In this analysis TM regions and matrix helices were considered to move independently of each other.

Further analysis showed substrate and inhibitor effects on the observed dynamics (see Figure 2.13, bottom). While exchange in the matrix helices did not seem to be significantly affected by the presence or absence of ADP or CATR, TM helices were affected. The exchange rate with 10 mM ADP was increased two-fold in TM helices, while 3 mM CATR slowed down the exchange nearly six-fold. The populations, on the other hand, did not seem to be affected. These results were interpreted to reflect the exchange of AAC3 between m- and c-states in an asymmetric manner. As the populations were not affected by the presence or absence of the ligand, it was suggested that the substrates of AAC do not affect the stability of the two states, and

instead the ligand modulates the energy of the transition state. This interpretation is in agreement to the proposed model of AAC transport based on biochemical assays, however the asymmetric nature of the exchanges disagrees with the models proposed for the structural rearrangements (200).

## 2.6 Motivation to perform NMR studies on mitochondrial carriers

The current paradigm which describes the transport mechanism of mitochondrial carriers and is based on a large amount of scientific evidence, suggests that mitochondrial carriers transport their substrates by cycling between two conformations, which are open either to the cytoplasmic or the mitochondrial matrix sides. Furthermore, in both states the substrates are thought to bind to a single binding site, located approximately mid-way along the bilayer normal. Three-fold symmetry of the transporters also suggests that conformational changes, which the proteins undergo during the transport cycle should follow the same symmetry.

The exact molecular mechanism of the conformational exchange, however, is not known, as well as the structure of the m-state. Previous measurements of the inter-conversion rates suggest that the conformational exchange occurs on the time-scale of hundred times per second (204).

Such a system, therefore, seems to be ideal for Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion studies. These NMR experiments provide information about the dynamic processes occurring in the frequency range of 50-5000 s<sup>-1</sup> on single residue resolution. More specifically, considering a system which is exchanging between two well-defined states, the information about the population of each state, the rate of exchange between these two states and the modulus of chemical shift difference between the states can be extracted. Chemical shift difference might be interpreted as a structural term, as some nuclei, such as C<sub>α</sub> and C<sub>β</sub> provide information about secondary structures. Therefore, both thermodynamic and structural information is provided by CPMG experiments.

Mitochondrial carriers are large proteins by NMR standards and they are surrounded by a DPC micelle, giving a cumulative weight of ≈50-60 kDa. This poses several challenges in their study. First of all, the increase in particle size results to decreased tumbling rate of the macromolecule. In turn, the transverse relaxation is increased and the resonance cross-peaks observed on NMR spectra are broadened (for larger particles the broadening might result in the complete disappearance of the signals). This limitation can now be overcome to a certain extent by applying TROSY experiments, which select only slowly relaxing components and replacement of protons in the protein by deuterons (<sup>2</sup>H), which reduces the severity of transverse relaxation.



## 3 Materials and methods

### 3.1 Sample preparation

#### 3.1.1 Reagents

Gadodiamide was provided by Dr. Klaus Zangger (University of Graz, Austria), 1,1',2,2'-tetraoleyl cardiolipin and bovine heart cardiolipin extract was obtained from Avanti Polar Lipids, n-dodecylphosphocholine and 38-<sup>2</sup>H n-dodecylphosphocholine came from Anatrace, carboxylatractyloside (CATR), guanoside-5'-triphosphate (GTP) and adenosine-5'-triphosphate (ATP) from Sigma-Aldrich.

#### 3.1.2 Mitochondrial carrier constructs

All the plasmid constructs used for mitochondrial carrier expression were obtained from GeneScript. GGC1 and GGC1 (Pro138Arg, Pro230Arg) sequences containing C-terminal 6-histidine purification tags were cloned into pET21a+ vector via NdeI and XhoI sites. ORC1 construct containing N-terminal 8-histidine tag was cloned into pET28a vector via NcoI and XhoI.

#### 3.1.3 Protein expression

All samples studied by NMR were expressed in M9 medium<sup>22</sup> containing <sup>15</sup>NH<sub>4</sub>Cl and <sup>13</sup>C-glucose (or <sup>13</sup>C-glucose with all <sup>1</sup>H atoms exchanged to <sup>2</sup>H) as the only sources of nitrogen and carbon, respectively. Isotopic enrichment is required, as the natural abundance of the corresponding isotopes is  $\approx 1\%$  or less, and the common <sup>14</sup>N and <sup>12</sup>C isotopes are spin 1 or spin 0, which loosely speaking do not result in an observable NMR signal. Furthermore, as the system we studied is large for solution state NMR standards  $\approx 50$ -60 kDa, the protein was highly enriched with <sup>2</sup>H, to dilute the pool of aliphatic non-exchangeable <sup>1</sup>H atoms. <sup>1</sup>H, atoms due to their large gyromagnetic ratio and abundance, contribute strongly to dipolar couplings and in turn are one of the most important contributors to transverse relaxation rates (therefore the spectra of large proteins are of low sensitivity and resolution). This enrichment was achieved by growing the BL21(DH3) *E. coli* strains in  $\approx 98\%$  D<sub>2</sub>O. D<sub>2</sub>O is generally toxic to bacteria, therefore adaptation was required, by first inoculating the bacteria in Lysogeny Broth, then in subsequent steps the culture was grown in H<sub>2</sub>O M9 medium, 1:1 H<sub>2</sub>O:D<sub>2</sub>O medium and the final preculture was obtained by growing bacteria in 98% D<sub>2</sub>O. This preculture was used to inoculate 0.5 l of D<sub>2</sub>O-M9 medium. The temperature in all the steps was 37° C, and the shaking speed and flask sizes were chosen to allow adequate aeration (generally 1:10 for precultures and 1:4-1:6 culture:flask volume for the final cultures). After BL21(DH3) *E. coli* containing the expression vectors with mitochondrial carrier genes reached exponential growth phase (OD<sub>600</sub>  $\approx 0.6$ ), the protein expression was induced by addition of 1 mM IPTG. For GGC1 and AAC3 the proteins were expressed for 3-4 hours at 37° C and for ORC1 at 20° C overnight. We have also used an approach to facilitate the assignment, termed

---

<sup>22</sup>37.3 mM Na<sub>2</sub>HPO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 8.56 mM NaCl, 18.4 mM <sup>15</sup>NH<sub>4</sub>Cl, 0.2% glucose, 1 mM MgCl<sub>2</sub>, 100  $\mu$ M CaCl<sub>2</sub>, 100  $\mu$ M MnCl<sub>2</sub>, 50  $\mu$ M ZnCl<sub>2</sub>, 50  $\mu$ M FeCl<sub>3</sub>, vitamin supplements and either 100mg/l of ampicilin or 30mg/l of kanamycin

”reverse labeling”, which consists of adding one or more amino acids in unlabeled form. ORC1 cultures, where one of the amino acids were unlabelled 100 mg of V, R, K, T, L, I, F, containing natural abundance of isotopes was added 30 minutes prior the induction and the protein was expressed in H<sub>2</sub>O cultures. CH<sub>3</sub>-ALV-[<sup>15</sup>N, <sup>13</sup>C, D] GGC1 containing bacterial pellets were provided to us by Dr. Remy Sounier.

### 3.1.4 Protein purification (*S. cerevisiae* ORC1)

ORC1 purification protocol was largely developed by Dr. Peixiang Ma. Bacterial pellets from 500 ml of cell culture were resuspended in 40 ml sonication buffer (50 mM TrisCl pH 7.0, 1 mM PMSF, 5% glycerol, 1 µg/ml leupeptin, 1 µg/ml pepstatin, 1 mM TCEP and 1mM EDTA). The sample was divided into two parts and each part was sonicated separately to increase the efficiency. As we have noticed formation of higher molecular mass oligomers in ORC1, which we thought were originating due to the oxidation of some of the 4 Cys residues present in ORC1, we have replaced more common reducing agents DTT or β-mercaptoethanol with TCEP, which is a stronger reducing agent, does not form mixed disulfide bonds and is not reversible. EDTA was included in order to scavenge Fe<sup>3+</sup> ions, which are the main catalysts in oxygen-based oxidation of cysteines. Leupeptin and pepstatin are inhibitors of eukaryotic proteases, but he have utilized them to avoid protein degradation due to cross-contamination of the samples, as we were using common equipment in the laboratory.

The pellets were disrupted by sonic energy pulses, allowing long breaks between the pulses to avoid sample heating significantly above 4° C. Although ORC1 accumulated in inclusion bodies and was unlikely to be affected by denaturing temperatures, we were trying to avoid precipitation of some soluble proteins, which could interfere with the further the purification steps. The sonicated bacteria were centrifuged (40000g, 15 minutes at 4° C), and the pellets containing inclusion bodies resuspended in 3 ml of 3% sarkosyl, 1 mM TCEP, 1 mM EDTA and incubated for 2-3 hours at room temperature in the presence of lysozyme and DNase. Although lysozyme could potentially precipitate due to the reduction of its disulfide bonds, it was not observed. After the incubation the sample was centrifuged again to remove the insoluble debris (40000g, 30 minutes, 4° C). The sample was refolded by flash-dilution, slowly adding drops of sarkosyl-dissolved ORC1 to 35-40 ml of 50 mM NaH<sub>2</sub>PO<sub>4</sub>:Na<sub>2</sub>HPO<sub>4</sub> (pH 7.5), 100 mM NaCl and 1.5 mM DPC, 1 mM TCEP, 0.02 mg/ml 1,1',2,2'-tetraoleylcardiolipin buffer at 4° C. Refolding ORC1 at lower dilution factor (than 1:10) would result in the presence of higher molecular weight aggregates in the later purification steps. Refolded ORC1 containing His<sub>6</sub> purification tag was bound to 12 ml of NiNTA resin, previously equilibrated with the refolding buffer and incubated overnight at 4° C by gentle mixing. Afterwards, the resin was washed on a gravity flow column with 10 resin volumes of refolding buffer supplemented with 30 mM of imidazol, where the DPC concentration was increased to 3 mM and TCEP concentration reduced to 0.5 mM in order to avoid the reduction of Ni<sup>2+</sup> ions. The samples were eluted by incrementing imidazol concentration in two steps by applying 1 column volume of refolding buffer containing 200 or 500 mM of imidazol, 15 mM DPC and 0.5 mM TCEP at each step. After this step of purification,

the ORC1 sample was concentrated to no more than 3-3.5 mg/ml (otherwise the higher molecular weight, likely oligomeric peak would appear) and subjected to gel filtration on HiLoad Superdex 200 16/60 prep grade column using 50 mM MES pH 6.5, 100 mM NaCl, 1 mM EDTA, 1 mM TCEP, 1 mM PMSF, 1  $\mu$ g/ml leupeptin, 1  $\mu$ g/ml pepstatin, 3 mM DPC buffer. Typically, the gel filtration profiles indicating the membrane protein preparation of good quality would be obtained (Figure 3.1). The collected fractions were analysed by SDS-PAGE and the purest fractions would be pooled to obtain the final sample. The sample would afterwards be concentrated and desalted on a concentrating tube by using the gel-filtration buffer with no NaCl. Desalting was necessary, as the sensitivity of the cryo-probes of NMR spectrometers decreases sharply with the increasing salt concentration (above  $\approx$ 100 mM salt). The protein yield would be high, around 15 mg of ORC1 per liter of cell culture. Final concentrations of ORC1 used for NMR would be in a range of 500 to 1000  $\mu$ M. It has to be noted that we have used centrifugal tubes of different cut-offs in different purifications (10, 30 or 50 kDa). We could observe that some detergent would always pass through the Centricon filter into the flow-trough. The molecular weight cut-off of a centricon likely affected the amount of the detergent removed (for example measurements of infra-red absorbance of DPC micelles indicate, that DPC can fully pass through the 50 kDa centricon, while it is retained to a similar extent by 10 or 30 kDa centricons). Furthermore, some detergent was added into the desalting buffer (usually 1-1.5 mM DPC) to avoid diminishing of DPC concentration below critical micellar concentration and subsequent "precipitation" of the sample. The precipitation, when occurred, was not typical: instead of forming white flakes the protein would form transparent-looking fibrils. However, as we could observe by NMR, the difference of DPC concentration in the final samples did not affect the appearance of the spectra.

We have further investigated the oligomeric form of obtained ORC1 close to NMR sample conditions (concentration of the protein  $\approx$ 150  $\mu$ M. We have observed a single peak, on AUC profiles, with the observed behaviour in the best agreement with the monomeric state (Figure 3.2).

### 3.1.5 Protein purification (*S. cerevisiae* GGC1)

For GGC1 purification we have used the protocol from Dr. Remy Sounier. Bacterial pellets from 500 ml of cell culture were resuspended in 20 mM HEPES pH 7.5, 150 mM NaCl, 0.1 mM PMSF, 10 mM  $\beta$ -mercaptoethanol buffer and sonicated as ORC1. The pellets were collected by centrifugation at 40000g for 15-30 minutes at 4 $^{\circ}$  C. The resulting inclusion bodies were dissolved in a buffer containing 20 mM HEPES pH 7.5, 150 NaCl, 3M guanidinium hydrochloride, 1% Triton X-100 and 5 mM  $\beta$ -mercaptoethanol overnight at 4 $^{\circ}$  C upon gentle stirring. The insoluble debris were removed by centrifugation at 40000g for 15-30 minutes at 4 $^{\circ}$  C and supernatant applied to 12 mL NiNTA resin in the presence of 10 mM imidazole on a gravity-flow column. The flow-through was reapplied on the column again to increase the fraction of bound GGC1. The columns containing bound protein were washed with 3-times column volume with buffers in the following order: 1) 20 mM HEPES pH 7.5, 150 NaCl, 3M guanidinium hydrochloride, 1% Triton X-100, 5 mM  $\beta$ -mercaptoethanol;

2) 20 mM HEPES pH 7.5, 150 NaCl, 0.1% Triton X-100, 5 mM  $\beta$ -mercaptoethanol;  
3) 20 mM HEPES pH 7.5, 150 NaCl, 3 mM DPC, 5 mM  $\beta$ -mercaptoethanol. This way it could be expected that the denaturant gets removed and the Triton X-100 is replaced with DPC, therefore allowing the protein to refold. The protein was eluted with 3 column volumes of the buffer used in the step 3 of washing, but supplemented with 250 mM of imidazol. Afterwards, the protein was concentrated to  $\approx$ 5 mg/ml with a centrifugal concentration tube and applied on a HiLoad Superdex 200 16/60 prep grade gel filtration column, using 50 MES pH 6, 150 NaCl, 3 mM DPC and 2 mM  $\beta$ -mercaptoethanol buffer. The resulting fractions were analyzed by SDS-PAGE and selected based on purity for preparation of the NMR sample. The protein would be concentrated and desalted, as ORC1. The yield of GGC1 is even higher, amounting to 30-40 mg per liter of cell culture. GGC1mut (Pro138Arg, Pro230Arg mutant) was purified the same way as wild-type GGC1, and the behaviour was nearly identical (the yield was actually slightly higher).

CH<sub>3</sub>-ALV-[<sup>15</sup>N, <sup>13</sup>C, D] GGC1 was purified the same way as wild-type GGC1. After the gel-filtration, the sample was applied to NiNTA column and washed with 5 column volumes of a 50 mM MES pH 6, 100 mM NaCl, 3 mM <sup>2</sup>H-DPC. The sample was eluted with the same buffer supplemented with 200 mM imidazol. Imidazol was removed by multiple steps of concentration and subsequent dilution with a buffer containing no imidazol.

Investigations of GGC1 samples by AUC also indicated that the protein is monomeric at 150  $\mu$ M concentrations (Figure 3.2). However, the sedimentation coefficient appeared to be larger for GGC1 as compared to ORC1, despite the two proteins having very similar molecular weight. This observation might indicate that the hydrophobic regions of the proteins are solvated differently in the presence of DPC micelles.

### 3.1.6 Protein purification (*S. cerevisiae* AAC3mut)

AAC3mut was purified as reported previously (200) by Audrey Hessel, with several modifications to the protocol. Briefly, the cell pellets were sonicated in a 20 mM HEPES pH 7.4, 150 mM NaCl, 30 mM lysozyme and 1mM  $\beta$ -mercaptoethanol buffer. The inclusion bodies were collected by centrifugation at 20000 g for 30 minutes. The inclusion bodies were solubilized in 4ml of 20 mM HEPES, 150 mM NaCl, 1.5% sarkosyl, 0.05% polyethylene glycol 4000 and 1 mM  $\beta$ -mercaptoethanol buffer overnight at 4° C. The insoluble debris were removed by centrifugation at 40000 g for 30 minutes. Supernatant containing solubilized AAC3mut was refolded by flash-dilution into a buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl, 20 mM imidazol and 6 mM DPC, resulting in 10-fold overall dilution. Refolded AAC3mut was applied on NiNTA resin (2 ml for 1 l of cell culture). The protein was eluted with a buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl and 200 mM imidazol. AAC3mut was then dialyzed for 4 hours in 20 mM HEPES pH 7.4, 30 mM NaCl, 3 mM DPC buffer and applied on Hi-trap Q FF column. The non-bound fraction was applied on an analytical gel-filtration column (Superdex 10/300 GL) in the presence of 30 mM MES pH 6, 20 mM NaCl, 5 mM  $\beta$ -mercaptoethanol, 3 mM DPC buffer. The gel-filtration profile contained at least three overlapping peaks. The peak corresponding to the monomer of AAC3mut was re-injected on the gel-filtration column

and the resulting profile contained a single peak with a small shoulder, providing sufficient purity for NMR experiments (Figure 3.1). We could obtain 2 mg from one liter of D<sub>2</sub>O M9 culture.

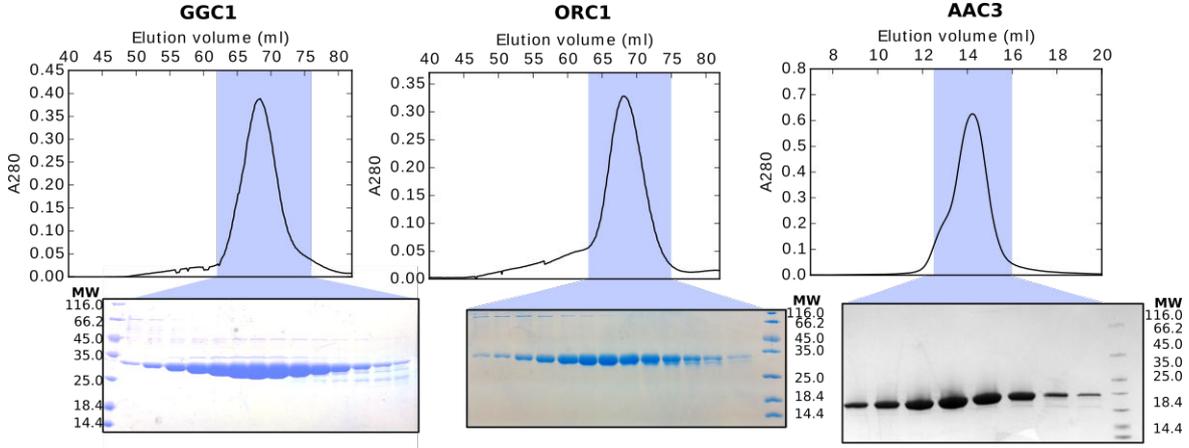


Figure 3.1: An illustration of size-exclusion chromatography profiles (top) and SDS-PAGE gels for GGC1, ORC1 and AAC3. HiLoad Superdex 200 16/60 prep grade columns were used for the two former proteins, while Superdex 10/300 GL column was used for AAC3, therefore the difference in elution volumes. Relatively sharp and symmetrical gel filtration peaks are suggestive of a homogeneous samples. Note that the contrast of SDS-PAGE gel for GGC1 was enhanced to make impurities more visible.

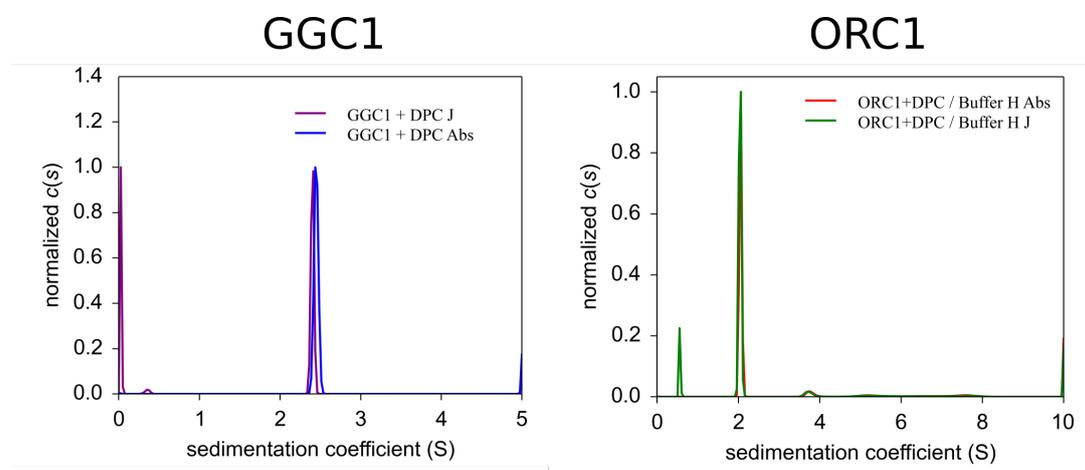


Figure 3.2: AUC data for GGC1 and ORC1. The samples appear to be homogeneous. Note that the sedimentation rate for GGC1 is higher as compared to ORC1. This could indicate different amount of detergent bound.

### 3.1.7 Titration of mitochondrial carrier samples with different molecules

Solutes (GTP, ATP, L-ornithine, L-proline, L-histidine, NaCl, CATR, gadodiamide) used for the titration experiments were dissolved in the flow-through buffer obtained after the last step of sample concentration. Maintaining buffer conditions, which

are as close as possible to the sample buffer is important in NMR, as even slight differences in pH or salt concentration can introduce chemical shift perturbations. The pH was carefully measured after addition of  $\approx 10$  times more of the intended solute to the buffer which did not contain protein, and if necessary was adjusted by NaOH (in all the cases the added solute would only acidify the buffer), so that the pH of the sample containing protein would not be affected. The solute dissolved in the flow-through buffer would then be added to the protein containing sample using accurate eVol (SGE Analytical Science) pipettes.

Cardiolipin and POPG stock solutions were in chloroform. To titrate these lipids, the chloroform was first evaporated under the nitrogen flow and the lipids were dissolved in ultrapure water ( $R=18.2 \text{ M}\Omega \cdot \text{cm}$ ) at known concentration. The NMR sample volume was carefully measured to estimate the required amount of lipid to achieve intended concentration, and an aliquote containing required amount of lipid was freeze-dried overnight. The lipid would then be dissolved in the NMR sample. We have measured if it affected the pH beforehand, dissolving an extra aliquote of lipid in a buffer containing no protein.

### **3.1.8 Reconstitution of mitochondrial carriers in liposomes (Performed by L. Capobianco)**

#### **Standard Protocol**

Inclusion bodies (IB) suspended in a buffer consisting of 10 mM Tris-HCl pH 7.4, were solubilized in 1.67% sarkosyl for 5 minutes at 0 °C. The solution obtained was diluted twenty times with buffer containing 10 mM PIPES, pH 7.0 and centrifuged at 12.000 rpm for 10 minutes at 4 °C. The supernatant (extract) was reconstituted into liposomes. The reconstitution mixture contained purified proteins, 1.14 mg/ml DPG, 1.3% Triton X-114, 1.3% phospholipids (as sonicated liposomes), 5 mM GTP, 20 mM PIPES at pH 7.0 and water to a final volume of 700  $\mu\text{l}$ . These components were mixed thoroughly, and the mixture was recycled through the same Amberlite column (Bio-Rad).

#### **Protocol with DPC**

IB were solubilized with buffer containing 1.67% (w/v) sarkosyl and 10 mM PIPES, pH 7.0. After solubilization at 4 °C for 20 min, the solution was diluted twenty times with a buffer containing 6 mM dodecylphosphocholine (DPC), 10 mM PIPES, pH 7.0 and centrifuged at 12.000 rpm for 10 minutes at 4 °C. For GGC1 in DPC the supernatant was reconstituted as described above.

To check if the protein was inserted into the liposomes (and not just aggregated on the surface) proteoliposomes were treated with an alkaline solution. Both the soluble and proteoliposome fractions were analysed by SDS-PAGE. Aggregated protein would be expected to be removed from the liposomes by neutralizing the charge of the lipids.

#### **Reconstitution procedure**

The recombinant proteins purified to homogeneity from IB is reconstituted into an artificial membrane system (liposomes). Liposomes were obtained by sonicating (Bran-

son 250, Branson Ultrasonics Corporation) egg yolk phospholipids, at a concentration of 100 mg/ml, for 10 minutes at 40 watts at 2 °C, with 1 second pulse followed by 2 seconds pause. Subsequently, a reconstitution mixture consisting of purified protein (100  $\mu$ l with 0.5-1  $\mu$ g of protein), 1.3% Triton X-114, 1.4% phospholipids (as liposomes), 20 mM PIPES, pH 7.0, 10 mM substrate, pH 7.0, is prepared (final volume 700  $\mu$ l). The reconstitution has also required the presence of 1.14 mg/ml DPG (about 3% of phospholipids). These components were mixed thoroughly the Amberlite column (Bio-Rad). Amberlite was pre-equilibrated with 20 mM PIPES at pH 7.0.

All operations were performed at 4 °C, except for the passages through Amberlite, which were carried out at room temperature.

External substrate is removed from proteoliposomes on Sephadex G-75 columns, pre-equilibrated with 20 mM NaCl and 10 mM PIPES at pH 7.0. The same buffer is used for elution. The first 600  $\mu$ l of slightly turbid eluate, containing the proteoliposomes, are collected. The amount of reconstituted protein is assayed by densitometric analysis (Model GS-700 Imaging Densitometer, Bio-Rad) on denaturing polyacrylamide gel, using carbonic anhydrase as protein standard.

## 3.2 NMR experiments

### 3.2.1 Acquisition and processing of NMR spectra

Most of the data was acquired on 600, 700, 850 MHz Bruker spectrometers equipped with cryogenic probes and controlled by Topspin 3.2 software. The spectra were processed using nmrPipe software unless the non-uniform acquisition was applied (as for the 4D spectra), where MDD software (230) was used.

### 3.2.2 Chemical shift perturbation analysis

The assignment of spectra after titration with a particular molecule was obtained after manually comparing the spectra with the spectra of apo protein using CcpNMR software (231). In case of larger chemical shift perturbations, as for cardiolipin, the spectra were recorded at multiple concentrations of the molecule. For GGC1mut assignment we have recorded additional HNC0, HN(CA)CO, HNCACB and  $^{15}\text{N}, ^1\text{H}$ -HSQC-NOESY- $^{15}\text{N}, ^1\text{H}$ -HSQC spectra.

The magnitudes of chemical shift perturbations were calculated using in-house scripts, with the equation:

$$\Delta\delta = \sqrt{(\delta_H^i)^2 + (-4.316/42.576)^2 \cdot (\delta_N^i)^2 + (10.705/42.576)^2 \cdot (\delta_C^{i-1})^2}$$

$\delta_X$  in the equation corresponds to the chemical shift perturbation on a single dimension. The constants are the gyromagnetic ratios of  $\gamma_N$  or  $\gamma_C$  and  $\gamma_H$ .

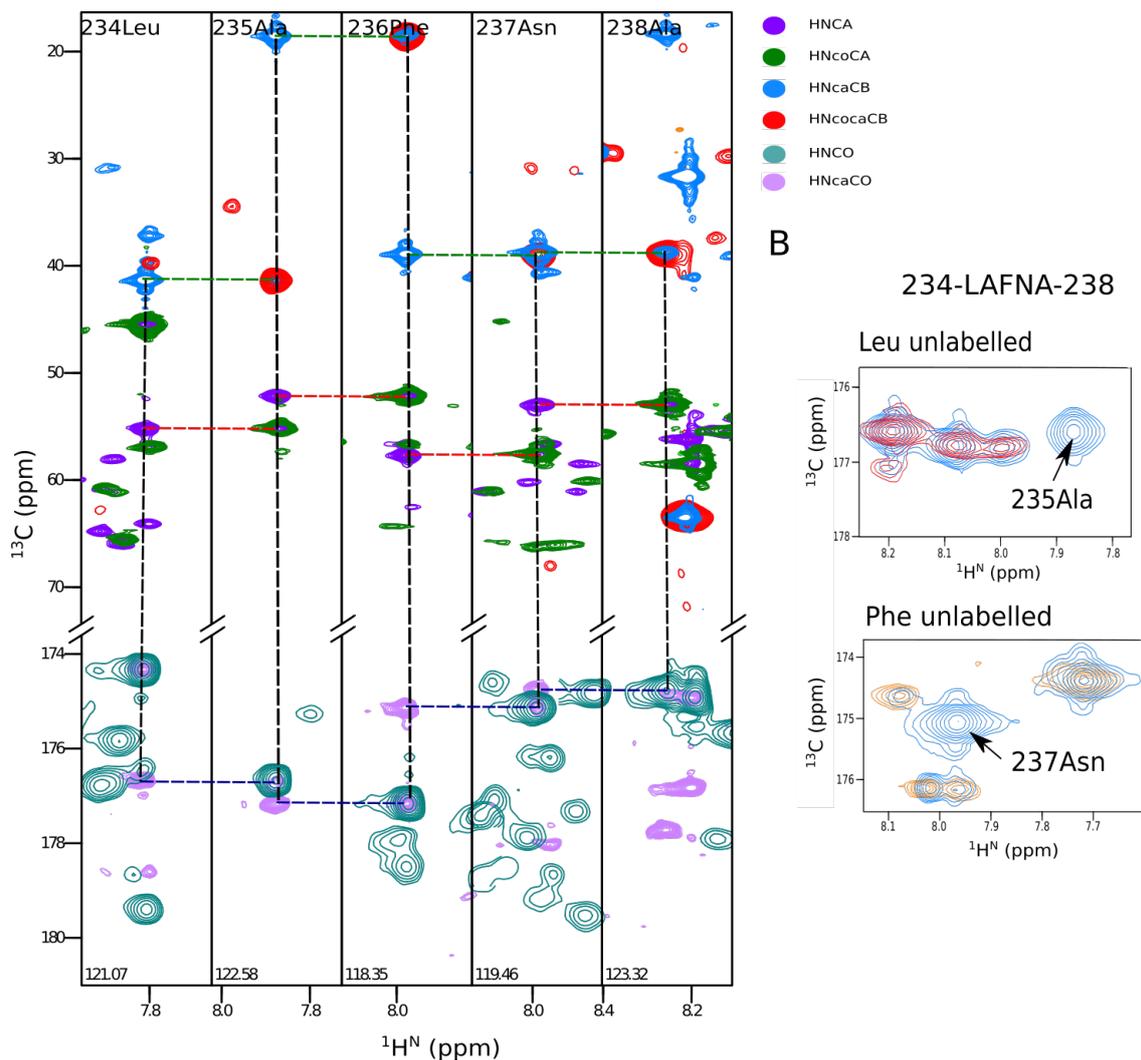


Figure 3.3: An illustration of assignment procedure for ORC1. On the left slices of overlay 3D assignment spectra (HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, HNCO and HN(CA)CO are depicted. All the peaks at a particular  $^1\text{H}$  (abscissa) and  $^{15}\text{N}$  (third dimension) frequencies in each slice correspond to a single amide group of a residue. Multiple peaks are visible as they correspond to different  $^{13}\text{C}$  atoms. HNCA (magenta), HN(CA)CO (pink), HN(CA)CB (cyan) spectra typically contain two peaks with  $^{13}\text{C}$  frequencies of a current and a previous residue. The remaining spectra contain only one peak corresponding to the previous residue. Matching the resonances frequencies of cross-peaks allow to obtain information about the sequential link of the resonances. The nature of the residues can be identified by the frequencies of  $\text{C}_\alpha$  and  $\text{C}_\beta$  (Ala  $\text{C}_\beta$  is at  $\approx 20$  ppm, for example). Another way to identify the residue type in selective unlabelling of some amino acids in the sample (right). Leu-unlabelled sample, for example, does not display a cross-peak of the next residue in sequence in an HNCO spectrum, such as in this case 235Ala. Another example is provided for Phe unlabelled sample.

### 3.3 Assignment of ORC1 backbone resonances

NMR spectra of proteins usually show a large number of resonances. Some information on the protein sample can be obtained by a very qualitative analysis of one-

or multidimensional spectra. However, for the characterization of a protein on a by-residue level, the NMR spectra have to be assigned, i.e. the correspondence between a given resonance frequency with a single NMR active nuclei in the protein has to be revealed. Resonance assignment is usually performed using multidimensional NMR spectra. In such spectra, resonance frequencies from different nuclei are correlated either through bonds by scalar couplings or through space (NOEs, dipolar interaction). These correlations show up as cross-peaks in the spectra whose coordinates correspond to the resonance frequencies of the correlated nuclei.

Protein resonance assignment is completed in two steps: first, backbone resonances are assigned using three-dimensional experiments in which the amide resonances ( $H^N$ , N) of each individual amino acid are correlated via scalar couplings with the carbon resonances ( $C\alpha$ ,  $C\beta$ , CO) of the same or the preceding residue. This information is completed by particular chemical shifts that allow in some cases to identify the residue type (peculiar chemical shifts of glycine  $C\alpha$ , alanine  $C\beta$ , threonine and serine  $C\alpha$  and  $C\beta$ ). Backbone assignment was used for interpretation of all experiments involving  $H^N$ , N and CO nuclei (CPMG,  $R_2$  measurements, ligand binding monitored by chemical shift perturbation of  $H^N$ , N, CO, and amide proton exchange in  $D_2O$ ).

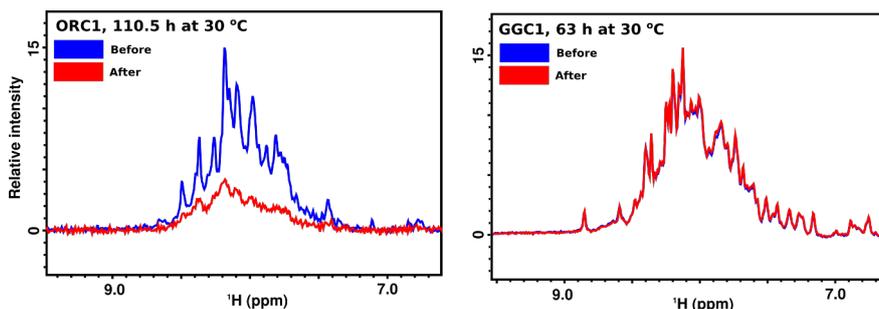


Figure 3.4: An illustration of the sensitivity loss in an ORC1 sample. On the left is a comparison of the same ORC1 sample ( $\approx 1$  mM) before (blue) and after (red) 4.5 day experiment at 30 °C (signal intensity scaled according to the number of scans). Substantial decrease in signal is evident. On the right is a comparison of GGC1 sample of similar concentration after recording an experiment for 2.5 days at 30 °C. Although the experimental time is shorter, there is no decrease in the signal intensity at all for GGC1.

In practice, however, there are several complications. The backbone assignment experiments include correlations between the nitrogen atoms and amide protons (as these type of experiments provide the highest ratio of information content vs. sensitivity), which are present in all residues except prolines. The prolines, therefore, are invisible in such spectra and the lacking information about their resonance frequencies induces breaks while attempting to find the sequential order of resonance cross-peaks. Another complication results from the limited dispersion of the frequencies in the spectra, which in turn results in peak overlap in larger proteins. Substantial peak overlap creates complications in the assignment, as spin systems can no longer be delineated unambiguously thus introducing multiple possibilities for the sequential arrangement of resonance-cross peaks. Additionally, the peaks are broader in large proteins and therefore overlap is more substantial.

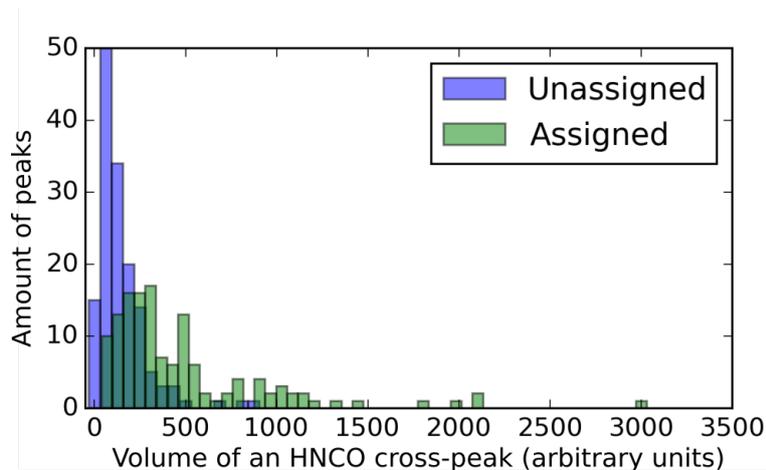


Figure 3.5: A histogram showing the distribution of HNCO peak volumes of assigned and unassigned peaks. Unassigned peaks are shifted to the lower volumes, indicating their lower intensity in an HNCO spectrum and thereby faster  $R_2$  relaxation. As HNCO is the most sensitive experiment in the whole assignment series, faster relaxation results in more dramatic effects in other experiments, leading to missing cross-peaks in the spectra.

The third complication arises if there are millisecond dynamics in a protein. These dynamics induces increased transverse relaxation and peak broadening such that some resonance cross-peaks become very weak or even invisible in the NMR spectra. The spectra acquired with mitochondrial carriers suffer from all of the aforementioned limitations. For the backbone assignment of ORC1 (illustrated in Figure 3.3) we acquired multiple spectra in order to extract sequential information. We have thus recorded BEST-TROSY versions of HNCA and HN(CO)CA ( $C\alpha$  correlations) HN(CA)CB and HN(CO)CACB ( $C\beta$  correlations), HN(CA)CO and HNCO (carbonyl correlations) and  $^{15}\text{N},(^1\text{H})\text{-HSQC-NOESY-}^{15}\text{N},^1\text{H-HSQC}$  spectra (through-space correlations between amide protons of residues close in space). These types of spectra are commonly used for backbone resonance assignments of proteins, however in the case of ORC1, they were not sufficient for the full assignment (one of the main reasons was missing resonance cross-peaks in the most informative, but least sensitive experiments, such as HN(CA)CB, where we could observe only 60% of resonances originating from  $C\beta$  atoms or HN(CA)CO were only 50% of the expected cross-peaks were visible. Even in HNCA experiments we identified only 80% of the expected cross-peaks). Therefore, we have attempted to get additional information about some residue types by performing experiments on samples with residue-specific unlabelling. In these samples, all amino acids were isotopically enriched with  $^{15}\text{N}$  and  $^{13}\text{C}$ , except either Ile, Leu, Phe, Thr, Val, Lys or Arg (compare Figure 3.3B). These residues were chosen because they are substantially abundant in ORC1 and are not subjected to metabolic scrambling (i.e. there are no metabolic pathways in *E. coli* which would convert these amino acids into any other kind of metabolic intermediate. Note however that this does not apply to amide nitrogen groups that are always scrambled due to the action of transaminases). Unfortunately, even these experiments did not provide enough information to improve the assignment. Finally,

we have recorded 4D HNCACO and HNCOCA experiments, but even these experiments were not sufficient to obtain more than 35% of assignment of the backbone resonance cross-peaks of ORC1.

The reasons for such complications in the assignment could have arose due to millisecond dynamics present in unassigned regions. Indeed, investigating the intensities of unassigned peaks reveal, that unassigned cross-peaks generally have smaller peak volumes in the HNCOC spectra (Figure 3.5) and more importantly, the low stability of the sample in our NMR conditions (Figure 3.4).

### 3.3.1 Assignment of methyl groups of GGC1

Backbone assignment of GGC1 was already available in Biological Magnetic Resonance Data Bank (228).

Mitochondrial carriers with their  $\approx 300$  residues are quite large proteins from an NMR point of view and analysis of three-dimensional spectra such as HNCOC is hampered by limited resolution and low signal intensity, as described above. Therefore, we also used sidechain methyl groups as specific probes for some NMR studies of mitochondrial carriers. In our group, much effort was spent on the development of specific  $^{13}\text{C}^1\text{H}_3$ -labelling of different methyl groups in an otherwise perdeuterated protein. Analysis of methyl signals offers some advantages: high signal intensity due to the presence of three equivalent protons, narrow lines because of the rapid rotation around the C-C bond and generally good chemical shift dispersion. Methyl resonances also have to be assigned to their specific residues. Classical sidechain resonance assignment takes advantage of the very efficient  $^{13}\text{C}$ - $^{13}\text{C}$  TOCSY transfer in CC(CO)NNH or HCCH-TOCSY like experiments. In these experiments, magnetization is transferred from sidechain protons onto carbons and spread along the entire sidechain before being detected on the amide group of the preceding residue or on individual sidechain CH-groups. These experiments do not work on large proteins, such as the mitochondrial carriers, which require deuteration.

In order to overcome this limitation, methyl assignment of Ala, Leu and Val sample was obtained using a set of 2D  $^{15}\text{N}$ -HSQC, 2D  $^{13}\text{C}$ -SOFAST, 3D  $^{13}\text{C}$ -HSQC-NOESY- $^{15}\text{N}$ -HSQC and  $^{15}\text{N}$ -HSQC-NOESY- $^{13}\text{C}$ -HSQC experiments. The NOESY spectra were acquired as 3D spectra without incrementing the indirect  $^1\text{H}$  dimension. The mixing time for NOESY experiments was 278 ms. To avoid the interference of aliphatic proton signals from DPC, we have used deuterated DPC with all  $^1\text{H}$  atoms exchanged to  $^2\text{H}$ . We could assign most of the Ala and all of the Val methyl groups. For Leu residues, however, the assignment was complicated by the absence of NOE cross-peaks between the methyl carbon ( $\text{C}\delta$ ) and amide nitrogen for most of the Leu residues. This is likely because of the larger distance in space between these two chemical groups and possibly because of the flexibility of Leu side-chain. The assignment procedure is illustrated in Figure 3.6 and the spectra of assigned methyl groups in Figure 3.7.

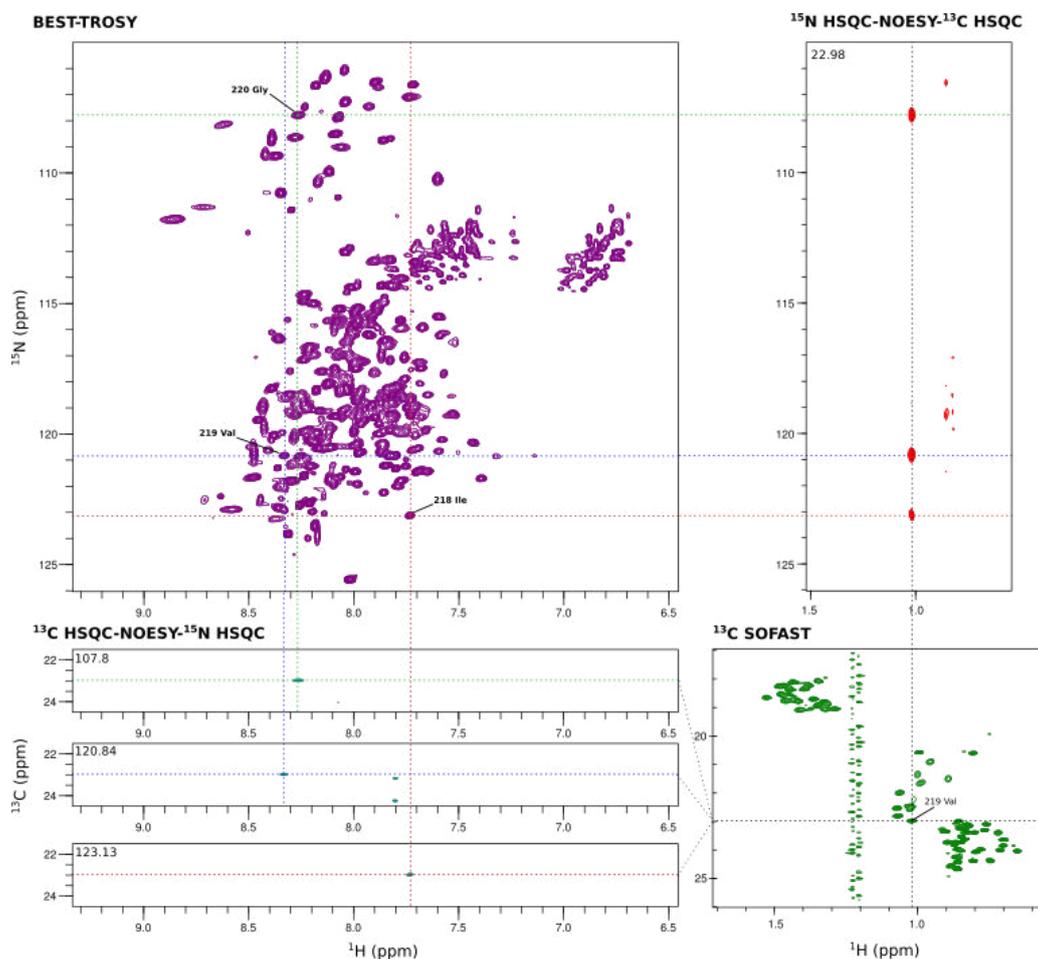


Figure 3.6: Illustration of the assignment procedure for methyl group labelled GGC1. 2D  $^{15}\text{N}$ -HSQC (A) and  $^{13}\text{C}$ -SOFAST-HMQC (D) spectra are illustrated in magenta and green, respectively. Observed correlations in these spectra correspond to  $^{15}\text{N}$ ,  $^1\text{H}$  amide and  $^{13}\text{C}$ ,  $^1\text{H}$  methyl groups, respectively. Methyl groups were then assigned by correlating the previously assigned  $^1\text{H}$ ,  $^{15}\text{N}$  backbone resonances with the  $^1\text{H}$ ,  $^{13}\text{C}$  methyl group resonances using two complementary 3D NOESY spectra. In the  $^{15}\text{N}$ -HSQC-NOESY- $^{13}\text{C}$ -HSQC spectrum, magnetization is transferred from amide protons to methyl groups close in space. The amide proton frequency is not edited and it is the amide nitrogen frequency that is correlated via the  $^1\text{H}^{\text{N}} - ^1\text{H}_{\text{methyl}}$  NOE to methyl groups, as shown in B. In the complementary  $^{13}\text{C}$ -HSQC-NOESY- $^{15}\text{N}$ -HSQC experiment, it is the  $^{13}\text{C}$  frequency of the methyl group which is connected via the  $^1\text{H}_{\text{methyl}} - ^1\text{H}^{\text{N}}$  NOE to backbone amide groups (C). The combined use of information provided by these two spectra allows unambiguous through-space connection of individual methyl groups with backbone amides, as depicted by the coloured lines. These connection in combination with  $^{13}\text{C}$  chemical shifts yielded assignment of nearly all Val and Ala methyl groups in GGC1.

### 3.3.2 CPMG data collection

All CPMG data was collected at 30° C. TROSY methods were used to optimize transverse relaxation and most of the data was recorded as a 3D HNC0 type experiments to avoid peak overlap and allow more accurate quantification of the peak volumes.

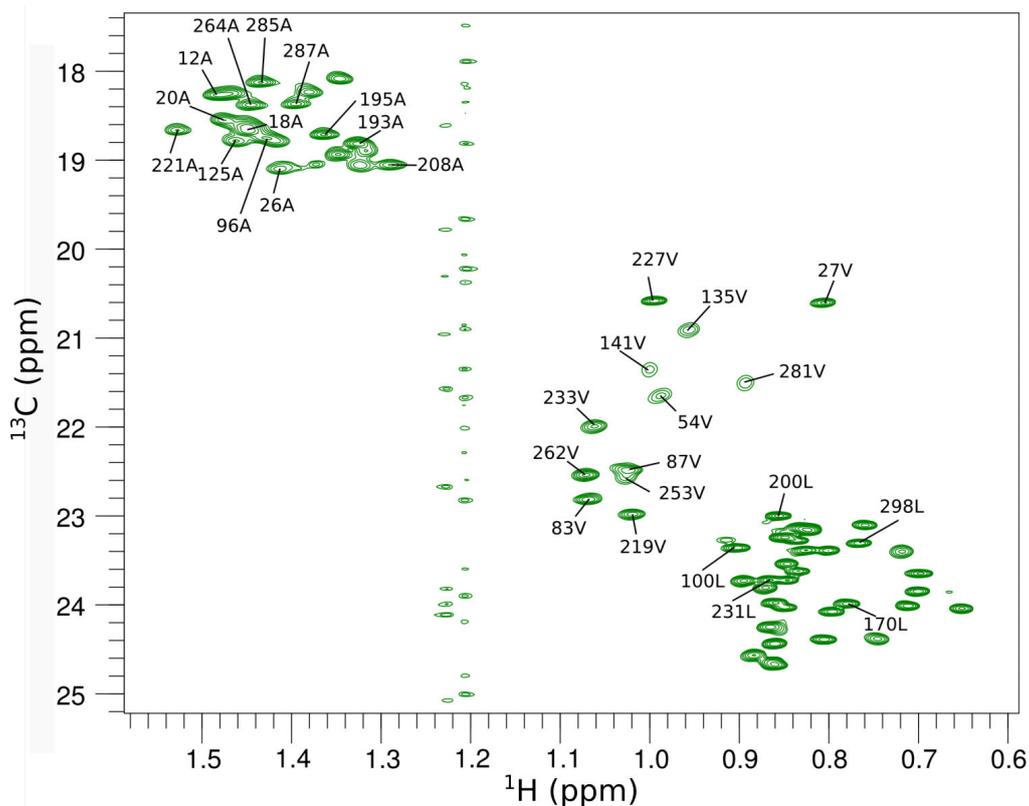


Figure 3.7: An illustration of  $^{13}\text{C}$ -SOFAST-HMQC spectrum with residue names indicating the assigned cross-peaks.

The pulse sequence was used similar to ref. (200), except that we have used shaped pulses to avoid excitation of aliphatic protons (232). Reference spectra for the calculation of  $R_{2,\text{eff}}$  were recorded in the absence of CPMG delay. 20 ms CPMG delays were used to record following CPMG data: 600, 700, 850 MHz GGC1-CH<sub>3</sub>, 600, 950 MHz AAC3mut, 600, 850 MHz GGC1<sup>2P→2R</sup> mutant with cardiolipin, 700, 850 MHz GGC1 with cardiolipin, 700 MHz GGC1 with cardiolipin and GTP.

While 40 ms delays for 600, 850 MHz GGC1 with GTP, 750, 850 MHz apo GGC1, 600, 850 MHz apo GGC1<sup>2P→2R</sup> mutant, 600 MHz GGC1 with cardiolipin, 600 MHz GGC1 with cardiolipin and GTP, 600, 700 MHz ORC1. CPMG frequencies in each case were between 50 and 1000. We could not observe the difference in the relaxation curves due to a different delay length, and therefore utilized shorter delays in the later experiments to reduce the effect of transverse relaxation.

### 3.3.3 CPMG data analysis

CPMG spectra were processed using nmrPipe. The peak volumes were quantified using nmrView Java software.  $R_{2,\text{eff}}$  was calculated using the equation:

$$R_{2,\text{eff}} = -1/T_C \ln(I_{\nu\text{CPMG}}/I_0)$$

$I_0$  - the volume (intensity) of an individual cross-peak extracted from the spectrum recorded without the CPMG delay while  $I_{\nu_{CPMG}}$  corresponds to the intensity of the same cross-peak at a particular CPMG frequency.

We have then selected residues displaying non-flat relaxation dispersion profiles by visual selection for global fitting of the data, which displayed good quality (not too noisy) dispersion curves on two or more fields used. Fitting by numerical calculation of Bloch-McConnell equation was performed using ChemEx software (version 0.6.0) (233). The software was developed by Guillaume Bouvignies and is available on <https://github.com/gbouvignies/chemex>. Amide CPMGs were fit using `n_trosy` method, which is suited to analyze  $^{15}\text{N}$  TROSY CPMG data. `ch3_mq` method was applied for the methyl data.

For amide data, the length of  $180^\circ$  CPMG pulses was defined as  $1\ \mu\text{s}$ . Equilibration time (used to compensate the difference between the  $R_2$  rate of two states) was set as 5.0 ms. CPMG delay  $T_{CP}$  was set as during the recording of experimental data, as well as the number of CPMG cycles used. Transverse and longitudinal relaxation rates of  $^{15}\text{N}$  were set to arbitrary numbers. HN coupling constant was set to -93 Hz. Chemical shift values for nitrogen atoms were extracted from the assignment table.

### 3.3.4 Jackknife analysis

As the fit parameters ( $|\Delta\omega|$ ,  $k_{ex}$  and  $p_B$ ) were varying between the fits of individual residues, we wanted to find some confirmation that the observed process is two-state exchange. For this we have used jackknife approach before fitting the data. A pool of residues showing non-flat relaxation dispersion profiles was randomly sampled taking 60-70% of residues for fitting. The sampling was performed by in-house scripts. Typically between 50 to 500 sets of randomly grouped residues were taken and fitted with ChemEx to a two state exchange model.



## 4 Results: Dynamics, Interactions and Structure of Mitochondrial Carriers in DPC detergent

### 4.1 Three different mitochondrial carriers undergo extensive millisecond dynamics

One of the key questions related to the function of mitochondrial carriers (MC) is the interconversion between different states, open to the matrix or the cytosolic side of the membrane (m- and c-state, respectively). As the time scale of transport, in lipid bilayers, has been shown to be in the millisecond regime, we wanted to probe motions on this time scale, as such motions may provide direct insight into functional turnover. We have studied in this study three different mitochondrial carriers from yeast, the ornithine carrier, ORC1, the GDP/GTP carrier, GGC1, as well as (a mutant of) the ADP/ATP carrier, AAC3. GGC1 as well as AAC3 have been studied

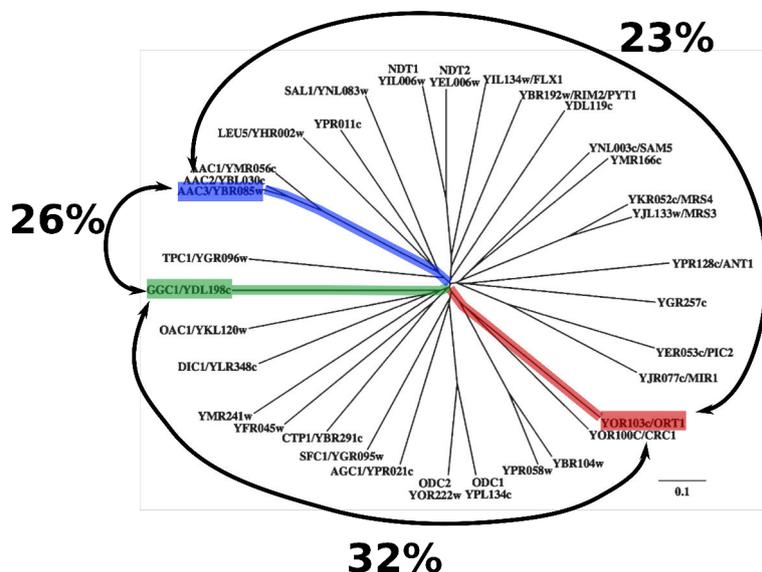


Figure 4.1: Phylogenetic relationship between yeast mitochondrial carriers. AAC3 is colored in blue, GGC1 in green and ORC1 in red. The numbers indicate sequence identity between the three carriers.

before (200, 228).

Figure 4.1 shows the phylogenetic relationship of mitochondrial carriers, which reveals that the three proteins we studied belong to different evolutionary branches of yeast MCs, and that their sequence identity is around 30% only. Studying these diverse proteins is interesting because the comparison may reveal common traits, which may be general to the entire MC family.

We have prepared the three proteins with  $^2\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  isotope labeling using a refolding procedure from inclusion bodies in DPC micelles (see Materials and Methods section 3.1.5). All three proteins yield spectra that display a single set of peaks (Figure 4.S1). The chemical shift dispersion of amide peaks, ranging from about 7.5 ppm to approximately 9 ppm is relatively narrow, which is commonly observed for

$\alpha$ -helical membrane proteins in detergent. The fact that a single set of resonances is observed for all three proteins shows that the proteins are structurally homogeneous, or that at least the interconversion of different states, if present, is faster than  $\approx 10$  milliseconds<sup>23</sup> or that possibly existing minor states are populated to only a small extent. That the proteins display homogeneous behavior is furthermore evidenced by the observation of a single sedimentation coefficient in analytical ultra-centrifugation (see Figure 3.2), Materials and Methods) and a homogeneous gel-filtration profile (Figure 3.1, Materials and Methods). We were able to obtain almost full assignment of GGC1 by transferring the peaks from GGC1 shift list previously deposited on the Biological Magnetic Resonance Databank (BMRB) (228) while ORC1 had to be assigned manually. Only  $\approx 35\%$  of ORC1 cross-peaks could be assigned despite large effort undertaken.

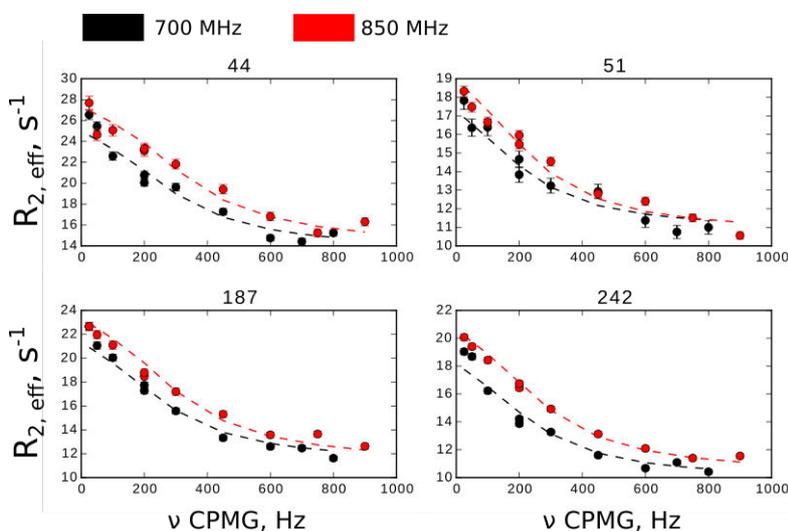


Figure 4.2: Exemplary CPMG relaxation dispersion profiles for several residues in GGC1. Black and red dots represent the data obtained at two different magnetic field strengths.

#### 4.1.1 GGC1 undergoes millisecond dynamics

In order to obtain more insight into the dynamics of these proteins, we next performed Carr-Purcell-Meiboom-Gill (CPMG) relaxation-dispersion (RD) experiments, using a 3D HNCO-variant, because significant resonance overlap in 2D spectra would otherwise hamper site-resolved analysis. Figure 4.2 shows exemplary RD profiles of backbone  $^{15}\text{N}$  sites in GGC1, obtained at two different static magnetic field strengths. Sizeable (non-flat) CPMG RD profiles, which are indicative of  $\mu\text{s}$ - $\text{ms}$  dynamics, are found for approximately 20% of the residues in GGC1. Interestingly, the residues which display exchange are located in only one half of the molecule, comprising helices 2 to 5 (see Figure 4.3A). We furthermore aimed at obtaining information also

<sup>23</sup>Exchange between different states on slower time scales would result in the observation of individual cross-peaks for each state.

Table 4: CPMG fit parameters for GGC1 under different conditions

Protein	Exchange rate, $s^{-1}$	Population (minor state), %
GGC1	1157±104	1.91±0.12
GGC1 (CH <sub>3</sub> )	1160±60	19±12
GGC1+GTP	1387±118	3.33±0.94
GGC1 <sup>2P→2R</sup>	1486±84	3.4±0.9
GGC1+CL	516±119	9.6±3.33
GGC1+CL+GTP	388±74	4.5±0.8
GGC1 <sup>2P→2R</sup> +CL	483±32	5.3±0.36

Table 5: CPMG fit parameters for AAC3

Protein	Exchange rate, $s^{-1}$	Population (minor state), %
AAC3	1500±166	1.17±0.34
AAC3 <sup>c-salt-bridge</sup>	1200±200	1.17 (fixed)
AAC3 <sup>TM</sup> (200) <sup>24</sup>	870±200	2±0.4
AAC3 <sup>matrix</sup> (200) <sup>25</sup>	940±170	2±0.4

about the side chains in GGC1, and labeled the Val ( $\gamma$ 1) Ala ( $\beta$ ) and Leu ( $\delta$ 1) with <sup>13</sup>CH<sub>3</sub> in an otherwise <sup>12</sup>C, <sup>2</sup>H labeled background, and performed multiple-quantum <sup>1</sup>H-<sup>13</sup>C CPMG RD experiments (234). We observed sizeable RD for  $\approx$ 15% of methyl sites (eight methyl groups; see Figure 4.S2). These methyls are located in the same part of the protein as the backbone sites for which exchange is detected (see Figure 4.3A).

In order to obtain quantitative insight into the underlying exchange process, we performed numerical fits of these RD profiles. The most common model for such a fit is a two-state model, whereby it is assumed that the exchange occurs due to the presence of a second conformation with a population  $p_B$ , that differs from the first state by a chemical-shift difference  $|\Delta\omega|$ , at an exchange rate of  $k_{ex}=k_{AB}+k_{BA}$ . In principle the dispersion of each residue could be fitted individually. However, when assuming that the underlying millisecond dynamics concerns several residues, one can attempt to fit several residues to one common exchange rate  $k_{ex}$  and minor-state population  $p_B$ , as well as individual chemical-shift differences  $|\Delta\omega|$ . This assumption needs to be validated, because whether a dynamic event is global or local has important consequences for interpretation. The rationale for assuming that the exchange process is a common one comes from the realization that a continuous structural part of GGC1 is concerned (Figure 4.3A).

In order to validate the common 2-state exchange model, we performed extensive jackknife simulations: randomly sampled sub-sets of residues were fitted separately, and the convergence of the fit parameters was investigated. This statistical analysis, presented in figure 4.4A reveals that the population and exchange rate for different sub-clusters is the same, supporting a view where all residues experience similar exchange dynamics. The fits reveal a minor-state population of about 2% and an exchange rate of 1200  $s^{-1}$  (Table 4).

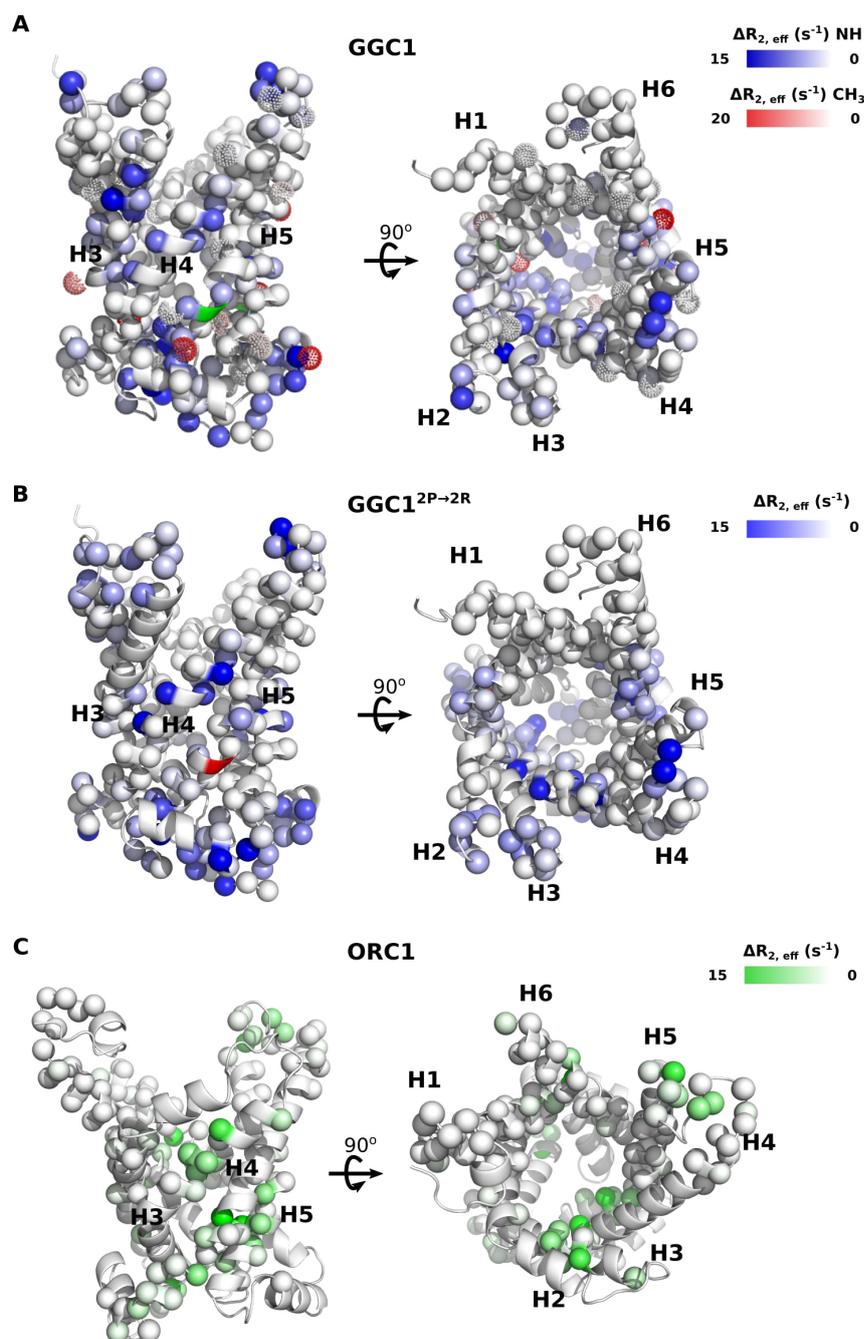


Figure 4.3: Distribution of residues displaying millisecond exchange on the homology models of ORC1 and GGC1, side (left) and top (right) views. Spheres represent assigned residues, for which CPMG data could be measured. The data was obtained on 850 MHz spectrometer. The horizontal color bars illustrate the size of observed dispersions ( $R_{2, \text{eff}}(\text{max}) - R_{2, \text{eff}}(\text{min})$ ). White spheres correspond to flat relaxation dispersion profiles. A - ORC1, B - GGC1. Dotted spheres indicate CPMG relaxation dispersions measured for Ala, Leu and Val methyl groups. A region colored in green correspond to one of the proline kinks. C, GGC1<sup>2P→2R</sup> mutant. Red color indicated one of the mutation sites.

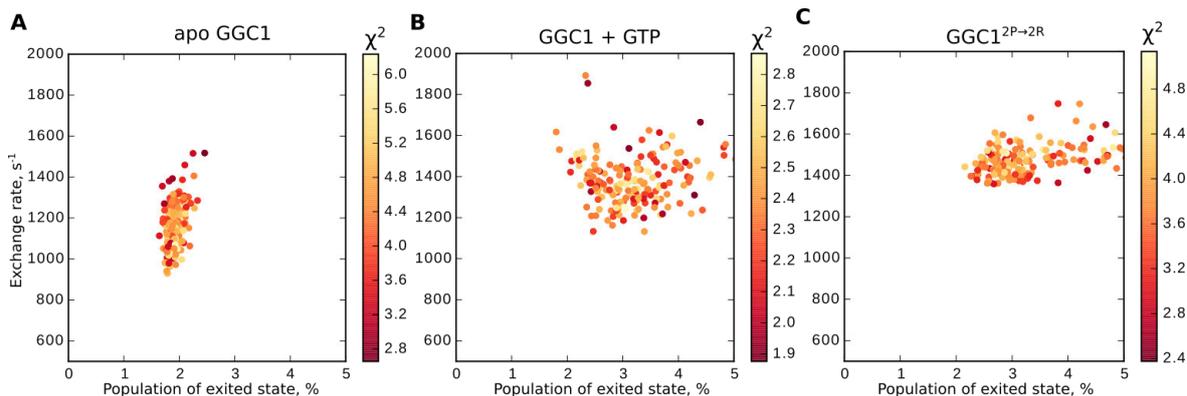


Figure 4.4: Results of RD curve fitting by jack-knife analysis of three different GGC1 samples.  $\approx 200$  fits are represented in each plot, where 60% of residues from a total pool of  $\approx 20$  residues displaying exchange were taken. The populations of the minor state and exchange rates are almost identical in all three cases. A, apo wild-type GGC1. B, wild-type GGC1 in the presence of 20 mM GTP. C, GGC1<sup>2P→2R</sup> mutant.

#### 4.1.2 Millisecond dynamics are also present in ORC1 and AAC3

It is interesting to see that the other two mitochondrial carriers reconstituted in DPC, AAC3 and ORC1, feature very similar dynamic properties. For the case of AAC3, Brüscheiler et al very recently reported also the presence of asymmetrically distributed dynamics, whereby residues located in helices 6, 1 and 2 show exchange (Figure 4.5A). In the absence of inhibitors or substrates, a common exchange model appears to fit the data: although in their analysis the residues in the TM helices and the matrix helices were fitted separately, the reported fit parameters from fitting the two groups are very similar ( $p_B = 2\%$  in both cases,  $k_{ex} = 870 \pm 200 \text{ s}^{-1}$  and  $940 \pm 170 \text{ s}^{-1}$ , respectively, Table 5). It is noteworthy here, however, that the reported fit parameters in Brüscheiler et al are quantitatively wrong, i.e. not in agreement with their experimental data, as discussed further below. We were able to obtain a part of the original data from Bruscheiler et al. Fitting this data with our approach resulted a much better agreement with the data and fit parameters (compare Fig S3) in an exchange rate of  $1500 \pm 166 \text{ s}^{-1}$  and the populations of minor state of  $1.17 \pm 0.34\%$  (Table 5;  $|\Delta\omega|$  values are reported in Table 4.S1).

Similarly, in ORC1 we also observe clear dispersions (Figure 4.3C). The spectral quality of ORC1 is generally lower, and we only have limited assignment. Nonetheless, the pattern that arises from the  $\approx 35\%$  of assigned residues also clearly shows an asymmetrical distribution of residues undergoing millisecond exchange. Due to lower data quality, the convergence of fits was less good than for GGC1. A jackknife analysis, similar to the one presented for GGC1 shows that the exchange rate constant is on the order of  $1000 \pm 400 \text{ s}^{-1}$ , and a minor population of  $1 \pm 0.36\%$ , again similar to both GGC1 and AAC3.

<sup>25</sup>Values reported by Bruscheiler et al. for the TM helices

<sup>25</sup>Values reported by Bruscheiler et al. for the matrix helices

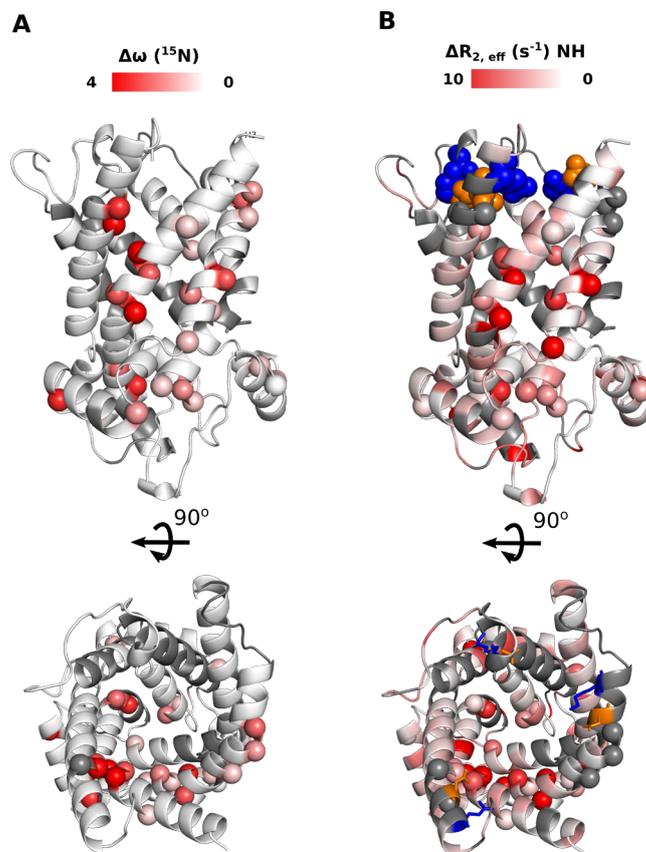


Figure 4.5: Distribution of residues displaying millisecond exchange on the crystal structure of AAC3 (190), where loops invisible in the X-ray structure have been modelled. Spheres correspond to the residues for which non-flat CPMG relaxation dispersion profiles were obtained by Bruschiweiler et al. Color bars indicate the chemical shift difference obtained after a global fit of all the residues displaying millisecond exchange. A, the data obtained for AAC3\* by Bruschiweiler et al. B, The data obtained by us on AAC3<sup>c-salt-bridge</sup> mutant. Orange and blue spheres correspond to the cytoplasmic salt-bridge network residues, which were mutated in AAC3<sup>c-salt-bridge</sup> mutant.

## 4.2 Substrates have very small impact on dynamics

We thus observe similar dynamics in all three mitochondrial carriers, despite the low sequence identity, and the different nature of the substrates they carry. In all three proteins there is an exchange process on a time scale of approximately 0.5-1 ms, which is similar to the exchange dynamics in membranes ( $\approx 2$  ms (204)).

If this dynamic process corresponds to a functionally relevant one, it might be altered by the presence of substrate. We have thus investigated the substrate effect on the observed millisecond dynamics in GGC1. Surprisingly, however, the presence of 20 mM GTP in a  $\approx 0.7$  mM GGC1 sample has, overall, only very small effects on CPMG RD profiles (Figure 4.S6). The most affected residues do not lie near the Pro kinks, as could be expected from the suggested model of exchange, but instead were located in peripheral regions of the protein. Among those, Ser206 is present in a loop

Table 6: CPMG fit parameters for ORC1

Protein	Exchange rate, $s^{-1}$	Population (minor state), %
ORC1	$1030 \pm 407$	$1.06 \pm 0.36$

linking H4 and H5 and Ser152 is located in mitochondrial matrix loop between H3 and H4.

A similar jackknife analysis as was done for the apo protein shows that the exchange parameters are only slightly different from the apo sample (Figure 4.4B and Table 4). We can see that the minor state populations appear to be slightly larger for GGC1 in the presence of GTP. The difference, however, is also reflected in lower  $|\Delta\omega|$  values for GGC1 in the presence of GTP (Figure 4.S4). Lower chemical shift differences would indicate that GTP induces a substantial change in either c-state or m-state, or both.

However, the populations and  $|\Delta\omega|$  are correlated and are difficult to accurately deconvolve for the cases where the exchange rate is fast. We therefore think that the observed differences in  $p_B$ s are not relevant, given very similar RD profiles for both samples (see Figure 4.S6).

It has to be noted that we have used in this jackknife analysis the subset of residues for which data of both samples were available. However, if we use all the residues in the analysis of each sample, we obtain very similar fit parameters, reflecting that the choice of residues is not significantly affecting the outcome of the fit (Table 4.S2).

For yeast AAC3 it has been reported that the addition of ADP does not have an effect on the populations, but it has a small effect on the exchange rate constant. (ADP *in vivo* transported from the cytosol to the matrix, equivalent to the role of GTP for GGC1.) In the analysis of Brüsweiler et al, the TM helices and matrix helices were fitted separately, and the exchange rate constant for the TM helices was reported to increase from  $870 \pm 200 s^{-1}$  to  $1800 \pm 350 s^{-1}$  upon addition of 10 mM ADP. In addition to the ADP experiments, Brüsweiler et al also reported CPMG experiments with a very strong inhibitor, carboxyatractyloside (CATR). They report a decrease of the exchange rate constant of TM helix residues from  $870 \pm 200 s^{-1}$  to  $150 \pm 110 s^{-1}$ .

Together, these findings were interpreted as a change of the transition-state energy barrier, where the substrate lowers this barrier whereas the inhibitor increases it.

It is surprising that CATR, a very strong inhibitor with nanomolar affinity, only resulted in very modest change of dynamics. Furthermore, we want to stress that it is not clear at this point whether the fit parameters reported in (200) are significant: in fact, when plotting simulated CPMG curves onto the experimental data reported in this study, it becomes evident that the fit parameters do not describe the data (see Figure 4.S3B). We have been able very recently to obtain the original data from the authors of this study (only data of the apo form), and re-fitted these data. The exchange rate constant and the chemical-shift differences for the apo form are very different from the reported values (Figure 4.S3A, Table 5), such that it is not clear at this point whether there is a significant difference in the different samples reported.

To conclude, the presence of substrates only has very minor effects on the dynamics of mitochondrial carriers in DPC micelles. We next wanted to investigate whether mutations which severely impact the function result in alteration of the millisecond dynamics.

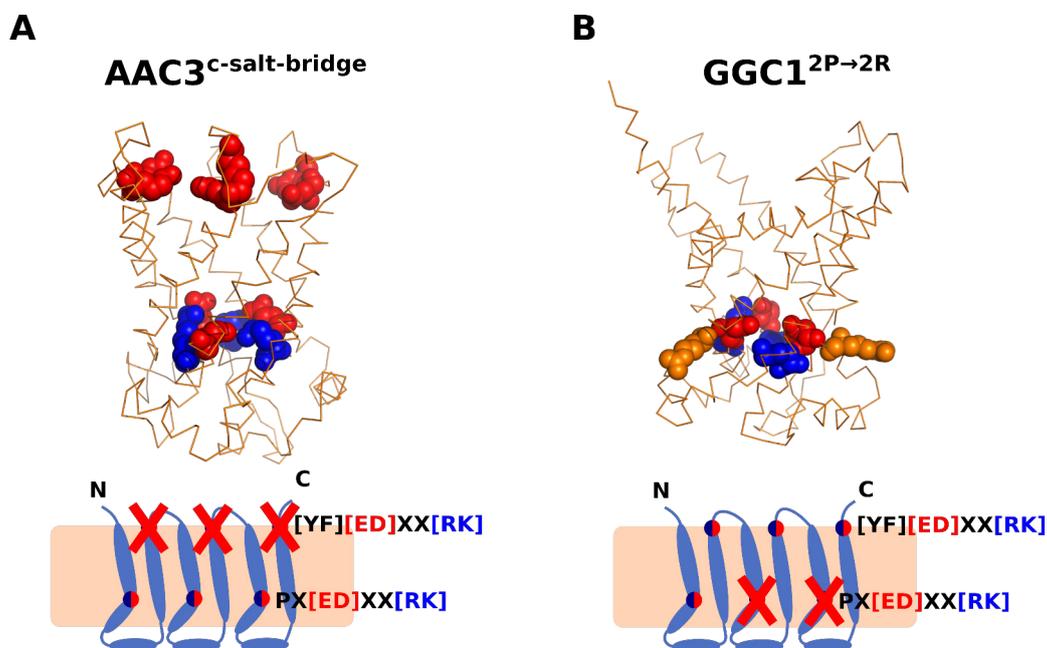


Figure 4.6: Inactivating mutations introduced to AAC3 and GGC1 MCs. A, the cytosolic salt bridge network in AAC3 was disrupted by replacing 3 positively charged residues in [YF][RF]XX[RK] motif to negatively charged amino acids (shown in red on top of the molecule). We refer to this mutant as AAC3<sup>c-salt-bridge</sup> mutant. In this figure the matrix salt bridge network is also shown in blue and red. B, the Pro kinks were disturbed in GGC1, by replacing Pro138 and Pro230 with Arg residues. As the proline residues are located close to the matrix salt bridge network and may be important in the conformational exchange of GGC1, we expected these substitutions to inhibit the function of the protein.

#### 4.2.1 Non-functional mutants display essentially identical dynamics as WT proteins

We designed two different mutants, which are expected to disrupt function. The first mutant we designed contains substitutions of two strongly conserved prolines (Pro138 and Pro230) to arginines (Figure 4.6B). These prolines are present in the three-fold repetitive PX[ED]XX[KR] motif, which due to its conservation among the mitochondrial carriers is known as mitochondrial carrier signature motif. The exact role of Pro residues is not known, however proximity to the matrix salt-bridge network, where they introduce sharp kinks into odd-numbered  $\alpha$ -helices has been implicated to be important for the conformational exchange of the carriers (221). We therefore expected the substitutions to be sufficient for the inhibition of the conformational exchange of GGC1. In the following we refer to this mutant as GGC1<sup>2P→2R</sup> mutant. The second mutant, a mutant of AAC3, addresses the cytosolic salt-bridge network,

Table 7: Substrate transport assays of MCs performed in proteoliposomes on wild-type and mutant proteins

Protein	Transport rate ( $\mu\text{mol} \cdot \text{mg protein}^{-1} \cdot 5 \text{ min}$ )
AAC3	$630 \pm 31.5$
AAC3 <sup>c-salt-bridge</sup>	$4.66 \pm 0.23$
GGC1	$1200 \pm 37$
GGC1 <sup>2P→2R</sup>	$26 \pm 4$

which in AAC3 is formed by residues 98D/101K, 201D/204K and 295D/298Q (Figure 4.6A). Although these residues do not form salt bridges in the known crystal structures, corresponding to the c-state, this network is thought to form in the m-state. Indeed, it has been shown before (190, 194) that mutants that cannot form this salt-bridge network have strongly reduced transport. We refer to this mutant in the following as AAC3<sup>c-salt-bridge</sup> mutant.

We first established that these mutations indeed disrupt the functionality of the carriers. Liposome transport assays were performed by our collaborators Paola Lunetti and Loredana Capobianco (University Salento, Italy). These experiments consist of reconstituting mitochondrial carriers, refolded from inclusion bodies in Triton-X100 detergent, into liposomes. Transport was then monitored by radioactivity measurements, using radiolabeled substrates (see Materials and Methods section 3.1.8 for details).

The results, shown in Table 7 and Figure 4.9B establish that these mutants indeed have negligible residual transport, compared to control experiments performed in parallel on WT proteins. It can be assumed that the exact mechanism by which these proteins become inactive may differ: while in GGC1<sup>2P→2R</sup> the proposed "hinge" regions are targeted, in AAC3<sup>c-salt-bridge</sup> mutant it is the salt bridge network that stabilizes the m-state which is disrupted. One may thus assume that the two mutants behave differently from the wild-type proteins in their dynamics, and that they also may have this different behavior for different reasons.

To understand how these mutations impact the dynamics of these proteins we next performed CPMG experiments under identical conditions as with WT proteins above.

#### 4.2.2 Millisecond dynamics in GGC1<sup>2P→2R</sup> mutant

The mechanism by which the mutations in GGC1<sup>2P→2R</sup> inhibits the protein's function is not known. We can, however, expect that the removal of Pro-kinks in two of the TM helices affects the structure of GGC1 in such a way that two well-defined conformational states (c- and m-states) are replaced by a single distorted state (Figure 4.7). The millisecond dynamics, therefore, should be either completely abolished or display a profoundly different characteristics.

The first interesting observation is that the NMR spectra of the GGC1<sup>2P→2R</sup> mutant are very similar to wild-type protein, which is somewhat surprising. Changes in

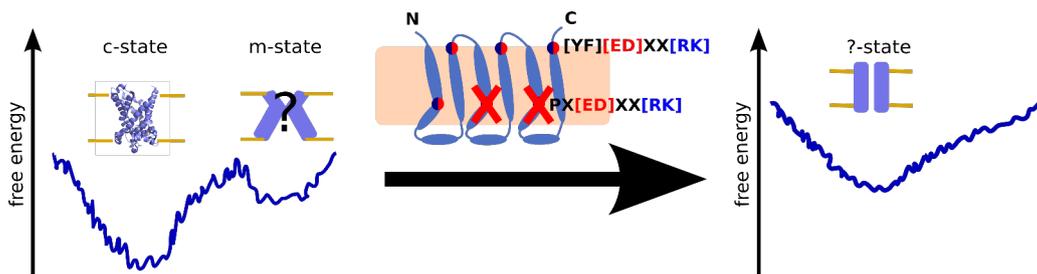


Figure 4.7: A possible mechanism which results in the inactivation of  $\text{GGC1}^{2\text{P} \rightarrow 2\text{R}}$

chemical shifts are very local around the mutation site, but not in other parts of the proteins, even those expected to be close in 3D structure (discussed further in section 4.9.1). Even more surprisingly, when we recorded CPMG RD experiments, we have observed that the millisecond dynamics were still present in this mutant protein. More importantly, the distribution of the residues displaying millisecond exchange remained nearly identical to the wild-type protein (see Figure 4.3B), and the profiles of the dispersions showed high similarity between  $\text{GGC1}$  and  $\text{GGC1}^{2\text{P} \rightarrow 2\text{R}}$  (Figure 4.58). Only two residues Lys235 and Asn242, located close to one of the mutation sites (Pro230) (on the matrix side of helix 5) displayed marked differences.

After jack-knifing based RD fitting, we could observe that the  $k_{ex}$  and  $p_B$  remained very similar between these two proteins (see Figure 4.4 and table 4).

The fact that a non-functional mutant with rather harsh mutations displays essentially the same dynamics strongly suggests that these dynamics are not crucial for function.

#### 4.2.2.1 Millisecond dynamics in $\text{AAC3}^{\text{c-salt-bridge}}$ mutant

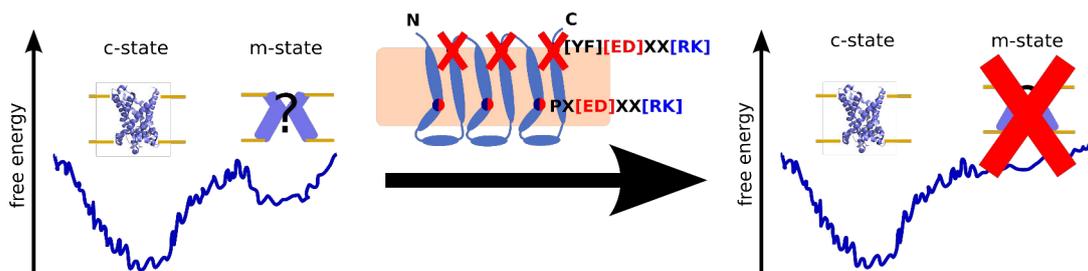


Figure 4.8: A possible mechanism which results in the inactivation of  $\text{AAC3}^{\text{c-salt-bridge}}$  mutant

We have further investigated the effect of mutations in  $\text{AAC3}$  on the millisecond dynamics. The mutations we have introduced in the cytosolic salt-bridge network would be expected to abolish the interconversion between the m- and the c-states of the protein by destabilizing the m-state (see Figure 4.8).

However, contrary to our predictions, and in accordance with the observations made with  $\text{GGC1}$ , the exchange was still present in  $\text{AAC3}^{\text{c-salt-bridge}}$  mutant. Moreover, we could observe very similar distribution of the exchange-displaying residues on

AAC3<sup>c-salt-bridge</sup> and wild-type AAC3 reported by Bruschiweiler et al. (compare Figure 4.5A and 4.5B)

The CPMG RD data of AAC3<sup>c-salt-bridge</sup> mutant was of limited quality (Figure 4.S10A) and we could not obtain convergent fits. However, if we have restricted one of the parameters ( $p_B$ ) to a value we have obtained by fitting wild-type AAC3 data, we could obtain very similar  $k_{ex}$  (see table 5) and  $|\Delta\omega|$  (Table 4.S1). More thorough comparison between the exchange profiles was not possible, as we only obtained a partial set of exchange data from Bruschiweiler et al and could not compare the dispersion profiles directly. Dispersion data of only two residues were published in the paper.

This observation strongly suggests that the millisecond dynamics remain similar between AAC3 and AAC3<sup>c-salt-bridge</sup> mutant, which is not active in the transport assays.

Therefore, mutations which inhibit the function of the mitochondrial carriers do not have an effect on the millisecond dynamics observed by NMR for the reconstituted proteins. This strongly suggests that the millisecond dynamics in MCs in DPC detergent does not correspond to a process where a carrier is interconverting between two functional states.

#### **4.2.3 Cardiolipin slows down the exchange in GGC1, but the effect is not biologically relevant**

We could not observe a link between the functional turnover of mitochondrial carriers and the millisecond exchange we could detect by NMR. However, all the aforementioned studies were performed in the absence of an important lipid, cardiolipin. Crystal structures of AAC, which is the only carrier with atomic resolution information available, reveal three cardiolipin molecules bound to a single molecule of AAC. Several studies have shown that cardiolipin is indispensable for the transport activity of AAC (210) and AAC and UCP1 was revealed to be highly unstable in the absence of cardiolipin in 10MNG (decyl maltose neopentyl glycol) detergent (235, 236). Although such studies were not performed on GGC1, the carrier contains strongly conserved [FY]XG repeats, which are implicated in cardiolipin head-group binding in the mitochondrial carriers. We would therefore expect that cardiolipin is also important for the function and stability of GGC1 and wanted to see if we could observe an effect of this lipid on GGC1 in DPC micelles.

Indeed, we could observe a large change in the CPMG RD profiles when 6 mM cardiolipin was present in GGC1 sample (Figure 4.S11). The regions displaying exchange remained the same as in the absence of cardiolipin. However, the dispersion profiles indicated much slower dynamics.

Jack-knifing simulations indeed indicated more than a two-fold decrease in the  $k_{ex}$ , as well as an increase in  $p_B$  (Table 4). The increase of populations, however, was correlated with a strong decrease in  $|\Delta\omega|$  values - reflecting therefore that these two parameters are entangled, such that they are less reliably interpreted.

To test if in the presence of cardiolipin the exchange of GGC1 is influenced by GTP, we have added 20 mM of this nucleotide to the GGC1 samples. However, the CPMG RD profiles remained virtually the same as in the absence of the GTP (Figure 4.S13). Jack-knife analysis also indicated statistically indistinguishable exchange rates (Table 4). The populations displayed a two-fold difference. However, given the similar RD profiles and the correlation between  $p_B$  and  $|\Delta\omega|$ , this difference is not significant.

This suggests that cardiolipin is not restoring the expected functional behaviour of detergent solubilized GGC1.

To confirm that the effect of cardiolipin is not specific, we have used GGC1<sup>2P→2R</sup>. The two Pro residues, which we mutated to Arg, are situated in the vicinity to the [FY]XG motif as can be judged from the structural model of GGC1. We could therefore expect the disruption of a local structural region by the substitutions and effect on cardiolipin binding. If the cardiolipin had a specific effect on slowing down the exchange rate of GGC1, GGC1<sup>2P→2R</sup> mutant dynamics should not be affected by cardiolipin.

This, however, is not the case. We have observed that cardiolipin affected the dynamics of GGC1<sup>2P→2R</sup> in the same way as the wild-type GGC1 (Figure 4.S12; Table 4). It is therefore most likely that cardiolipin affect GGC1 not via some specific interaction, but by the general change in the microscopic or macroscopic properties of the GGC1-DPC sample.

### 4.3 Summary of millisecond dynamics in mitochondrial carriers

We have observed the following features of mitochondrial carriers reconstituted in DPC micelles regarding the millisecond dynamics:

- Residues for which we observe millisecond dynamics are clustered on one side in all three MCs.
- All exchanging residues can be described by a similar time scale and population. This suggests the presence of a common exchange process, but it could also mean that local, uncorrelated processes occur with similar properties (in terms of kinetics and thermodynamics).
- The exchange rate in all three carriers is around  $1000\text{ s}^{-1}$ , with around 1-2% of minor state.
- Substrates (GTP) and substitutions which abolish the transport function of GGC1 and AAC3 do not affect millisecond exchange.
- Cardiolipin decreases the exchange rate 2 to 3 fold in both GGC1 and GGC1<sup>2P→2R</sup>. The effect thereby seems to be non-specific.

These observations suggest that DPC reconstituted carriers might not retain their native, functional conformation in DPC. We have thus continued our studies to try to confirm if the carriers were structurally compromised in this detergent.

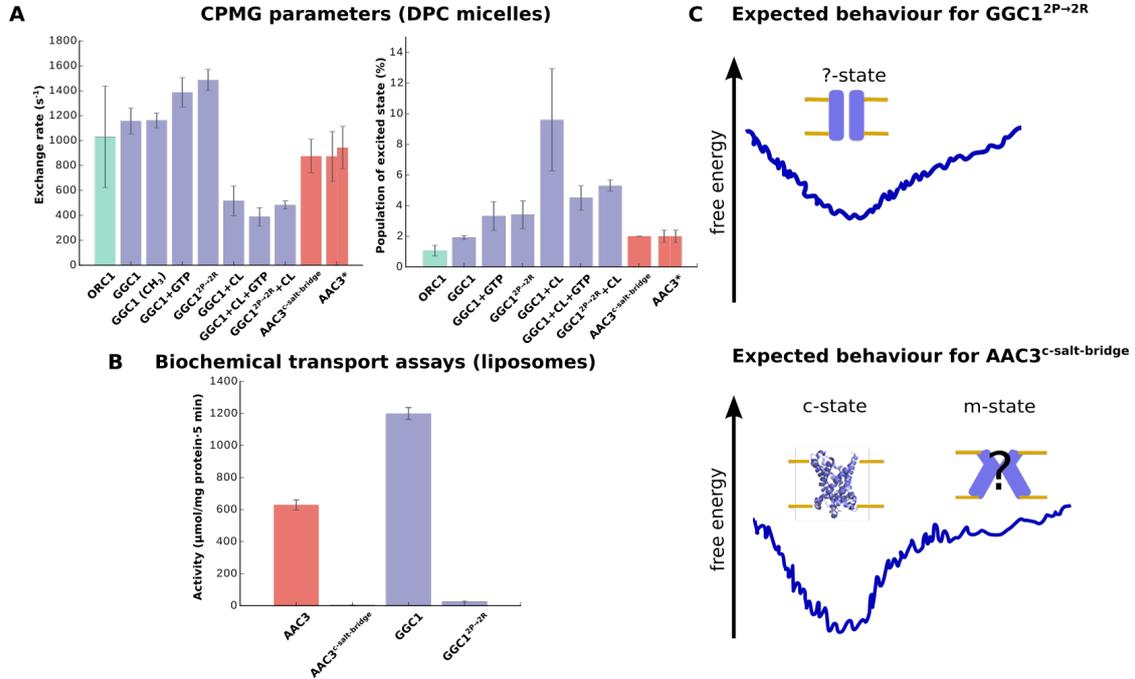


Figure 4.9: A, exchange parameters obtained after fitting of CPMG relaxation dispersion data for different carriers (exchange rates on the left and populations on the right). The parameters which were not extracted from jackknife fitting are: GGC1(CH<sub>3</sub>), GGC1<sup>2P→2R</sup> with cardiolipin and AAC3<sup>c-salt-bridge</sup>. AAC3\* data was taken from (200). B, transport activity of wild-type carriers and their mutants determined by liposome transport assays. C, an illustration displaying expected effect on the thermodynamic properties of ground and minor state populations of GGC1<sup>2P→2R</sup> (top) and AAC3<sup>c-salt-bridge</sup> mutant (bottom).

#### 4.4 DPC refolded carriers are not specific for their substrates

All the mitochondrial carriers we studied were refolded in DPC micelles. Detergent environment is very different from cellular membranes, where these proteins naturally reside. It is therefore crucial to confirm that these membrane proteins retain their biologically active structure in detergent.

The most direct assessment of biological relevance of the carriers would be the measurements of transport activity. This, however, is not possible in detergent-solubilized transporters.

We have therefore investigated the substrate binding ability of the carriers. To our advantage, NMR spectroscopy is a very powerful tool to probe the interactions between molecules.

It can provide not only residue-resolved information about the interaction, but indicate the time-scale of the interaction as well. For example, if the binding is fast on NMR time-scale, averaged chemical shifts of the residues in the bound and the free

protein would be seen (which in practical terms is represented as a change in the position of the peaks - chemical shift perturbations). On the other hand, if the binding is slow, two peaks are visible for each residue, one corresponding to the bound protein, and the other one for the free protein. Lastly, if the interaction occurs on the intermediate time scale, only a single peak for each residue is visible, which may experience both chemical shift perturbation and a decrease in intensity.

#### 4.5 ORC1 in DPC micelles is not able to bind its substrate L-ornithine

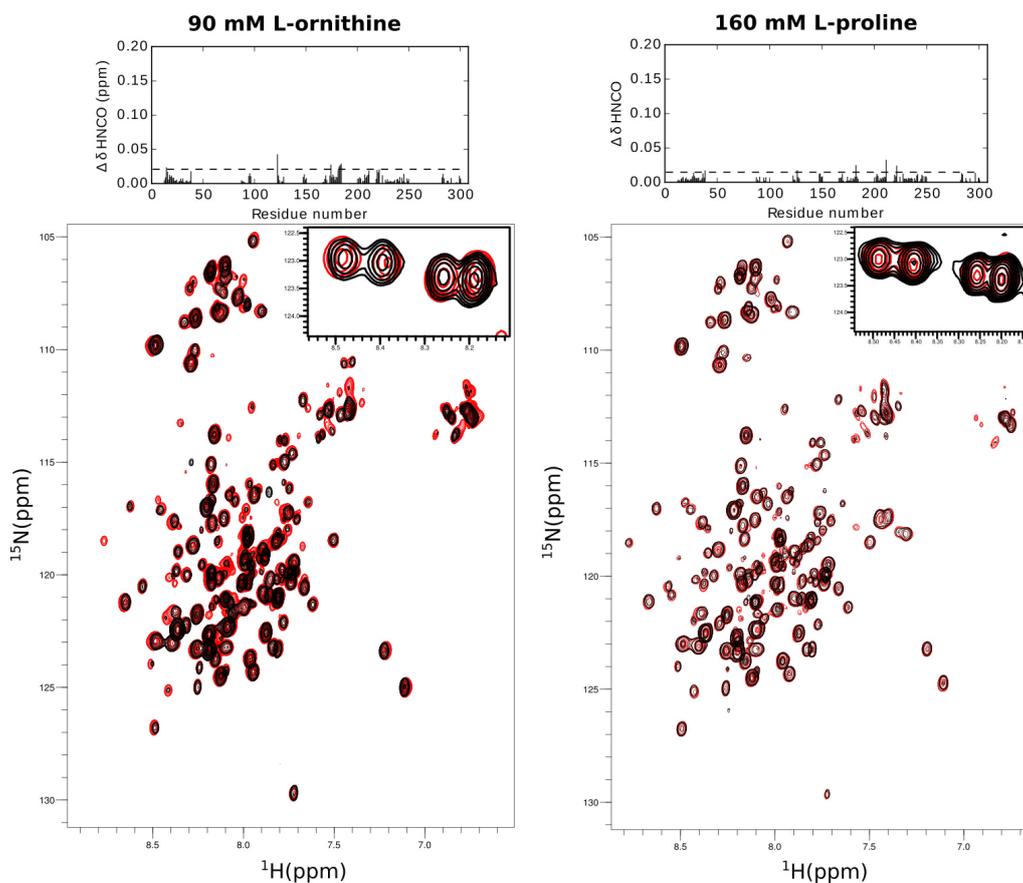


Figure 4.10: Data of yeast ORC1 in the presence (black) and absence (red) of 90 mM L-ornithine (left) and 160 mM L-proline (right). Top: residue-wise chemical shift perturbations extracted from the BT-HNCO spectra below, normalized to the chemical shift perturbation of  $^1\text{H}$ . Horizontal dashed line represents 3 standard deviations extracted by analysing all chemical shift perturbations for assigned residues. The scale of the ordinate in the bar-plot is kept the same as in the previous examples for GGC1 binding, to allow easier visual comparison of the binding data for the two proteins. Bottom: overlay of BT-HNCO spectra in the presence and absence of aforementioned solutes. Insert indicates an example of a visible chemical shift perturbation for L-Orn, which is absent in the presence of L-Pro.

Using the ligand titration to probe the concentration dependent changes in the appearance of the NMR spectra, we have investigated the interaction between yeast

ORC1 and L-ornithine. Previous reports (77) show, that L-ornithine is the main substrate for yeast ORC1, and this amino acid also displays the strongest inhibition of the transporter. The binding of L-ornithine would therefore be expected, inducing substantial changes in the NMR spectra.

Observation of the ORC1 cross-peak positions in the 3D HNC0 spectra at increasing L-ornithine concentrations, however, revealed diminishingly small chemical shift perturbations and only a global change in the peak intensities (due to increasing salt concentration and dilution of the sample) could be observed. Even at extreme concentrations of L-ornithine (120 mM,  $\approx$ 150-fold excess over the protein) the chemical shift perturbations were negligible (Figure 4.10, left). We have tested such large concentrations of the substrate, as we reasoned that the presence of N-terminal His<sub>8</sub> purification tag, having a positively charged nature, could manifest in a strong binding to the negatively charged cavity due to its very high effective concentration.

To understand the nature of these small CSPs we have also tried titrating ORC1 with L-pro, L-his and NaCl. Neither L-pro, nor L-his is a substrate of yeast ORC1, and these molecules do not inhibit the transport (77), as reported previously.

We could observe that in the presence of 160 mM L-pro, a non-charged molecule, the shifts were even weaker (Figure 4.10, right). Very small differences ( $<0.02$  ppm) between the CSPs induced by L-orn and L-pro could be observed at the N-terminal part of helix 1 (Figure 4.10, inserts) and the matrix loop linking helices 2 and 3.

Similar shifts as for L-ornithine, however, were also induced by L-his and NaCl (Figure 4.S15, also inserts) - indicating that these CSPs do not have biological relevance and are rather induced by the charged nature of titrated molecules.

We therefore conclude, that DPC solubilized ORC1 is not able to bind the substrates, and likely is in a biologically non-relevant state.

## 4.6 GGC1 in DPC micelles does not show binding specificity

### 4.6.1 GGC1 is able to bind GTP

We have next investigated the ability of GGC1 to bind the substrates. The main substrates for the carrier are 5'-guanosine di- and triphosphates (GDP and GTP, respectively) as reported in the transport studies (78) and they also show the strongest competitive inhibition of the carrier. The inhibition efficiency was measured by performing transport assays of radioactively labeled GTP in the presence of unlabeled molecules. Only if a molecule is able to bind inside the cavity of the carrier a competitive inhibition would be expected. Strong inhibition therefore indicates strong binding. We were therefore expecting to observe binding of GTP to GGC1, which would induce changes in the NMR spectra.

We have first investigated if we could observe the interaction between GGC1 and GTP by looking at the backbone atoms of GGC1. In agreement with our expectations, addition of 20 mM of GTP to  $\approx$ 0.5-0.7 mM GGC1 sample induced pronounced chemical shift perturbations (CSPs), confirming the interaction between GGC1 and its substrate (Figure 4.11A). We have then investigated these CSPs more closely, by

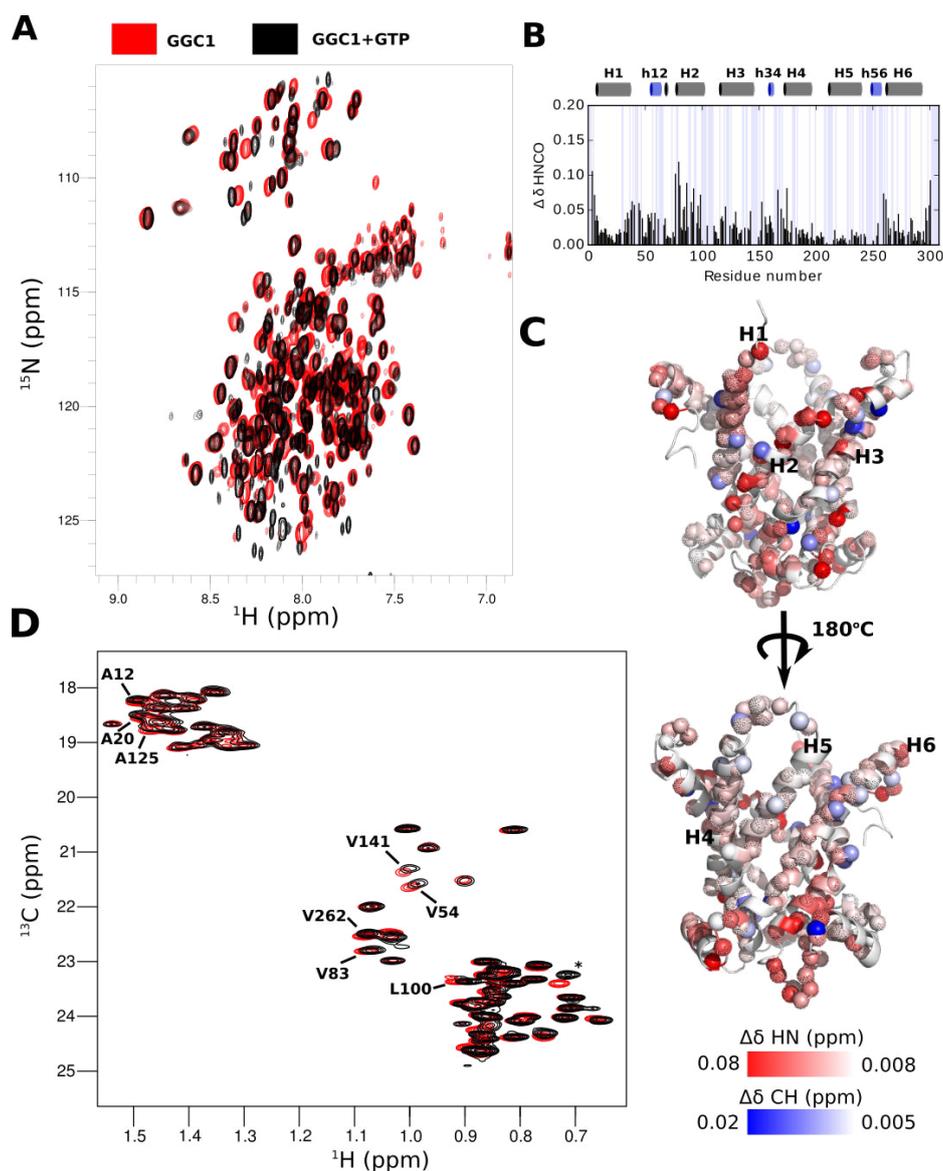


Figure 4.11: GTP binding data for GGC1. A, overlay of GGC1 3D BT-HNCO spectra in the absence (red) and presence (black) of 20 mM GTP. B, Data obtained from the comparison of peak positions from the spectra in A. ( $^{15}\text{N}$ ,  $^{13}\text{C}$  and  $^1\text{H}$ ) CSPs are normalized to the chemical shift perturbation of  $^1\text{H}$ . Blue shaded background represents residues for which we could not determine chemical shift perturbation (due to the lack of assignment or disappearance of the peak). Gray and blue cylinders above correspond to the TM and matrix helices respectively, determined from the homology model of GGC1. C, the data in B plotted on the homology model of GGC1 (white-red spheres) and the chemical shift data extracted from the methyl-SOFAST spectra (white-blue spheres). A good qualitative agreement is observed between the two independently obtained measurements. Horizontal color-bars represent the magnitudes of chemical shift perturbations for amides (red) and methyls (blue). D, the methyl-SOFAST spectra of GGC1 in the absence (red) and presence (black) of 20 mM GTP, used to extract the data in C). Some more strongly perturbed residues are indicated.

mapping the magnitude of perturbation<sup>26</sup> for each peak onto the homology model

<sup>26</sup>As the chemical shifts of all the observed nuclei may be effected ( $^{15}\text{N}$ ,  $^{13}\text{C}$  and  $^1\text{H}$ ), we have calculated a combined CSPs, weighted by the gyromagnetic ratio of the three nuclei.

of GGC1 (see Figure 4.11C). Although the proposed binding site of GGC1 (and other mitochondrial carriers) is approximately mid-way across the bilayer (203), we observed that the effect of GTP extended to other regions of GGC1 as well. The largest CSPs (0.05-0.1 ppm) were observed for the residues all along the H2, and helices H1 and H6 were affected to similar extent on their cytoplasmic sides (Figure 4.11B,C). Weaker, but substantial perturbations were also observed on all the matrix helices. This binding pattern has already been reported for GGC1 in DPC (228).

To get more information about the GTP binding, we have investigated how the substrate affects the methyl groups of A, L and V in GGC1. The CSPs we have observed on the methyl groups were in agreement with those observed on the backbone, as the same regions of the protein were affected (Figure 4.11C,D).

We have thereby shown the interaction between GGC1 and its substrate, GTP. The pattern of the interaction we have observed is in a very good agreement with one reported in a previous study of GGC1 (228), where the saturation of the GTP binding to GGC1 was also shown.

#### 4.6.2 DPC reconstituted GGC1 bind ATP in an identical manner as GTP

Previously discussed interaction between GGC1 and GTP could already be considered as a confirmation that GGC1 reconstituted in DPC retains its native-like state (as was assumed in the previous papers of several mitochondrial carriers (200, 225, 228)). However, the discrepancy between the expected conformational exchange model and the CPMG parameters we have observed, prompted us to perform a more stringent study.

We therefore wanted to investigate the substrate specificity of GGC1. Like all of the mitochondrial carriers, GGC1 possesses a remarkable ability to differentiate between very similar molecules. GTP, for example, differs only in C6 and C2 positions of its nucleobase from ATP. The reported transport assays, however, indicate that even at ten-fold larger concentrations, ATP is not able to inhibit GTP transport by GGC1 in liposomes (78).

We have confirmed these observations by asking Dr. Loredana Capobianco to perform GGC1 transport assays in liposomes, where the GTP transport would be measured in the presence of an excess of ATP. In agreement with the previous report, ATP showed only slight inhibition on our GGC1 construct Table 8. We therefore were expecting no interaction, or only very weak interaction between GGC1 and ATP.

The NMR spectra we observed after addition of 20 mM of ATP to the  $\approx 0.5$ -0.7 mM sample of GGC1 struck us by surprise. Even a visual analysis of the spectra indicated nearly identical CSPs induced by ATP as we have observed for GTP (Figure 4.12A). Performing residue-wise analysis of the CSP magnitudes (as before) we have observed nearly identical CSPs induced by ATP or GTP, for all the residues in GGC1 we could assign (Figure 4.12B). The excellent agreement between the CSPs

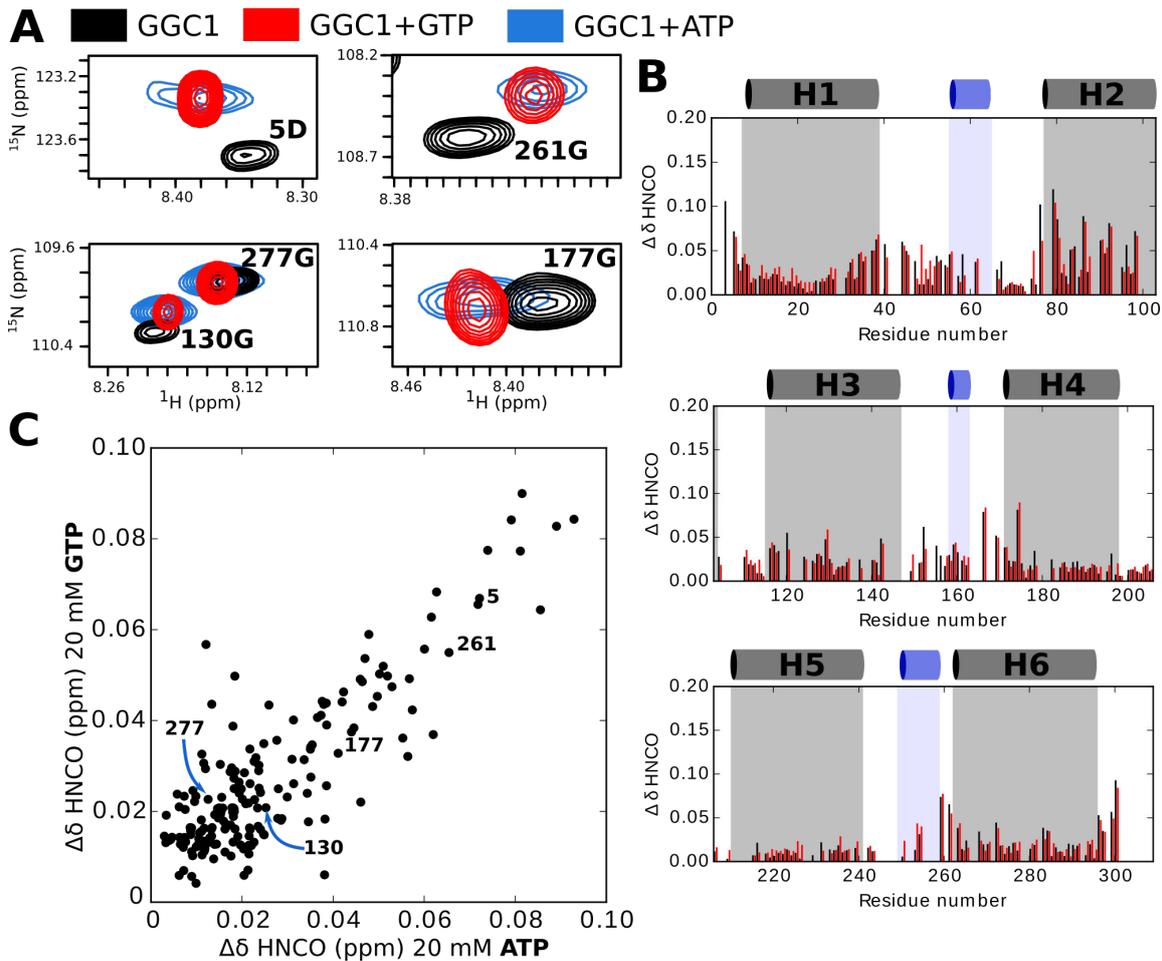


Figure 4.12: A, an overlay of GGC1 spectra in the absence (black) or presence of 20 mM of GTP (red) or ATP (blue), indicating the magnitudes and directions of observed chemical shifts. B, a plot displaying combined chemical shift perturbations of  $^{15}\text{N}$ ,  $^{13}\text{C}$  and  $^1\text{H}$ , normalized to the chemical shift of  $^1\text{H}$ . Black bars correspond to perturbations induced by 20 mM of GTP (black) or ATP (red). Cylinders above correspond to TM and matrix helices (gray and blue, respectively) extracted from the homology model of GGC1. C, the same data as in B but displayed as a correlation plot. Numbers indicate the position of residues in A in respect to the correlation plot.

can be seen in the correlation plot, where the perturbations of GTP and ATP are plotted on y- and x-axes, respectively (Figure 4.12C).

These findings strongly suggest that DPC reconstituted GGC1 no longer has the specificity for purine nucleotides, in contrast to the liposome-embedded carrier.

#### 4.6.3 GGC1 shows interaction with carboxyatractyloside

ATP and GTP are very similar in structure, and likely only small structural perturbations in the mitochondrial carrier could result in the loss of the ability of the protein to discriminate between these two molecules. Therefore we wanted to investigate the binding of a very different molecule, CATR, to GGC1. CATR is able to inhibit

Table 8: Inhibition of GGC1 by ATP. Transport of 50  $\mu\text{M}$  of external  $^3\text{H}$ -GTP in the presence or absence of 150  $\mu\text{M}$  of external ATP, in proteoliposomes preloded with 5 mM GTP.

	Transport rate ( $\mu\text{mol} \cdot \text{mg protein}^{-1} \cdot 2 \text{ min}$ )
-ATP	$0.67 \pm 0.044$
+ATP	$0.6 \pm 0.033$

AAC at nanomolar concentrations in liposomes. However, the effect of this inhibitor on GGC1 is very weak and likely non-specific: it is reported that 100-fold excess of CATR over GTP only inhibits the transport by only  $\approx 20\%$  (78) in proteoliposome transport assays. Our collaborators have confirmed that CATR has essentially no effect of inhibition on GGC1 (data not shown). CATR binding to GGC1 is therefore expected to be very weak and likely non-specific.

To test if CATR is interacting with GGC1, we have measured a 2D HSQC spectra at different concentrations of CATR in a GGC1 sample (1 mM, 3 mM, 5 mM). Surprisingly, even at 1 mM concentrations CATR induced substantial CSPs on the GGC1 spectra (Figure 4.13A). It is very interesting, as CSPs of similar magnitude were reported for AAC3 reconstituted in DPC micelles (200). Although we did not investigate in detail, the majority of CSPs showed complex change of magnitude upon increasing concentrations of CATR. We could observe saturation resembling behaviour at lower concentrations of CATR (1-3 mM), however, at 5 mM of CATR the shifts magnitudes increased again.

Mapping the CATR induced CSPs on the GGC1 structural model revealed that similar regions were affected by this molecule as compared to ATP or GTP (Figure 4.13B). Namely, the strongest effects were observed on the termini of GGC1 (H1 and H6 parts, which would face intermitochondrial space in GGC1), on the helix 2 and on the matrix loop connecting helices 3 and 4 (Figure 4.13C). This could indicate that the binding sites of GGC1 to GTP, ATP and CATR share some similarity.

We have therefore shown that GGC1 reconstituted in DPC is able to bind GTP, but also ATP and CATR in a non-specific manner, which is in a complete disagreement with the biochemical transport assays performed in liposomes.

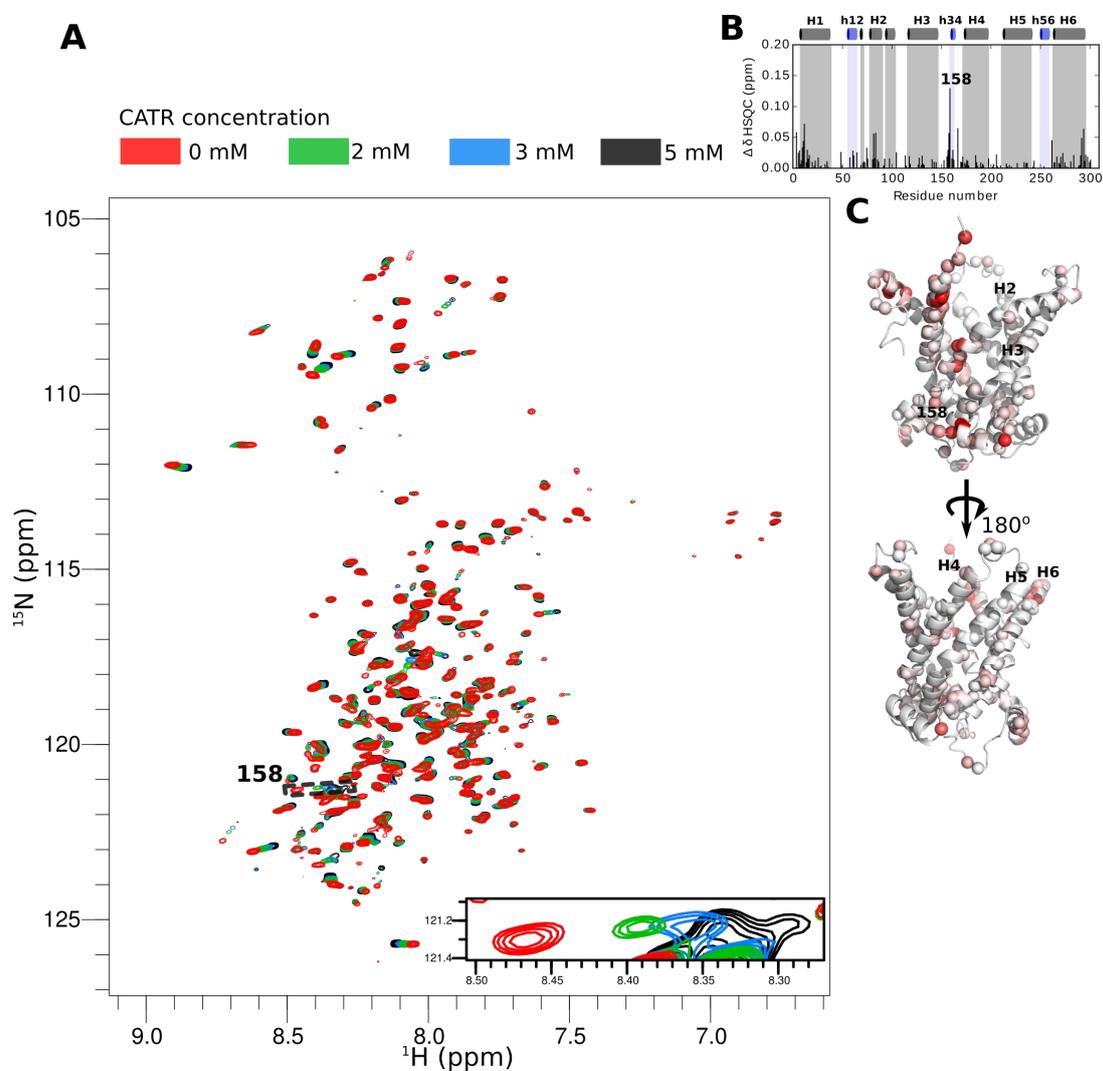


Figure 4.13: Data illustrating observed GGC1 chemical shift perturbations in the presence of carboxyatractyloside (CATR). A, spectra indicating wide-spread chemical shift perturbations upon addition of increasing concentrations of CATR (red, 0 mM; green, 1 mM; blue, 3 mM; black, 5 mM). Largest chemical shift perturbation is indicated by a dashed box and an insert. B, the data from A represented as a bar-plot, where chemical shift perturbations were normalized to  $^1\text{H}$  chemical shift. Gray and blue cylinders and shaded backgrounds represent TM and matrix helices, respectively. C, the data in A plotted on a 3D homology model of GGC1. The locations with strong chemical shift perturbations broadly agree with the regions with strongest CSP for GGC1 in the presence of GTP.

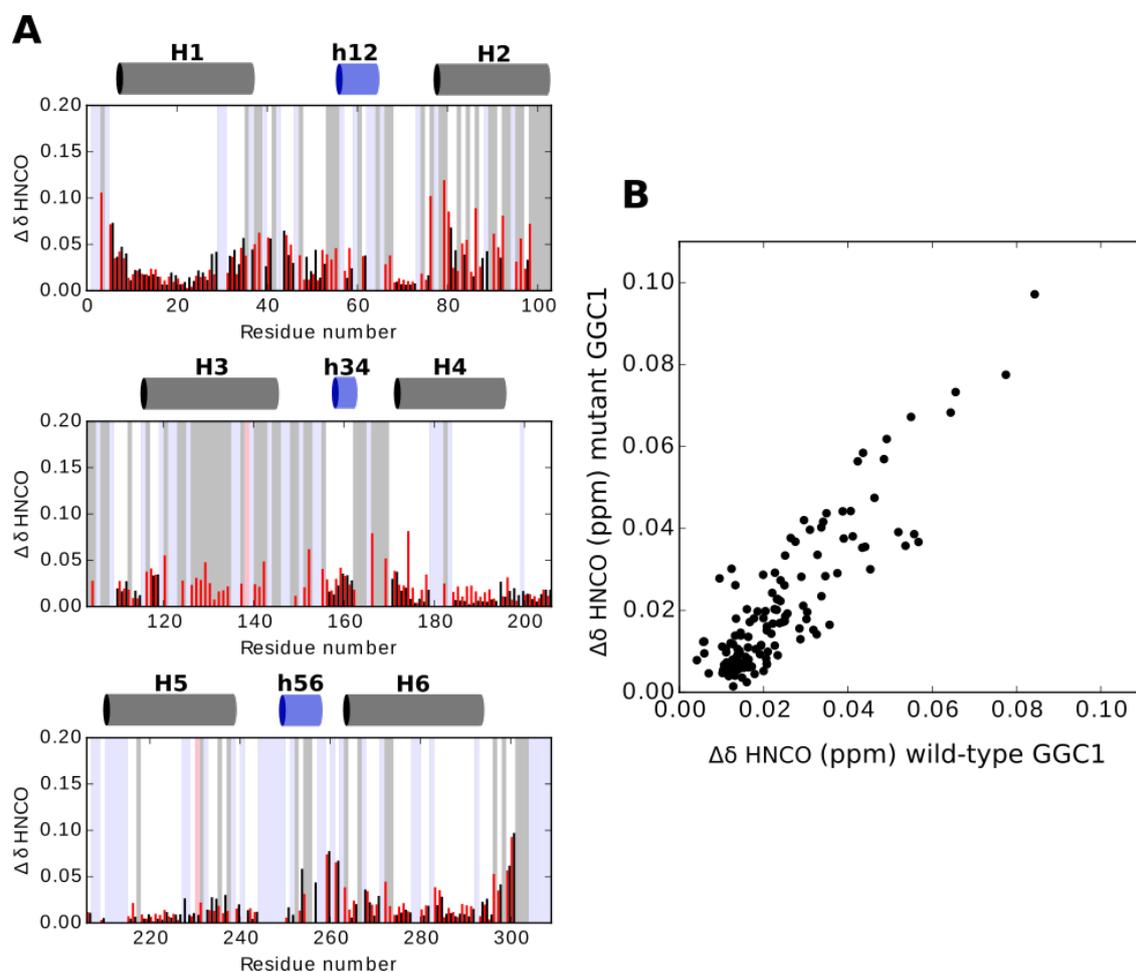


Figure 4.14: Comparison of chemical shift perturbations upon addition of 20 mM GTP for wild-type GGC1 (red bars) and GGC1<sup>2P→2R</sup> (black bars). A, residue-wise chemical shift perturbations. Ordinate represents a chemical shift perturbation in the three dimensions (<sup>15</sup>N, <sup>13</sup>C and <sup>1</sup>H) normalized to the chemical shift perturbation of <sup>1</sup>H by multiplication of chemical shift perturbation with the ratio of gyromagnetic radii. Abscissa indicates residue number. Light blue shaded background indicated regions where chemical shift perturbations are not known for wild-type GGC1. Grey shaded background indicated missing information about the chemical shift perturbations for GGC1<sup>2P→2R</sup>. Red shaded regions indicate the position of mutations (Pro138Arg and Pro230Arg). Gray and blue cylinders above indicate TM and matrix helices as determined from the homology model of GGC1. B, the same data as in A plotted in a different way to represent strong correlation between the wild-type GGC1 and GGC1<sup>2P→2R</sup> for GTP binding.

#### 4.6.4 GGC1<sup>2P→2R</sup> mutant binding to GTP

We have already shown that wild-type GGC1, when refolded in DPC micelles is no longer able to differentiate between different molecules, and all the tested interaction partners appear to have similar binding sites on the GGC1 structure. We further wanted to know if GGC1<sup>2P→2R</sup> mutant, which contains large substitutions at the conserved Pro kinks, would display a different mode of interaction with GTP. The difference could be expected, as the mutations in GGC1<sup>2P→2R</sup> should perturb the

orientation of TM helices in GGC1.

We have performed the same interaction studies with GGC1<sup>2P→2R</sup> as with the wild-type GGC1, where we observed CSPs after the addition of 20 mM GTP. No longer surprisingly, we could observe nearly identical CSP profile for this mutant as for wild-type GGC1, for all the peaks that were assigned (Figure 4.14A). The similarity of the binding is even more obvious on a correlation plot of CSP for wild-type GGC1 and the GGC1<sup>2P→2R</sup> mutant protein (Figure 4.14B).

It is not easy to interpret such a similar binding profile for a wild-type protein, and a protein which contains severe perturbations at two positions, where Pro residues have been shown to induce sharp kinks in the helices, as observed in the crystal structures. A more detailed interpretation of the data will be presented below in the context of the tertiary structure of the carriers.

#### 4.6.5 Cardiolipin binds to GGC1 non-specifically

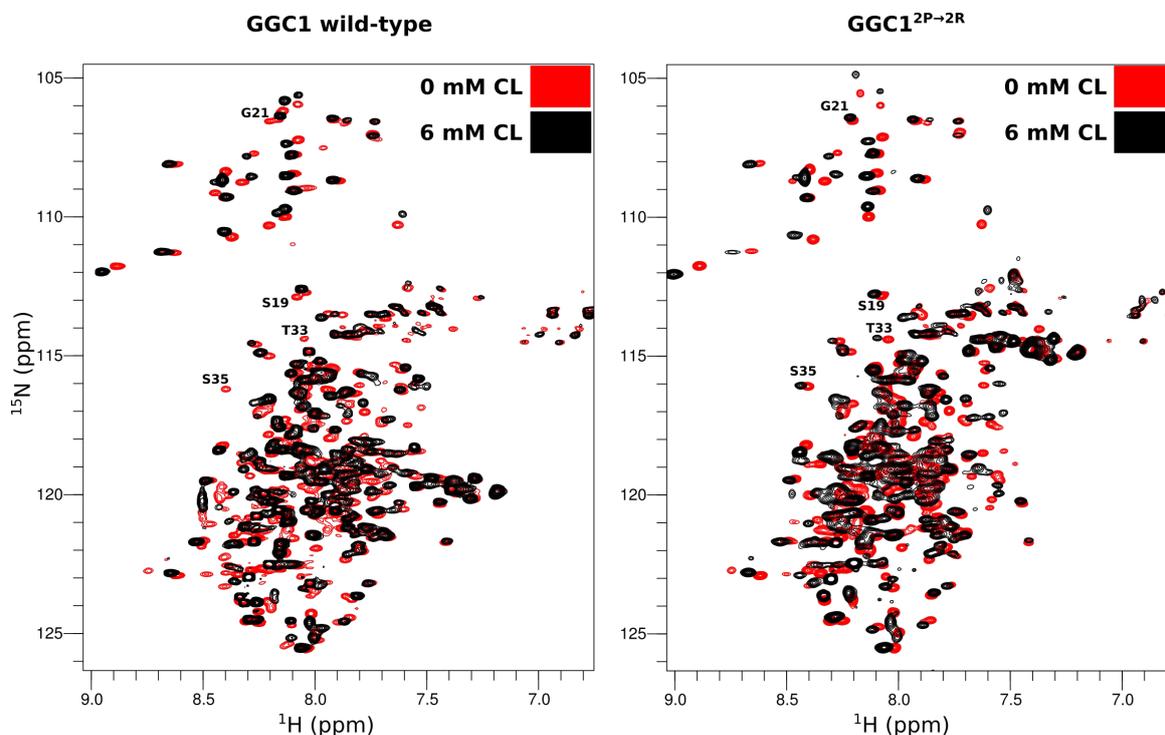


Figure 4.15: 2D HSQC spectra of wild-type GGC1 (left) and GGC1<sup>2P→2R</sup> (right) in the absence (red) and presence (black) of 6 mM cardiolipin. Similar magnitudes and directions of chemical shifts are obvious. Some residues disappearing in wtGGC1, but still present in GGC1<sup>2P→2R</sup> are indicated. Differences in the reference spectra of GGC1<sup>2P→2R</sup> and wtGGC1 likely correspond to differences resulting from the mutations.

It appears, so far, that DPC reconstituted GGC1 does not discriminate between three different molecules (ATP, GTP and CATR), in contrast to the observations on proteoliposome transport assays. In all these binding studies, however, we used GGC1 purified and refolded in the absence of cardiolipin. As mentioned before, car-

diolipin is reportedly important for the function of mitochondrial carriers.

Addition of 6 mM cardiolipin to a  $\approx 0.6$  mM GGC1 sample resulted in wide-spread CSPs, indicating interaction between GGC1 and cardiolipin (Figure 4.15, left). Although we would expect the strongest perturbations to occur near the regions where the negatively charged cardiolipin head-group is bound, for example [FY]XG motif, we could not observe such a trend (data not shown). Besides the chemical shift perturbations, we have observed a substantial decrease in intensities on the helix 1 (Figure 4.S16). This helix is close to one of the cardiolipin binding sites. The binding, therefore, may be biologically significant.

We have tested the potential biological significance of cardiolipin binding by performing the same experiments of GGC1<sup>2P→2R</sup>. The CSPs induced by 6 mM of cardiolipin on GGC1<sup>2P→2R</sup> mutant were, however, nearly identical to the wild-type protein (Figure 4.15, right). This is suggestive that cardiolipin induced effects are non-specific and not biologically relevant.

We have strengthened our previous conclusion by investigating the GGC1 binding to another negatively charged lipid, POPG. POPG is structurally similar to cardiolipin, as both lipids contain a glycerol head-group. However, two phosphatidic acid groups are attached to the glycerol moiety in cardiolipin, whereas POPG contains only one. It should be noted that cardiolipin binds very strongly to MCs, and even extensive washing of MCs extracted from the native membrane cannot eliminate CL, although it would remove all other lipids (such as POPG). Therefore, we expected different behaviour of CL and POPG.

After analysing the spectra in the presence and absence of 6 mM POPG we have indeed observed smaller chemical shift perturbations as compared to 6 mM of cardiolipin (Figure 4.S18). However, these perturbations were almost identical to the perturbations induced by 3 mM of cardiolipin on GGC1. This effect is more likely related to the difference in charge (CL<sup>2-</sup> as compared to POPG<sup>1-</sup>) than to specificity.

It seems, therefore, that the binding of GGC1 to cardiolipin is also not representative of the behaviour which could be expected from the carrier containing its native structure. We have to note here, however, that cardiolipin was only able to induce strong decrease of the peak intensities for the residues in the helix 1 in wild-type GGC1. We did not observe the decrease in intensities neither in GGC1<sup>2P→2R</sup> in the presence of cardiolipin, nor in GGC1 in the presence of POPG (Figure 4.15, S18).

#### **4.6.6 GGC1 displays similar features refolded from the inclusion bodies or expressed by a cell-free system**

Cardiolipin did not appear to restore the function of GGC1. Thus we were wondering if the refolding procedure itself might be a problem. Therefore, we have expressed GGC1 using an *in-vitro* translation system in the presence of Brij-35 detergent. Cell-free expression system allows the protein to refold inside the detergent, as soon as the nascent protein leaves the ribosome, without undergoing aggregation into in-

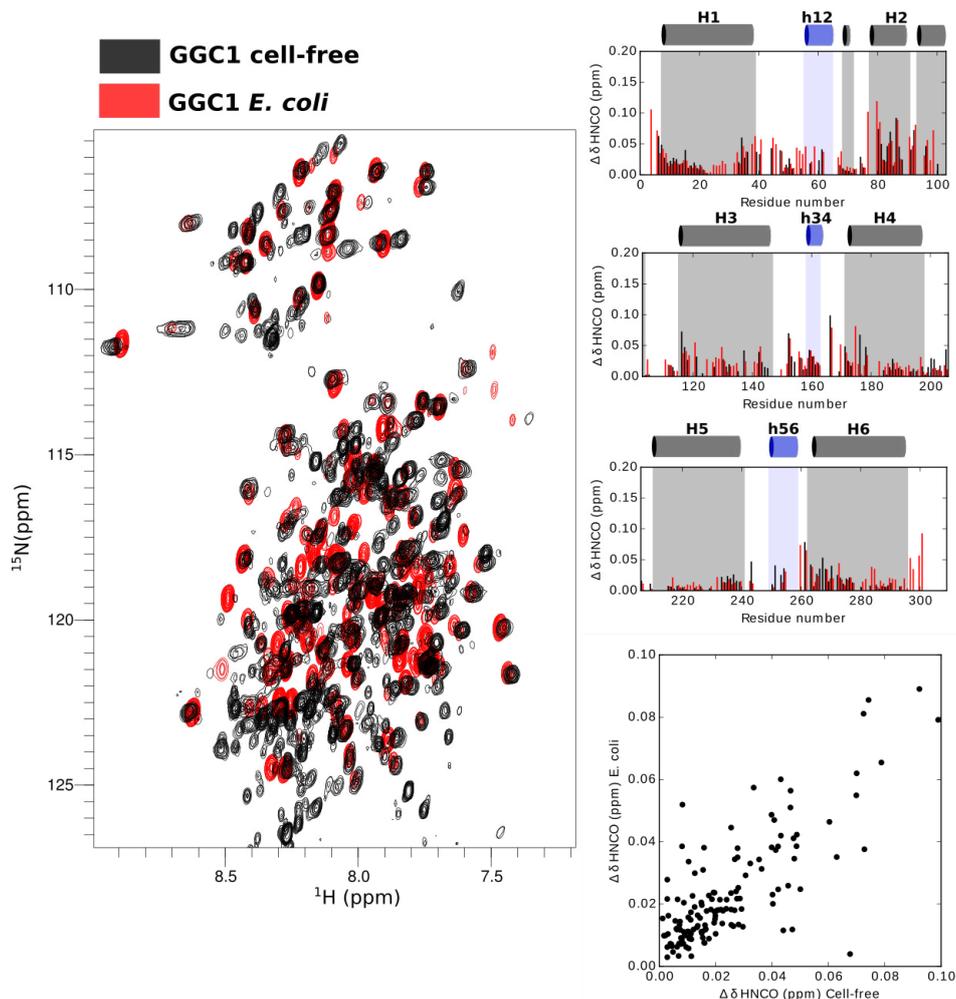


Figure 4.16: Comparison of GGC1 refolded from inclusion bodies (red) and expressed using cell-free expression system. Note that in the cell-free expression, abortive fragments were present, and thus additional cross-peaks from these fragments are expected. Left: overlay of combined planes of HNCOSY spectra. Good overlay of isolated peaks indicates that the topology of two proteins remains similar. Right: comparison of GTP binding properties between refolded GGC1 (red) and cell-free expressed GGC1 (black). On the bottom correlation plot of the data on top can be seen.

clusion bodies. We have expected to see a differently folded GGC1, which might have even regained its biological function.

After expressing GGC1 in Brij-35 detergent, we have exchanged the detergent to DPC and investigated this sample by NMR. Interestingly, the HNCOSY spectrum appeared to be very similar to the one of GGC1 refolded from inclusion bodies (see Figure 4.16).

Some differences might have arose due to the fact that this protein has an N-terminal His<sub>6</sub> purification tag: on one hand the positions of C- and N-terminal peaks are different, and on the other hand we observe here some peaks from the translation

truncation products<sup>27</sup>. Of course there could be additional factors affecting the peak positions - for example some remaining Brij-35 detergent in the sample.

To check if these two samples were indeed similar, we have tested the binding of cell-free GGC1 sample to 20 mM GTP. We have observed a very similar binding profile as for the wild-type protein (see Figure 4.16).

It thus appears that the procedure of sample preparation does not effect the structure of GGC1 substantially. We have also obtained very similar results for ORC1 (data not shown). We conclude that the mitochondrial carriers we have investigated (ORC1, GGC1) adopt the same structural state irrespective whether they were refolded from inclusion bodies or directly inserted into Brij-35 detergent micelles followed by exchange of the detergent to DPC.

#### 4.7 AAC3<sup>c-salt-bridge</sup> mutant in DPC does not have binding specificity

So far we have shown that DPC solubilized ORC1 is not able to bind the substrates, while GGC1 shows non-specific interactions with several molecules. We further questioned AAC3 binding by performing the experiments on AAC3<sup>c-salt-bridge</sup> mutant. AAC, similarly to GGC1, is very specific towards ATP. Human AAC has been reported to only be inhibited by only <30% by a 1000-fold excess of GTP. As there was no information available for yeast AAC3, we have asked Dr. Loredana Capobianco to perform similar measurements on this protein. The results indicated no substantial binding.

We could confirm the binding of ATP to AAC3<sup>c-salt-bridge</sup> mutant by performing 2D HSQC experiments on the protein in the absence and presence of 20 mM ATP. However, we have observed much smaller CSP magnitudes for AAC3<sup>c-salt-bridge</sup> as compared to GGC1 (see Figure 4.17A and Figure 4.S14 for the full spectrum).

There was already an NMR study published on a near wild-type AAC3 in DPC by Bruschiweiler et al. However, we could only compare a CSP of a single residue (Arg141) of our AAC3<sup>c-salt-bridge</sup> mutant with AAC3\*, as only so much binding information was given in the previous study by Bruschiweiler et al. (200) regarding the ATP binding. The CSP of this residue between the two proteins in the presence of 20 mM ATP is very similar (Figure 4.17A, insert).

We have further measured CSPs of AAC3<sup>c-salt-bridge</sup> mutant peaks in the presence of 20 mM GTP. The spectrum of the carrier in the presence of GTP was nearly identical to the one recorded in the presence of ATP - in a very good agreement with the aforementioned observations on GGC1 (see Figure 4.17B). Due to the limited peak-resolution in 2D HSQC spectra (which we recorded as the sample concentration was limited), we did not attempt to obtain complete assignment of AAC3<sup>c-salt-bridge</sup> cross-peaks in the presence of substrates. However, a limited amount of cross-peaks we could assign indicated a strong correlation between the two molecules (Figure 4.17C) and showed that the strongest CSPs by both purine nucleotides were induced in matrix helices and loops (Figure 4.17B). This does not seem to be in agreement with a very well established model of AAC binding to ATP, based on many biochemical experiments of AAC in mitochondrial inner membranes or liposomes and

---

<sup>27</sup>Due to the low protein amount obtained after cell-free expression, we did not perform size exclusion chromatography on this sample

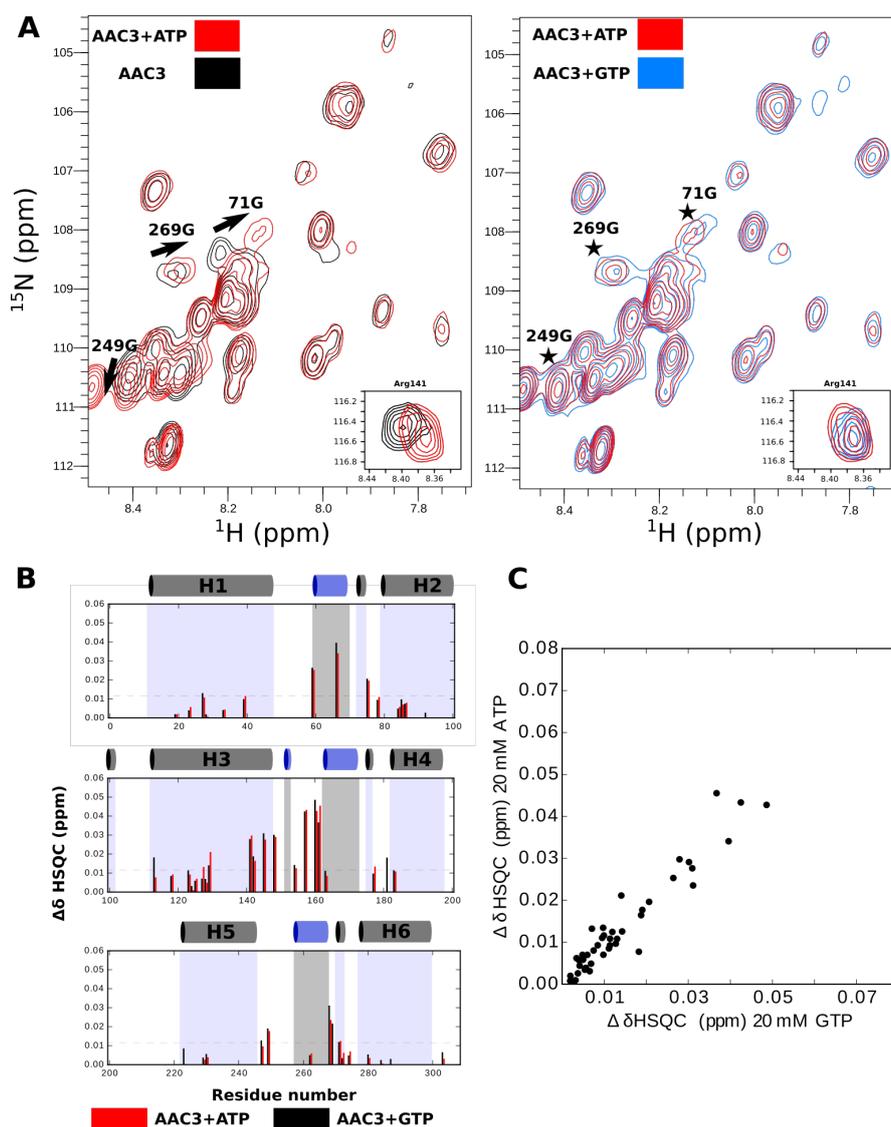


Figure 4.17: A, glycine region of 2D BT-HSQC spectra depicting AAC3<sup>c-salt-bridge</sup> mutant binding to two different nucleotides. Left - overlay of apo AAC3<sup>c-salt-bridge</sup> mutant spectra (black) with the spectrum of AAC3<sup>c-salt-bridge</sup> mutant in the presence of 20 mM of ATP. The more pronounced chemical shift perturbations are indicated by arrows. Right - overlay of AAC3<sup>c-salt-bridge</sup> mutant spectra in the presence of either ATP (red) or GTP (blue). The residues displaying most pronounced chemical shift perturbations are indicated by stars. B, the data extracted from A, plotted as a residue-wise combined chemical shift perturbations of  $^{15}\text{N}$  and  $^1\text{H}$ , normalized to the chemical shift of  $^1\text{H}$ . Red bars correspond to AAC3<sup>c-salt-bridge</sup> mutant in the presence of ATP and the black bars - GTP. Gray and blue cylinders indicate the TM and matrix helices, respectively, as obtained from the crystal structure of AAC3. From the limited assignment it appears that the largest chemical shift perturbations occur not in the TM helix region. C, the same data as in B, but displayed as a correlation of ATP and GTP induced chemical shift perturbations.

features observed from in the AAC structures (203).

#### 4.8 The lack of substrate binding/specificity to the carriers in DPC micelles suggest perturbed tertiary structure

We have observed the following features of mitochondrial carriers reconstituted in DPC regarding their binding to the substrates:

- ORC1 does not bind L-ornithine. Small CSPs appear unspecific and related to charge interactions, as NaCl induces similar effects.
- GGC1 is able to bind GTP, but also ATP in an identical manner; it also binds CATR
- GGC1 is able to bind cardiolipin, but also POPG in likely non-specific manner
- GGC1<sup>2P→2R</sup> mutant binds GTP to the same extent as wild-type GGC1
- GGC1<sup>2P→2R</sup> mutant binds cardiolipin to nearly the same extent as wild-type GGC1
- AAC3<sup>c-salt-bridge</sup> mutant is binding ATP and GTP in an identical manner; it also appears that AAC3<sup>c-salt-bridge</sup> mutant is binding these molecules similarly as wild-type AAC3.

We also know from the literature the following information about the carriers in DPC:

- GGC1 shows saturating binding to GTP, in the  $K_D$  range of 5.5-7.5 mM (*228*)
- in AAC3 at least one residue shows saturation to ATP binding (Arg141) (*200*)

We can easily interpret the inability of ORC1 to bind L-ornithine as structural perturbations caused by DPC. However, GGC1 is still able to bind GTP, albeit without expected specificity, and apparently the binding shows saturation. How this behaviour could be interpreted?

AAC3 and GGC1 have a multitude of positively charged residues inside the cavity, which are not involved in the salt-bridges. Prediction of electrostatic potential indicates, that the cavities of mitochondrial carriers transporting negatively charged substrates contain large positive electrostatic charge (see Figure 4.18). Negatively charged molecules, therefore, could be expected to bind inside this positively charged cavity, even if the proteins experienced some structural perturbations.

This interpretation seems plausible, as it largely explains our observations. At pH 6.0 (the conditions in which we performed our NMR experiments), ATP exists in an equilibrium between  $ATP^{3-}$  and  $ATP^{4-}$ , with the former species being predominant by  $\approx 80\%$  (*237*). Similar distribution can be expected for GTP. In addition, it has been shown that  $GDP^{2-}$  induces weaker CSPs in GGC1-DPC samples than GTP, while  $AMP^{1-}$  only induces very weak CSPs (*228*).

Similarly, cardiolipin and POPG likely exist in  $CL^{2-}$  and  $POPG^{1-}$  ionization states, respectively, given the  $pK_a$  value of  $\approx 2.2$  for the phosphate groups in cardiolipin (*238*). We also observe charge (or alkyl chain) dependent chemical CSPs between POPG and CL (the shifts induced in GGC1 by POPG closely correspond to the

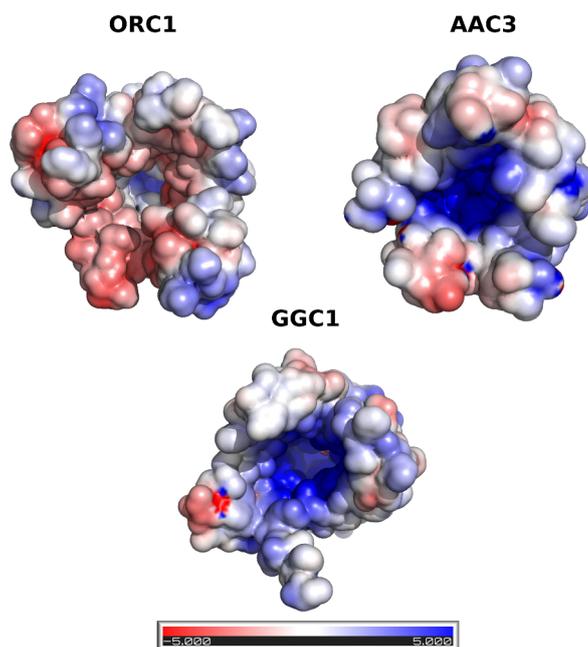


Figure 4.18: Electrostatic potential maps for the three mitochondrial carriers we studied. In the case of ORC1 and GGC1, electrostatic potential is mapped on homology models.

shifts induced by cardiolipin at twice lower concentration).

The  $pK_a$  values of two ionizable carboxyl groups in CATR have been reported as 3.8 and 6.8 (*239*), thereby this molecule also exists in an equilibrium between  $CATR^{2-}$  and  $CATR^{1-}$ .

Thereby it appears that there is some correlation between the charge and binding efficiency to mitochondrial carriers in DPC. However, molecules of different nature appear to induce very different CSPs. This is expected, as additional factors would very likely affect the binding. For example cardiolipin and POPG would be expected to partition inside the hydrophobic phase of the micelle, as opposed to ATP or GTP. These factors are too numerous to attempt any interpretation of why different molecules induce different CSPs with our current data.

The aforementioned interpretation, however, does not explain why GGC1 and  $GGC1^{2P \rightarrow 2R}$  mutant display nearly identical CSPs in the presence of GTP and cardiolipin. The two proteins would be expected to differ substantially due to the structural perturbations introduced by substitutions of residues forming Pro kinks, and thereby change the binding profile. To investigate this behaviour we wanted to obtain some structural information about mitochondrial carriers reconstituted in DPC micelles.

## 4.9 Mitochondrial carriers do not appear to have well defined structure in DPC micelles

### 4.9.1 GGC1 in DPC micelles lacks compact, well folded structure

As we have observed compromised substrate binding ability, we wanted to understand underlying structural perturbations. A good tool to get such information is the measurements of Nuclear Overhauser Effects (NOEs). These effects arise when two nuclei are in a close proximity and can be used to determine distances and arrangement of secondary structure elements in the protein. Typically NOEs are measured between the  $^1\text{H}$  atoms in the protein. In the case of large proteins, high-resolution spectra can only be obtained with deuterated proteins and partial back-protonation (in our case: protonation of amides, and some methyl groups). Therefore, we recorded NOE experiments between methyls, and between amides.

For this, we have used  $^1\text{H},^{13}\text{C}$ -ALV  $\text{CH}_3$  labeled  $^1\text{HN},^{15}\text{N},^{12}\text{C},^2\text{H}$ -GGC1 in  $^2\text{H}$ -DPC micelles. Although some methyl groups in this protein would be expected, as judged from the structural model, to be separated by less than  $5\text{\AA}$  (distance between C atoms), we could not observe any NOE cross-peaks between the methyl groups on an  $^1\text{H}$ -NOESY-HSQC experiment (data not shown). This is surprising and indicates that a well defined structure might be lacking in this protein.

We have gained additional information on the tertiary structure of GGC1 in DPC by comparing the chemical shift perturbation between wild-type and mutant proteins.

GGC1 and GGC1<sup>2P→2R</sup> differ in a way, that 2 Pro residues, inducing the kinks in the TM helices are replaced by 2 Arg residues. These regions, as can be judged from the structural model, are very well-packed. Such large substitutions, where a neutral residue is replaced with a charged one, would be expected, therefore, to induce chemical shift perturbations which would propagate along the whole structure of GGC1. The spectra of GGC1 and GGC1<sup>2P→2R</sup>, however, looked surprisingly similar (see Figure 4.19B). Mapping the CSPs between GGC1 and GGC1<sup>2P→2R</sup> on the structural model of GGC1 reveals, that although large perturbations are observed ( $\approx 0.3$  ppm) (Figure 4.19A) and some peaks cannot be assigned any more<sup>30</sup>, all the concerned regions are in the same helices as the substitutions introduced. In other words, we could not observe any global effect of the substitutions on the protein structure, even in the regions which are close in space to the mutation sites.

The aforementioned experiments indicated, that GGC1 might be in a very flexible state in DPC micelles. Helices in TM proteins are expected to be very rigid, due to increased strength of the backbone peptide hydrogen bonds. In addition, regions surrounded by the detergent micelle should be inaccessible to water molecules.

---

<sup>29</sup>Arginine side chains have very high nitrogen chemical shifts and should appear outside of the spectrum. However, discrete sampling of the frequencies and a limited minimum sampling interval results in assignment of faster oscillations to lower frequencies, which in NMR is defined as folding. Therefore, the position of these side chains are dependent on the sampling intervals and therefore may depend between different experiments.

<sup>30</sup>due to the large chemical shift differences or induced millisecond dynamics

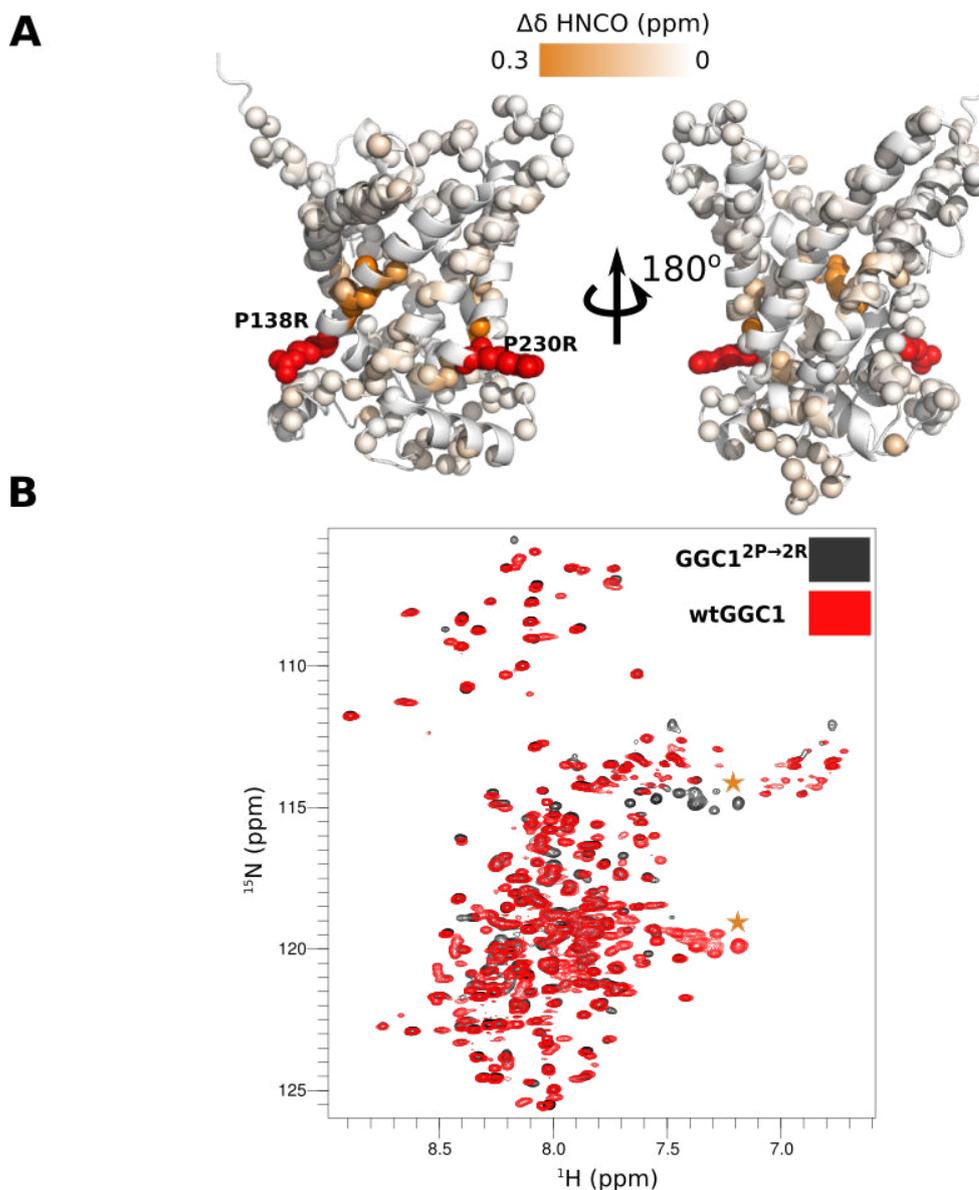


Figure 4.19: Chemical shift perturbations observed in GGC1 after introducing two proline to arginine substitutions (Pro138Arg, Pro230Arg), thereby creating GGC1<sup>2P→2R</sup>. A, chemical shift perturbations plotted on the homology model of GGC1. Spheres indicate the regions for which CSPs were determined. Colors correspond to the combined chemical shift perturbations in 3D BT-HNCO spectra, normalized to <sup>1</sup>H chemical shift, and are related to the color-bar above. Substituted residues are indicated as red spheres. B, Overlay of wild-type GGC1 (black) and GGC1<sup>2P→2R</sup> spectra (red). Orange stars indicate the arginine side-chain amides ( $\approx 7$  cross-peaks)<sup>29</sup>.

Therefore, a very slow exchange of protein amide and the solvent hydrogen atoms should be expected for a well-folded protein. To confirm that the structure of GGC1 might be very flexible we have used hydrogen-deuterium exchange experiments. These experiments are based on the fact that <sup>2</sup>H atoms are not detected in <sup>1</sup>H-based experiments, and all the regions where <sup>2</sup>H atoms are attached become invisible in

the NMR spectra.

GGC1 was transferred from H<sub>2</sub>O to D<sub>2</sub>O buffer by diluting the sample with a large excess of D<sub>2</sub>O. All the amide <sup>1</sup>H atoms, which are not covered by the detergent and are not involved in strong hydrogen bonds, would therefore be expected to exchange fast with the <sup>2</sup>H atoms (even at low pH of 6.0), and their signal would therefore disappear. On the other hand, no change is anticipated for the buried residues or those which are involved in stable secondary structures.

Initial GGC1 sample in H<sub>2</sub>O displayed a strong signal as can be seen from 1D-

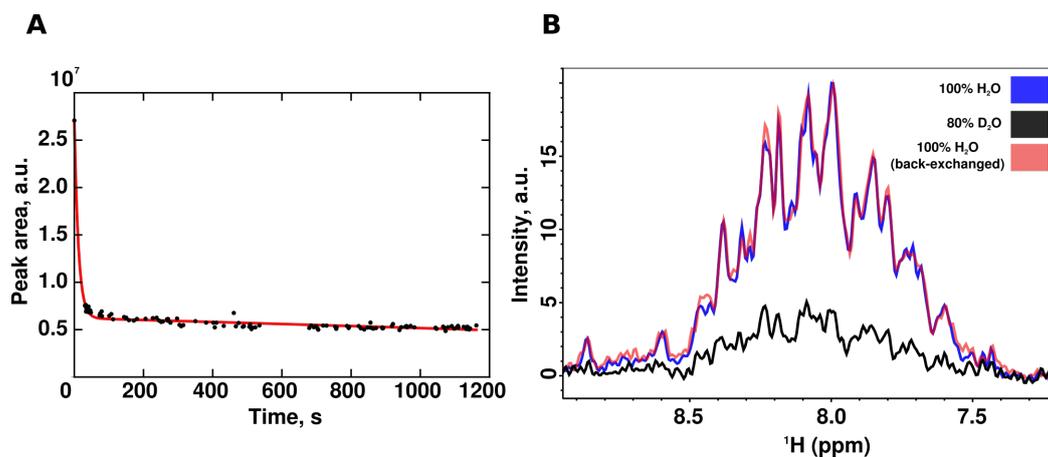


Figure 4.20: Hydrogen-deuterium exchange data. A, time-depend decay of GGC1 signal in the presence of 80% D<sub>2</sub>O. Due to the very fast H/D exchange, only 1D spectra could be recorded. B, recovery of GGC1 signal after dilution of D<sub>2</sub>O buffer to water. As can be seen, the signal recovery is complete. This shows that the signal drop in (A) is due to exchange of amide protons to deuterium, and not due to protein aggregation.

SOFAST experiment (Figure 4.20B, blue line). However, after we have diluted the sample with a large excess of D<sub>2</sub>O buffer, the signal disappeared almost immediately (<30 s) (Figure 4.20A, 4.20B black line). This could indicate that all the <sup>1</sup>H atoms are exchanged by <sup>2</sup>H atoms, or that the sample precipitated. To confirm that the latter is not the case, we have diluted the protein sample in D<sub>2</sub>O with H<sub>2</sub>O and concentrated the protein to the initial concentration. After this procedure, which took  $\approx$ 30 minutes, we observed that the signal intensity returned to the initial value, as measured before exchange to deuterium (see Figure 4.20B, red line).

This experiment supports our initial observations that GGC1 is very flexible, and might have a substantial population of unfolded state.

#### 4.9.2 AAC3 in DPC micelles lacks compact, well-folded structure

We have further questioned, if GGC1 is exceptional among the mitochondrial carriers, or the tertiary structure is missing in other carriers as well.

We have started the comparison of our AAC3<sup>c-salt-bridge</sup> mutant with the AAC3 pro-

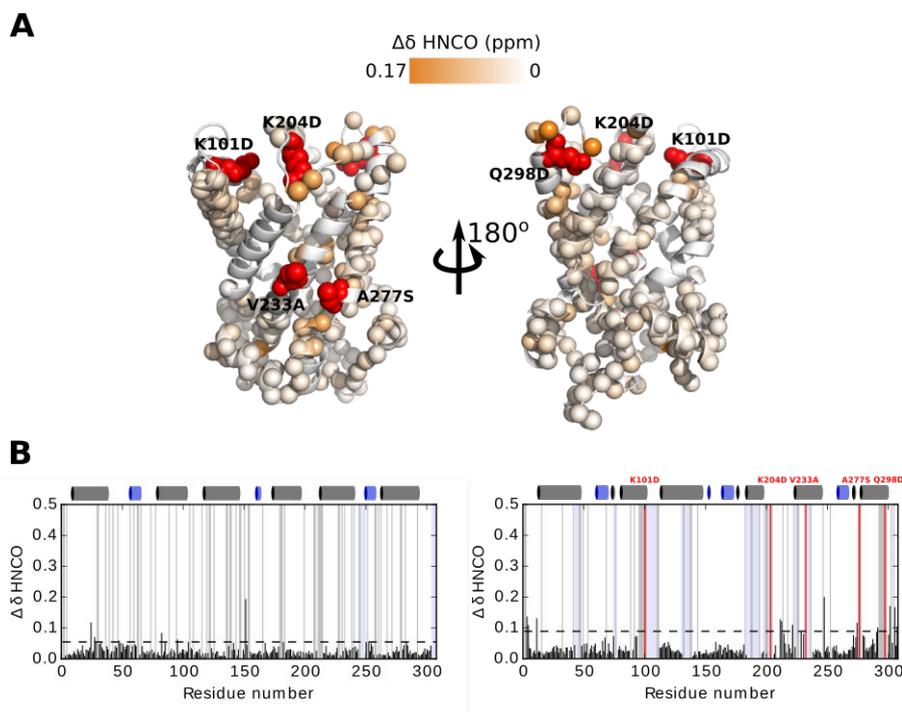


Figure 4.21: Chemical shift perturbations resulting from the five mutations introduced into AAC3<sup>c-salt-bridge</sup> mutant, relative to the construct used by Bruschiweiler et al (200) (the Cys substitutions are not mentioned in the paper, but can be seen in the BMRB submission). A, a representation of chemical shift perturbations on a X-ray structure of AAC3. The magnitude of combined chemical shift per <sup>15</sup>N, <sup>13</sup>C and <sup>1</sup>H CSPs, normalized to the <sup>1</sup>H chemical shift is represented by the color-bar. Red spheres indicate the mutation sites, while the regions without spheres correspond to the residues for which assignment could not be transferred. B, chemical shift differences between the BMRB provided shift lists and our samples for GGC1 (left) and AAC3<sup>c-salt-bridge</sup> mutant (right). Gray and blue cylinders represent TM and matrix helices, respectively. Red lines in AAC3<sup>c-salt-bridge</sup> mutant plot indicate mutation sites and the corresponding substitutions.

tein. The assignment of AAC3 was available on the Biological Magnetic Resonance Data Bank (BMRB). We were therefore able to transfer the assignment from close to wild-type AAC3 to our mutant protein.

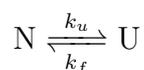
Besides the three charge inversion substitutions at the cytosolic salt-bridge network, AAC3<sup>c-salt-bridge</sup> mutant differed from AAC3\* in two positions in transmembrane helices. Brushweiler et al. replaced the 4 cysteines in AAC3 with either Val or Ala residues. In AAC3<sup>c-salt-bridge</sup> mutant these cysteines were replaced by Ala or Ser residues. Therefore, there were two additional differences between the constructs: V233 and A277 in Bruschiweiler et al. construct were replaced with A233 and S277 in ours.

Despite differences in 5 sites, 3 of which contained charge inversion mutations (K→D or Q→D), the chemical shifts differences were very small between the two constructs (<0.2, Figure 4.21A, B). The overall difference in chemical shift between AAC3 and AAC3<sup>c-salt-bridge</sup> was very small, comparable to the differences we have observed while transferring the peaks of GGC1 deposited on BMRB to an identical GGC1

construct (in Figure 4.21B compare right and left). Stronger chemical shift differences were only present in the regions close in sequence to the mutation sites (Figure 4.21B) and no significant perturbations were observed in the regions which are close in space to the substitutions.

This observation is similar to what we found for GGC1. We wanted to further confirm that AAC3 lacks a compact structure, and performed thermal unfolding experiments by NMR.

The simplest model of a protein unfolding is a two-state unfolding. According to this model, protein might exist in either a compact fold, with a very well defined  $\Delta G$  minimum, or as an unfolded coil, and these two states are in equilibrium at any given temperature:



The folding and unfolding rates,  $k_f$  and  $k_u$  depend on the temperature according to the Arrhenius equation:

$$k = e^{-E_a/(RT)}$$

$R$  is universal gas constant,  $T$  is the temperature and  $E_a$  is the activation energy. A single activation energy is given, as all the tertiary interactions in the protein are assumed to be broken in a single unfolding step (global unfolding). The populations of folded ( $N$ ) and unfolded ( $U$ ) states are thereby non-linearly dependent of temperature and a sigmoid unfolding curve is observed.

The process of folding is generally slow on the NMR time-scale <sup>31</sup> and upon heating the peaks corresponding to the folded state should decrease in intensity, while new peaks corresponding to the unfolded state should appear.

In contrast to the soluble proteins, the folding of membrane proteins is poorly understood, as generally the folding process is non-equilibrium and folding intermediates might exist. However, it was previously reported (194) that a mitochondrial AAC carrier from a thermophilic fungus in DDM micelles displays thermal unfolding with a clear sigmoidal transition from a folded to an unfolded state.

We have performed such unfolding studies on AAC3<sup>c-salt-bridge</sup> mutant and ORC1 in DPC micelles by NMR. The experiments were based on incrementing the temperature in discrete steps from 30 °C to 60 °C and recording a 2D HSQC spectrum at each temperature point. The lowest temperature correspond to the temperature at which we have performed all our NMR experiments.

ORC1 spectra displayed a gradual shift towards higher chemical shift values, which is consistent with the lengthening of the amide N-H bonds. No disappearance (or appearance) of the cross-peaks was observed, which could correspond to slow exchange

<sup>31</sup>On a 600 MHz spectrometer, an exchange process in which two populations are involved, with chemical shift difference of  $\approx 0.5$  ppm on <sup>1</sup>H between the two populations, will result in two sets of peaks if the exchange rate is  $< 300$  per second. Protein folding is generally slower (for example (240))

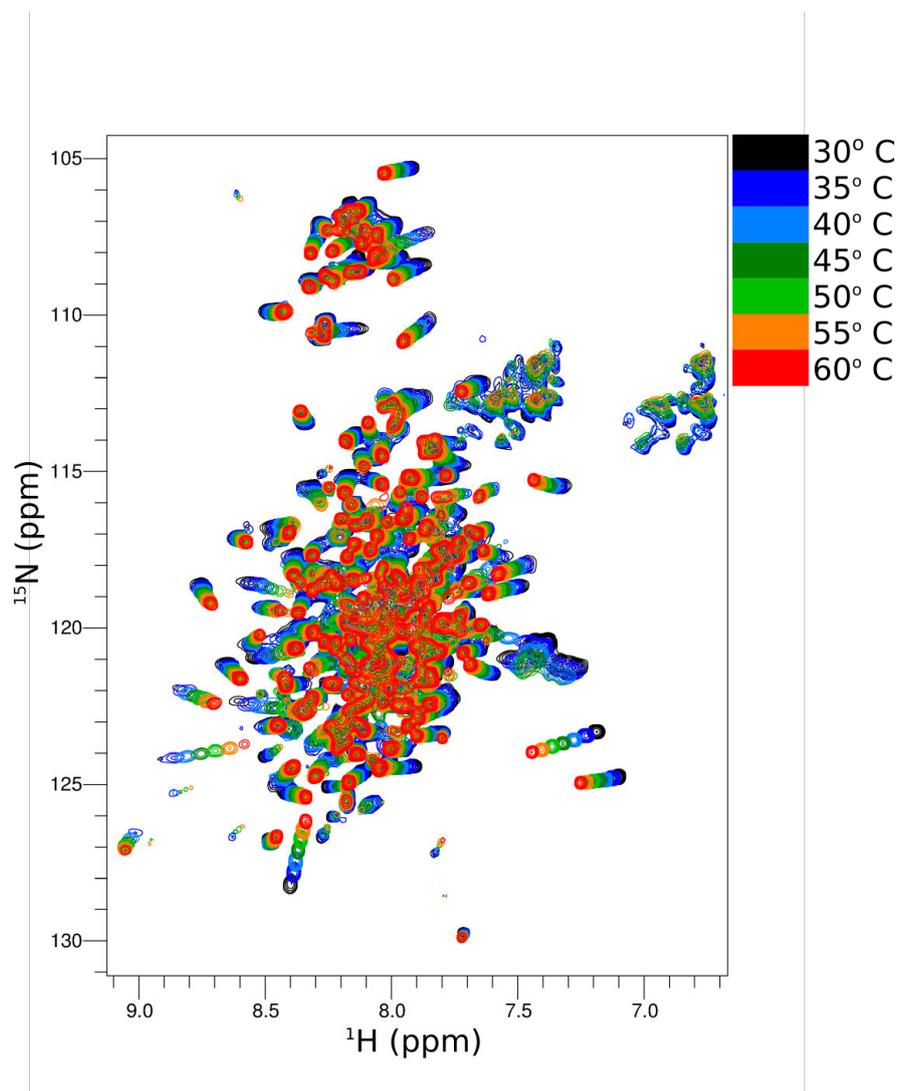


Figure 4.22: Temperature-dependent chemical shift perturbations of yeast ORC1. Overlay of several 2D HSQC spectra at different temperatures, as indicated in the vertical color-bar. The spectra were referenced according to a set of least shifting residues, so the systematic shift due to the temperature-dependent changes of amide-hydrogen dissociation rates can be neglected. As can be seen, the spectra become less disperse in the  $^1\text{H}$  dimension, indicating the loss of structure. However, all the shift magnitudes are linear, suggesting that no unfolding transition occurs.

process. After superimposing spectra at different temperatures, we could observe that the spectra became narrower at increasing temperatures, but all the chemical shift perturbations appeared to be linear (Figure 4.22). To confirm this, we have assigned all the residues for which we could follow CSPs and calculated combined  $^{15}\text{N}$ - $^1\text{H}$  CSP. Plots of CSPs against temperature for each individual peak we could assign indicated a linear dependence on temperature, and not a sigmoid one which would be representative of a cooperative unfolding. Plots showing one residue in each TM helix are displayed in Figure 4.23 as an example.

This behaviour of ORC1 is not consistent with a global unfolding process in any region of the protein. It is more likely, that the protein exists in multiple conformations and the resulting CSPs represent different populations of these conformations at different temperatures.

We have performed the same thermal unfolding experiments on AAC3<sup>c-salt-bridge</sup>

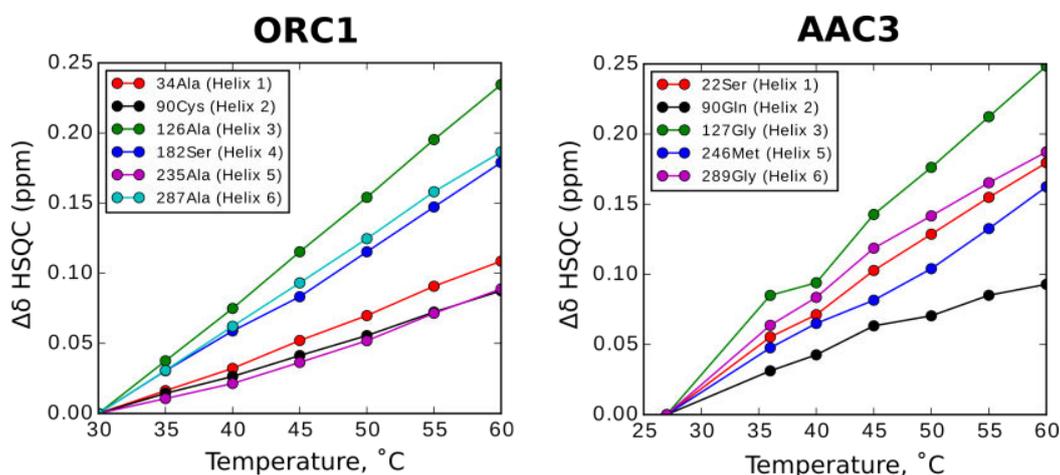


Figure 4.23: Data as in the figure above plotted as a combined  $^1\text{H}$   $^{15}\text{N}$  chemical shift perturbation dependence on the temperature for several residues present in different TM helices in ORC1 (right) and in AAC3<sup>c-salt-bridge</sup> (left). No sigmoidal transitions, which would indicate the unfolding of the tertiary structure of the protein are present. Besides the data shown, virtually all of the peaks which could be followed across the temperature range indicated linear transitions.

mutant (Figure 4.23, right) and the behaviour we observed was essentially the same as for ORC1, suggesting that this protein does not have a compact tertiary structure. We have further compared the spectra of AAC3<sup>c-salt-bridge</sup> before and after the heating cycle. 2D HSQC spectrum indicated that even after prolonged incubation at high temperatures, AAC3 regained the same "structure" as we have observed before the heating, as all the peaks returned to the initial positions (Figure 4.24A). Some protein, however, was lost (most likely due to the aggregation), as the signal intensity was lower after the heating (see Figure 4.24B).

#### 4.9.3 Summary of the observations on the tertiary structure of mitochondrial carriers in DPC

Our experimental observations strongly indicate that all studied mitochondrial carrier lack tertiary structure in DPC:

- Both GGC1 and AAC3 do not experience global CSPs upon introduction of large substitutions
- GGC1 displays a very fast hydrogen-deuterium exchange
- Both AAC3 and ORC1 display linear thermal unfolding transitions

Table 9: NMR observable which can provide structural information

Observable	Information
Chemical shifts	Secondary structures
Nuclear Overhauser effects	Short range (<5-6Å) distance information
Residual dipolar couplings	Orientation of structural elements
Solvent PREs	Solvent exposure
Paramagnetic labels	Long range distance information
Relaxation measurements	Flexibility information

These observations are in accord with a previous study on UCP1 (236), where the protein did not display a sigmoidal unfolding profile in 6 mM DPC. We have to note, that the concentration of DPC in our samples is at least 10 to 20 fold larger as in the previous study.

NMR provides many observables for the investigation of protein structures. We therefore wanted to obtain more detailed understanding about the structural perturbations occurring in mitochondrial carriers in the presence of DPC.

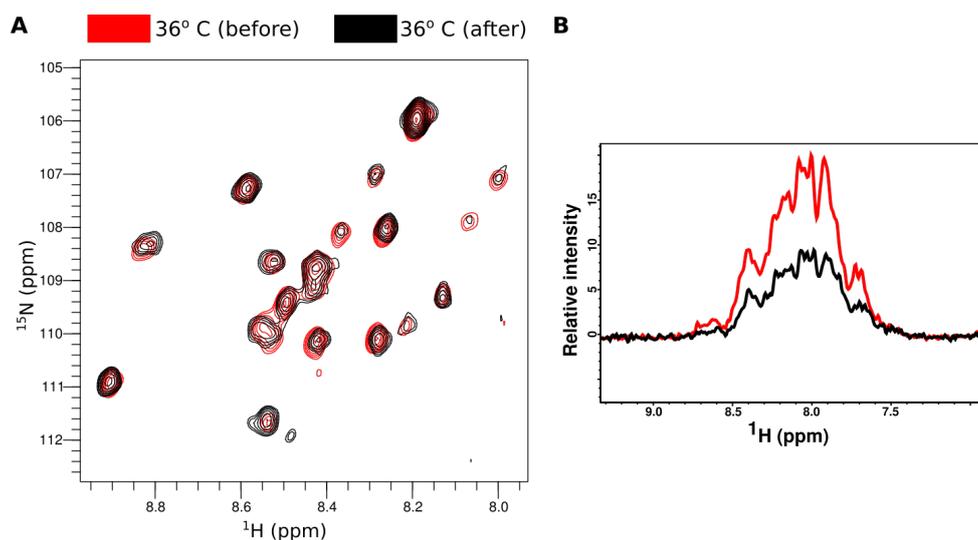


Figure 4.24: A, Gly region of AAC3<sup>c-salt-bridge</sup> mutant spectra before (red) and after (black) heating cycle (up to 60° C). The peak positions in the two spectra are in excellent agreement, indicating that the protein returned to the initial structure (or ensemble of structures) after the heating. However, some of the protein became undetectable, presumably due to the formation of aggregates, as shown by the drop in intensity of the NMR signal. B, comparison of the first point of the 2D BT-HSQC spectra representing the peak intensities of AAC3<sup>c-salt-bridge</sup> before (red) and after (black) heating.

#### 4.10 Investigation of mitochondrial carrier structure in DPC micelles

NMR provides several observables, which can be used for the structural investigations:

We have tried to use these observables to provide information about mitochondrial carriers in their DPC denatured state.

As was discussed in the beginning of the section 4.9.1, we could not observe NOEs between methyl groups of GGC1. Although we could not get any tangible information on the structure of GGC1 from this observation, it indicated that the TM helices in GGC1 are very flexible and/or that the helices are not in close contact. This observation is in agreement with the previously reported residual dipolar coupling (RDC) data on different TM helices in GGC1 (228). RDC data indicated, that all TM helices in GGC1 are very dynamic on nanosecond-millisecond time scale, as a common alignment tensor could not be found for a stretch longer than 7 residues in the same  $\alpha$ -helix. It has further been shown, that TM helices had very different generalized order parameters from each other. Considering a compactly folded protein, the regions in close spatial proximity would be expected to have similar dynamics if these two regions are linked by tertiary contacts (van der Waals, hydrogen bonds, etc.). The order parameter variation between adjacent helices in a protein thereby suggests that there are no well packed structural regions.

#### 4.10.1 Secondary structures of mitochondrial carriers

NMR spectroscopy can provide accurate information about the secondary structures of proteins, such as  $\beta$ -strands,  $\alpha$ -helices and disordered regions. This is possible, because the chemical shifts of several nuclei (N,  $H^N$ ,  $H^{C\alpha}$ ,  $C\alpha$ ,  $C\beta$ ,  $C^{\text{carbonyl}}$ ) depend on the  $\phi$  and  $\psi$  dihedral angles of the peptide bond. As different secondary structures differ largely in their peptide bond dihedral angle distribution, it is possible to obtain secondary structure information from the chemical shifts.

Talos+ is a commonly used software to perform such predictions. It utilizes a database of high-resolution NMR structures and the corresponding secondary chemical shift<sup>32</sup> information. Sequence context and secondary chemical shifts from the database are compared to the corresponding information of a given protein, and if a good agreement is found, the secondary structure is assigned. Such predictions display near 90% accuracy for soluble proteins.

In addition to secondary structure predictions, Talos+ also provides predictions on the flexibility of different regions in the form of order parameter ( $S^2$ ). Rigid regions of a protein typically have high  $S^2$  values and they decrease due to increasing flexibility. As disordered regions typically have very low order parameters, secondary chemical shift information can be used to provide the estimation of disorder (241). It is, however, not clear if Talos+ predictions remain accurate for a transmembrane protein in detergent. We have therefore performed predictions for two proteins, TSPO and DsbB, for which NMR structures were solved in DPC. The obtained predictions were in an excellent agreement with the NMR structures. This provides some reassurance that Talos+ can also be used for prediction of membrane protein secondary structures. However, Talos+ secondary structure predictions were used in the structure determination procedure of both proteins, therefore we cannot reject the possibility

---

<sup>32</sup>Secondary chemical shifts correspond to the difference between the observed chemical shifts and chemical shifts of a random coil

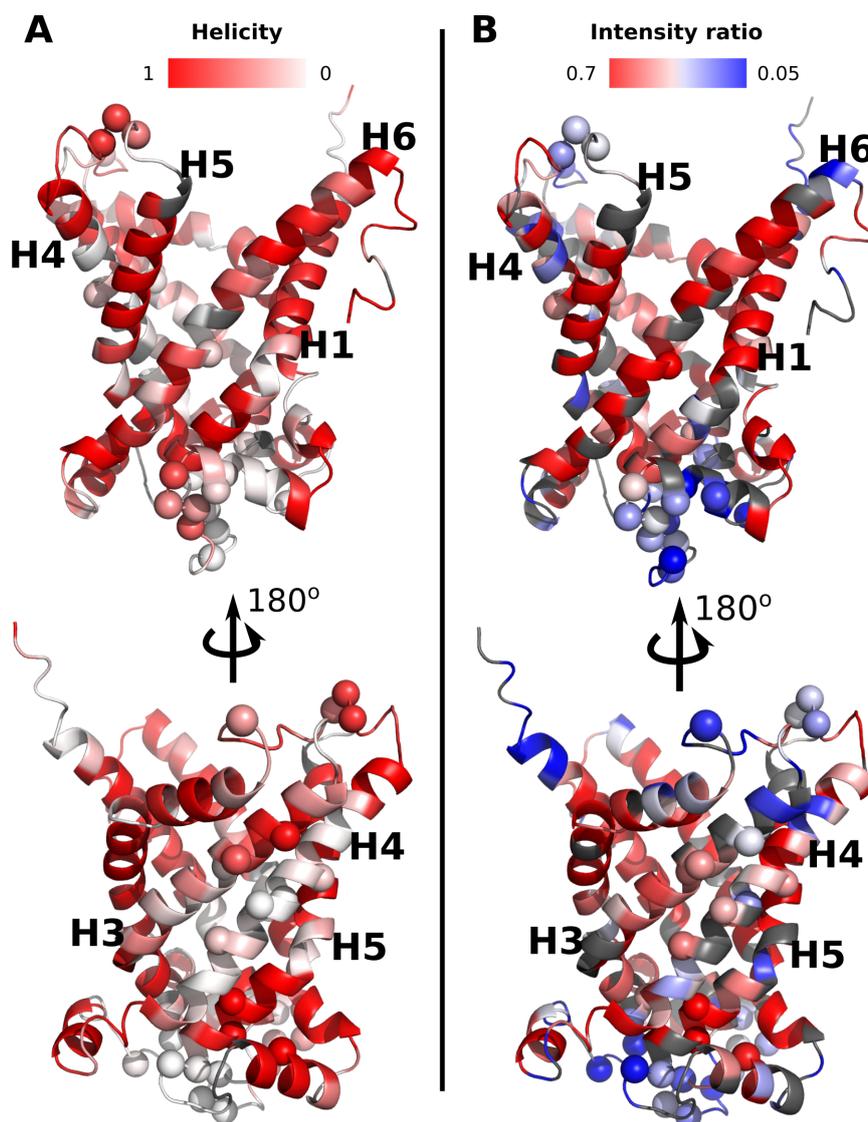


Figure 4.25: A, helical content predicted by Talos+ according to the  $C\alpha$ ,  $C\beta$ ,  $CO$ ,  $N$  and  $H$  chemical shifts, displayed on a homology-based structural model of GGC1. As depicted in the horizontal color-bar, larger amount of helical content is represented in stronger red shade. Gray regions indicate residues for which prediction was not available due to the missing assignment. Spheres represent residues displaying non-flat relaxation dispersion profiles. B, Ratio of peak volumes after and before addition of 1.5 mM of gadodiamide to GGC1 sample. The intensity loss corresponds to the color-bar above. A good relationship can be seen between the helicity content and solvent exposure.

that TSPO and DsbB structures themselves were not influenced by Talos+ predictions.

As the  $N$ ,  $H^N$ ,  $C\alpha$ ,  $C\beta$ ,  $C^{\text{carbonyl}}$  chemical shift information for mitochondrial carriers was available from the backbone resonance cross-peak assignment, we have used secondary structure predictions by Talos+ to get some structural information about

the carriers in DPC.

### **GGC1**

Although from the binding studies, NOE measurements and hydrogen-deuterium exchange experiments we inferred that a compact tertiary structure is lacking in GGC1, the predicted secondary structures are largely in agreement with the structural model of GGC1, as can be seen from Talos+ helicity values (See Figure 4.25A). We have observed, however, some local regions in all TM  $\alpha$ -helices, which were predicted to be disordered. Interestingly, most of these regions are in the vicinity of Pro residues (see Figure 4.S23). As Pro residues disrupt the helical structure, such disorder may be anticipated and flexibility around these Pro residues might even be involved in the biological function of the protein.

We have observed, however, that helices 3 and 4 are predicted to be disordered even in the regions where there are no Pro residues (see Figure 4.25, bottom). Such disorder is highly surprising, since TM segments exposed to lipid or detergent environment have never been observed to adopt random coil conformation, and most of the available biophysical evidence suggest that they should not exist.

We have therefore concluded, that besides displaying perturbations in the tertiary structure, GGC1 is also affected on secondary structure level by DPC detergent.

### **ORC1**

In contrast to GGC1, the secondary structures in DPC solubilized ORC1 differ significantly from those observed in the structural model (Figure 4.26A). Even with a limited assignment we can observe that the whole TM region of ORC1 appears to be disordered, involving helices 1, 3, 5 and 6. This extent of disorder in TM regions is highly unlikely. It also explains why we could not observe any substrate binding in ORC1 - the electrostatic cavity in this protein likely does not exist due to such extensive perturbations.

In addition, the exposure of hydrophobic TM segments of ORC1 could explain the strong aggregation propensity we have observed for this protein. The overall peak intensity of ORC1 sample decreases substantially after a week of measurements at 30 °C, while for example GGC1 samples could be measured for a several weeks without displaying noticeable reduction in sensitivity (As shown in the figure 3.4 in materials and methods)

### **AAC3**

The comparison of secondary structures of AAC3 between DPC solubilized protein and the crystal structure solved in DDM reveal, that AAC3 shows intermediate (between GGC1 and ORC1) secondary structure perturbations (Figure 4.27A). Some regions which would be expected to form TM segments appear to be disordered, even when there are no Pro residues in the vicinity. The largest perturbations appear to be in the unassigned helices 4 and 5. Although very limited information is available for these helices, several assigned residues indicate that there is a large drop in the predicted order parameter in helix 4 (Figure 4.27B) as well as in helix 5. Helix 3 furthermore displays disorder of its cytosol-facing end. Helix 2, in addition, displays

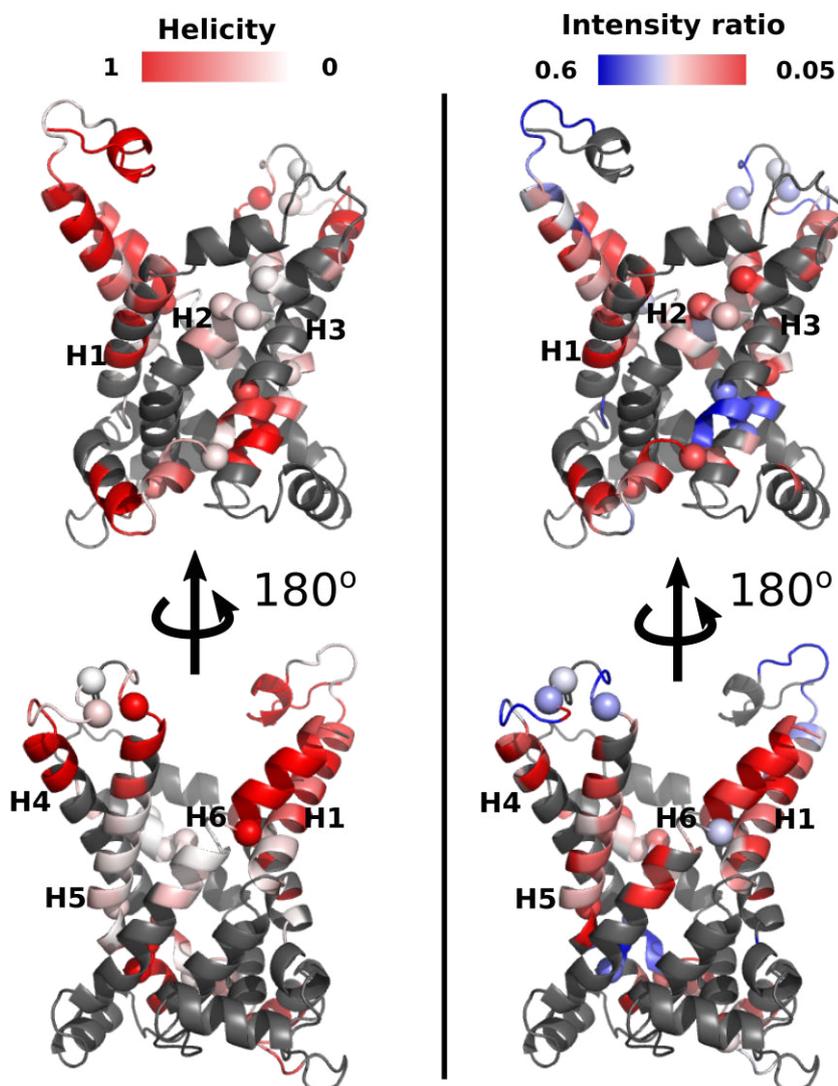


Figure 4.26: A, helicity content predicted by Talos+ according to the  $C\alpha$ ,  $C\beta$ , CO, N and H chemical shifts, displayed on a homology-based structural model of ORC1. As depicted in the horizontal color-bar, larger amount of helicity content is represented in stronger red shade. Gray regions indicate residues for which prediction was not available due to the missing assignment. Spheres represent residues displaying non-flat relaxation dispersion profiles. B, Ratio of peak volumes after and before addition of 3 mM of gadodiamide to GGC1 sample. The intensity loss corresponds to the color-bar above. A good relationship can be seen between the helicity content and solvent exposure.

several regions along its span which appear to be disordered. There is only a single proline in the middle of helix 2.

We have therefore observed deviations of secondary structure propensity of all three mitochondrial carriers in the presence of DPC micelles, as compared to their predicted models. Interestingly, the amount of overall perturbations seems to corre-

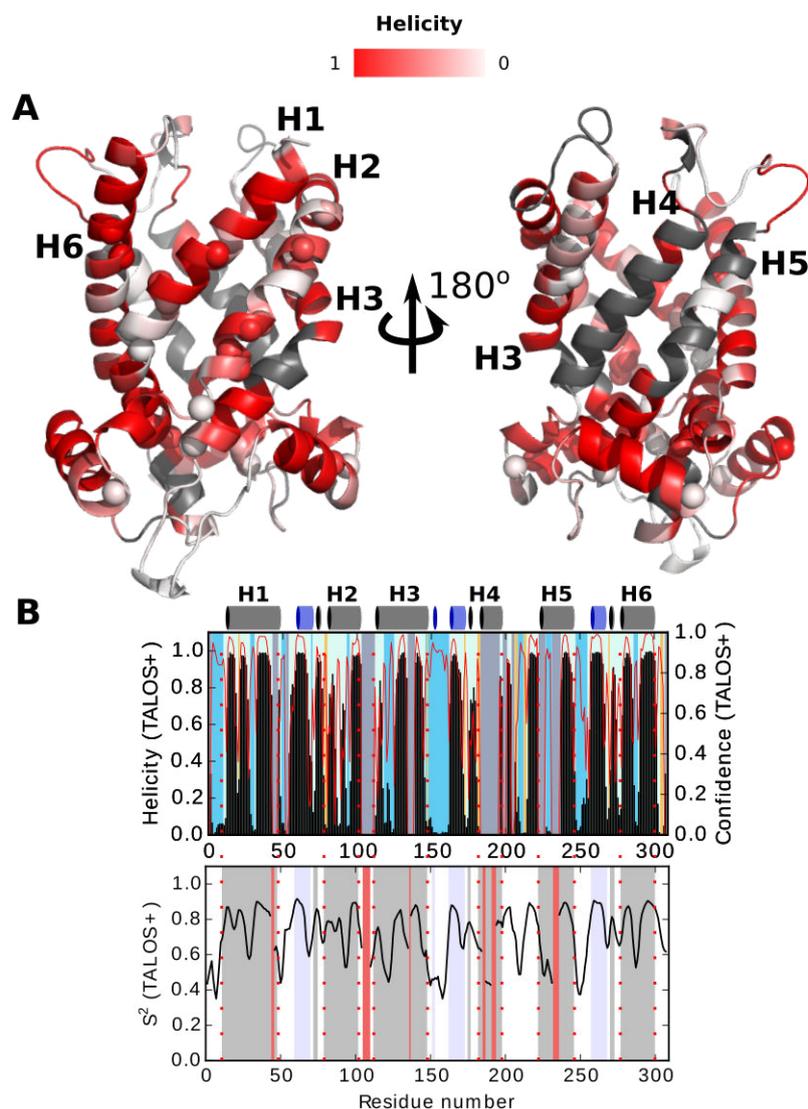


Figure 4.27: A, helical content predicted by Talos+ according to the  $C\alpha$ ,  $C\beta$ ,  $CO$ ,  $N$  and  $H$  chemical shifts, displayed on a crystal structure of AAC3\*, where the loops which otherwise are not visible in the X-ray structure have been modeled. Gray regions correspond to regions for which predictions are not available due to the lack of assignment. Spheres indicate non-flat relaxation dispersion profiles as reported by Bruschiweiler et al. B, Talos+ predicted helicity content (top) and order parameter (bottom). Cylinders correspond to TM (gray) and matrix (blue) helices. Black bars indicate the predicted helicity. Blue background corresponds to dynamics residues, which are inherently unstructured. Light green areas correspond to unambiguous predictions, while yellow areas indicate ambiguity in prediction. Gray regions represent residues for which the prediction is not available. Red line corresponds to the confidence of prediction. Red regions in the order parameter plot represent residues for which  $S^2$  prediction is not available. Dotted lines delimit the helices.

late slightly with the observed CSPs upon ligand titration.

For example, GGC1 displayed largest chemical shift perturbations in the presence of purine nucleotides, and it appears that it has the least unexpected perturbations

in the secondary structures. ORC1, on the other hand, displays no binding of L-ornithine and has the most perturbed structure. As for AAC3, we have limited information. If we assume that the AAC3<sup>c-salt-bridge</sup> mutant has the same CSPs as the wild-type AAC3<sup>33</sup>, which is most likely the case, then it appears that observed weak CSPs upon purine phosphonucleotide addition are in agreement with relatively large perturbations in the secondary structure.

#### 4.10.2 Transverse relaxation rate measurements for GGC1

As presented above, by performing Talos+ chemical shift based analysis of the secondary structures, we have observed that expected TM  $\alpha$ -helical regions are disordered in DPC reconstituted mitochondrial carriers. This observation is largely surprising, as  $\alpha$ -helices are expected to be very stable in a hydrophobic environment. We wanted to confirm if the disorder predictions of Talos+ were accurate. For this purpose we have measured transverse relaxation rates  $R_2$  of  $^{15}\text{N}$  of GGC1<sup>2P→2R</sup><sup>34</sup>.  $R_2$  relaxation rates are directly proportional to the correlation time  $\tau_C$  and the local order parameter  $S^2$ . Small molecules which tumble at a fast rate show low  $R_2$  values. However,  $R_2$  increases proportionally to the size of the molecule. Large molecules (such as GGC1 in DPC micelles), therefore, have very large  $R_2$  relaxation rates. However, even a large molecule might possess flexible loops. These loops move very fast with respect to the rest of the molecule, and thereby their local  $\tau_C$  is much smaller, resulting in smaller  $R_2$  rates.

After performing  $^{15}\text{N}$   $R_2$  measurements, with a pulse sequence adapted to TROSY

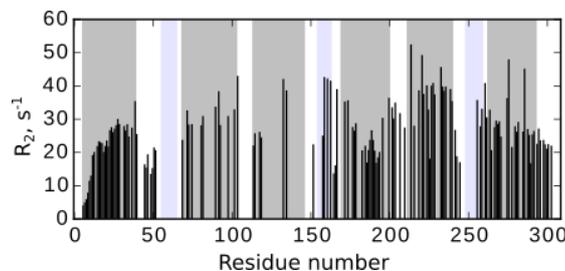


Figure 4.28: Residue-wise distribution of transverse relaxation rates for GGC1<sup>2P→2R</sup>. The agreement is very good with the previously published data for wild-type GGC1 (228)

coherence selection (242), we have observed that most of the TM segments in GGC1<sup>2P→2R</sup> display  $^{15}\text{N}$   $R_2$  rates of 30 to 40  $\text{s}^{-1}$ , which are consistent with the relaxation rates of a large, deuterated protein in DPC micelles (Figure 4.28). However, the helix 4, which appeared to be disordered in secondary structure predictions, displayed a lower  $R_2$  rates ( $\approx 20 \text{ s}^{-1}$ ). This observation therefore suggests that in

<sup>33</sup>for which CSP for only a single residue was reported

<sup>34</sup>At the time of the measurement, a sample of wild-type GGC1 was not available; due to large similarity of GGC1 and GGC1<sup>2P→2R</sup> spectra, we expected similar behaviour between these two proteins.

agreement with Talos+ predictions, it contains large degree of disorder. These observations on GGC1<sup>2P→2R</sup> were in a very good agreement with previous report on the <sup>15</sup>N R<sub>2</sub> rates of wild-type GGC1 (228). We have therefore confirmed likely disordered nature of one of the TM helices of GGC1 in DPC micelles.

#### 4.10.3 Probing solvent exposure by solvent PREs

Our secondary structure predictions and <sup>15</sup>N R<sub>2</sub> measurements have indicated, that mitochondrial carriers in DPC likely contain disordered regions in their TM helices. This behaviour is very unlikely for a helical segment in a hydrophobic environment. For example, TM segments of all membrane proteins with known structure are either  $\alpha$ -helical or  $\beta$ -sheet. These two folds allow formation of the maximum number of peptide hydrogen bonds. It is thought that this behaviour is dictated by the large free energy penalty of partitioning non-hydrogen bonded polar groups into a hydrophobic phase.

Lipid exposed hydrogen bonds also appear to be shorter by approx 0.1Å as compared to water exposed bonds.

If these regions were exposed to the solvent, however, disorder could be easily explained. Indeed, hydropathy predictions using Kyte and Doolittle scale (243) indicate that helix 4 in GGC1 is slightly less hydrophobic than the remaining helices. However, this information is not very conclusive, as the helix 2 also appears to be less hydrophobic, but it is well structured.

To probe the solvent exposure of the mitochondrial carriers, we have employed a soluble paramagnetically active compound, gadodiamide. Gadolinium ion, Gd<sup>3+</sup>, contains an unpaired electron which induces relaxation of nuclear spins if positioned in close proximity. The increased relaxation manifests as peak broadening (and therefore decrease of peak intensity).

Peaks corresponding to the atoms in solvent exposed regions (e.g. loops), would therefore be expected to be broadened by the paramagnetic effect of Gd<sup>3+</sup>. On the contrary, atoms of TM segments surrounded by detergent should not be substantially affected by the PRE agent.

#### GGC1

We have recorded a 3D BEST-TROSY HNCO spectrum of GGC1 in the presence of 1.5 mM of gadodiamide. After calculating the ratio of peak volumes in the presence and absence of PRE agent, we could observe that the regions where peaks were broadened (solvent exposed regions) were in a good agreement with the structural model of GGC1. All the TM regions of GGC1 showed less peak broadening as compared to the loop regions (see Figure 4.S24B).

Loops, which were located on both cytoplasmic and matrix side of the protein displayed a large decrease in intensity ( $\approx$ 80-90%). The majority of the peaks in helices H1, H2, H5 and H6, as well as the matrix helices showed relatively low peak-broadening. The decrease of intensity in these regions was only  $\approx$ 20-30% as compared

to GGC1 without gadodiamide.

Helices H3 and H4 were, however, more strongly affected by gadodiamide as compared to the other transmembrane helices. The decrease in intensity observed in these helices was  $\approx 60\%$ .

Comparison of the secondary structure data from Talos+ and the PRE measurements has indeed revealed high level of similarity (Figure 4.28, Figure 4.S20). The loop regions, for example, which were predicted to be disordered by Talos+ appeared to be strongly affected by gadodiamide. Well structured regions, on the other hand, such as H1, H5, H6 showed much lower effect by the paramagnetic ion. The peaks in the disordered segments in helices 3 and 4, displayed more severe broadening. The observed peak-broadening in these helices was not as dramatic as for the loop regions. This, however, can be easily explained. As NMR provides information as an average of the whole molecules in the sample, regions with intermediate broadening most likely exchange between solvent exposed and solvent accessible states.

In addition to H3 and H4, disordered regions close to proline residues also indicated much more substantial solvent exposure. This raises a question whether these sites are indeed flexible in a natural lipid environment.

Another striking feature we have observed in GGC1 is the distribution of CPMG RD displaying residues. The majority of the residues displaying non-flat RD profiles are located either in disordered or solvent exposed regions. For example, residues displaying some of the largest dispersions (A178, A182, G184, F186, A187) are all located on the helix 4.

These dispersions thereby might reflect some DPC induced unfolding effects. We cannot, however, provide an accurate interpretation of the exact nature of conformational process which is the cause of millisecond dynamics in these regions.

### **ORC1**

We made similar observations while analyzing ORC1 peak broadening in the presence of 3 mM gadodiamide (see Figure 4.26). Cytosol facing loops, for example, were predicted to be disordered by Talos+ and they have displayed a substantial amount of broadening. In addition, solvent exposure was evident in some TM regions, especially in the helix 3, helix 4 and helix 5, which were also predicted by Talos+ to lack stable structure.

Furthermore, the residues displaying non-flat relaxation dispersion profiles were clustered around potentially disordered and solvent exposed regions. We have to note that among unassigned residues of ORC1,  $\approx 20\%$  displayed non-flat relaxation dispersion profiles. This might indicate much more pronounced dynamics in ORC1 that we can observe.

**AAC3** We could not perform solvent PRE measurements on AAC3 sample, as the concentration (and thus sensitivity) was not sufficient for 3D HNC0 experiments. The peak overlap was too large in AAC3 2D spectra to attempt accurate integration of peak volumes.

We however have noticed, that the non-flat RD profiles displaying residues were also clustered in proximity of structurally perturbed regions, in agreement with the two other mitochondrial carriers.

#### 4.11 Supplementary figures

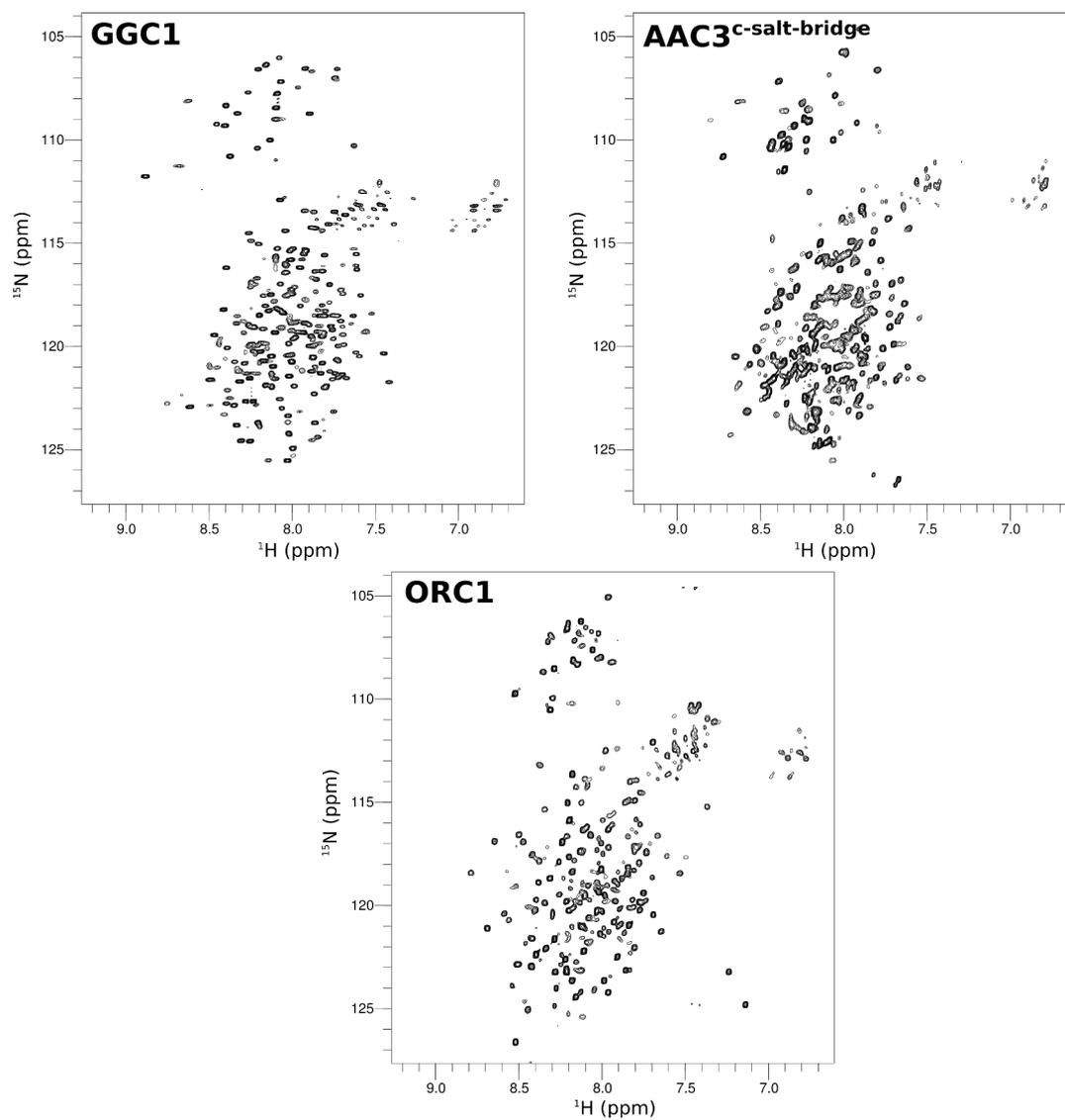


Figure 4.S1: 2D BEST-TROSY HSQC spectra of the three carriers in this study. A single set of peaks is present for all three carriers.

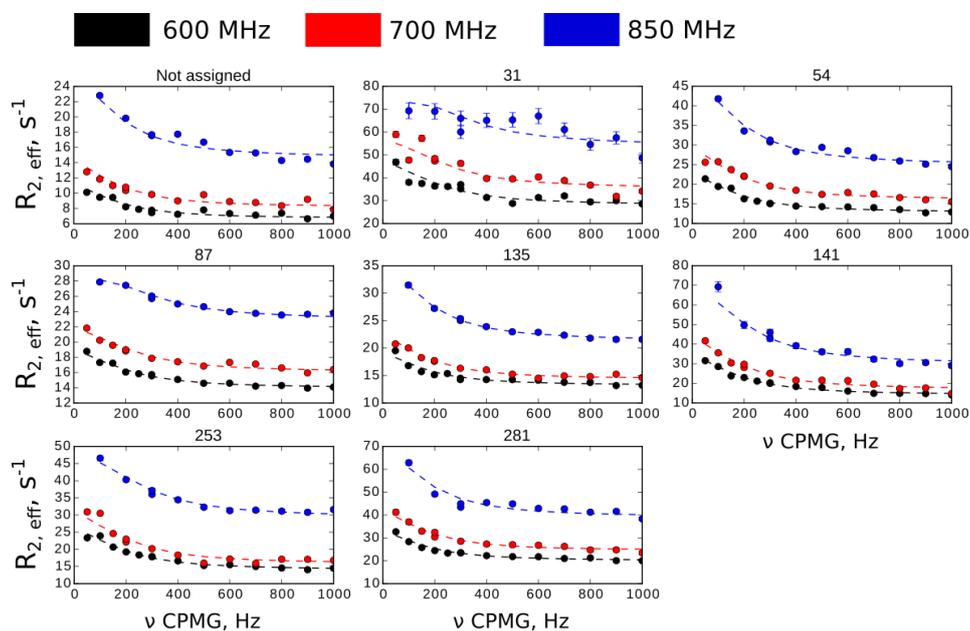


Figure 4.S2:  $^{13}\text{C}$  CPMG relaxation dispersion data for apo GGC1 recorded on 600 (black dots), 700 (red dots) and 850 (blue dots) MHz NMR spectrometers and the corresponding fits (dashed lines) to a global two-state model. Numbers above each panel indicate residue numbers of GGC1. The first panel indicates a Leu residue, for which assignment could not be obtained.

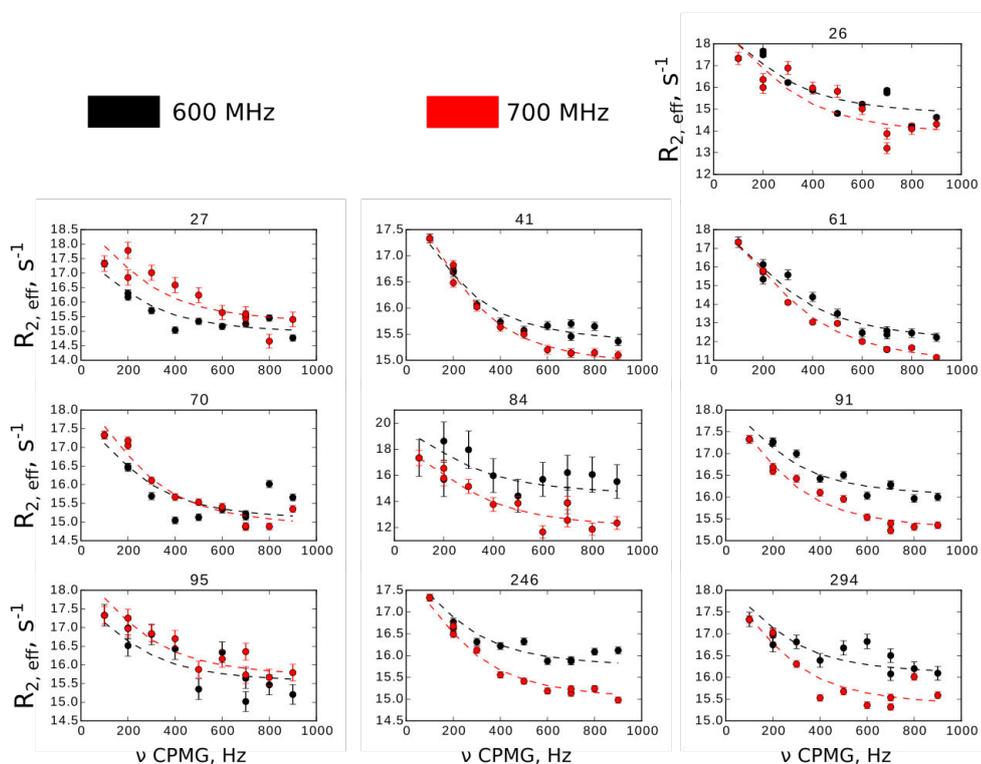
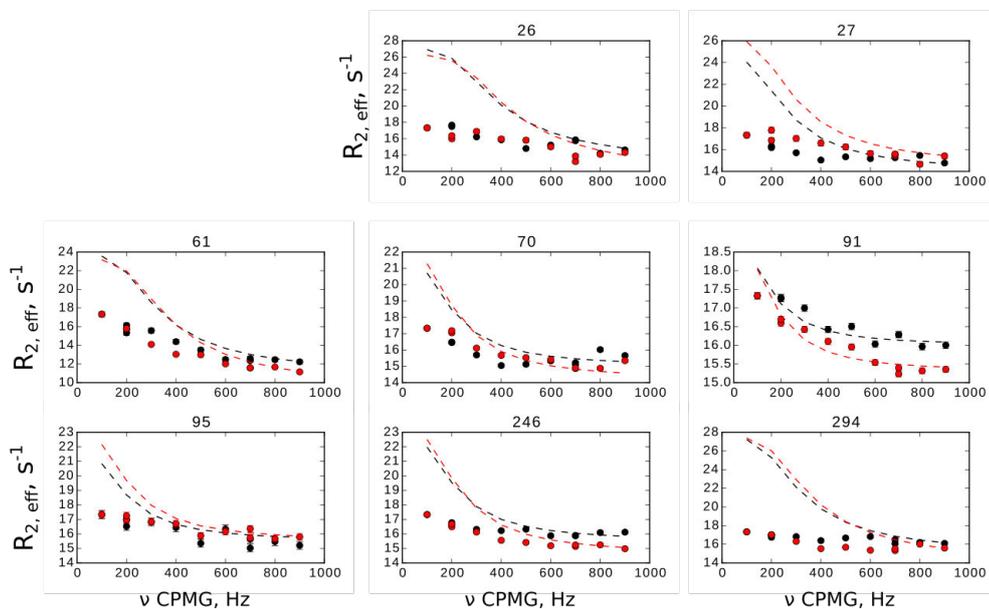
**A****Our fits****B****Bruschweiler et al fits**

Figure 4.S3: AAC3 fits reported in ref. (200) are quantitatively wrong. A, CPMG RD data reported by Bruschweiler et al, fitted with Chemex. These fits are more consistent with the data, as compared the ones reported in (200) (B).

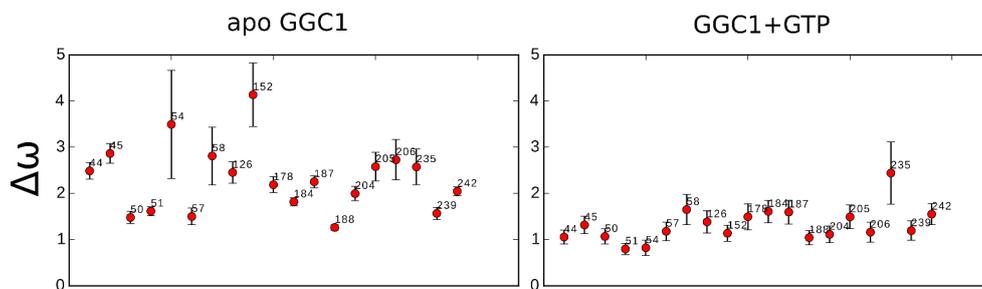


Figure 4.S4: Chemical shift difference ( $|\Delta\omega|$ ) obtained from fits of apo GGC1 (left) and GGC1+GTP (right). Numbers indicate residue number in GGC1. As can be seen, larger  $|\Delta\omega|$  values are attributed to apo GGC1. However, apo GGC1 also displays smaller populations of the exited state. The dispersion curves between the two proteins are very similar. Therefore, this difference likely occurs due to inability of the fitting software to deconvolute  $|\Delta\omega|$  and  $p_B$  values.

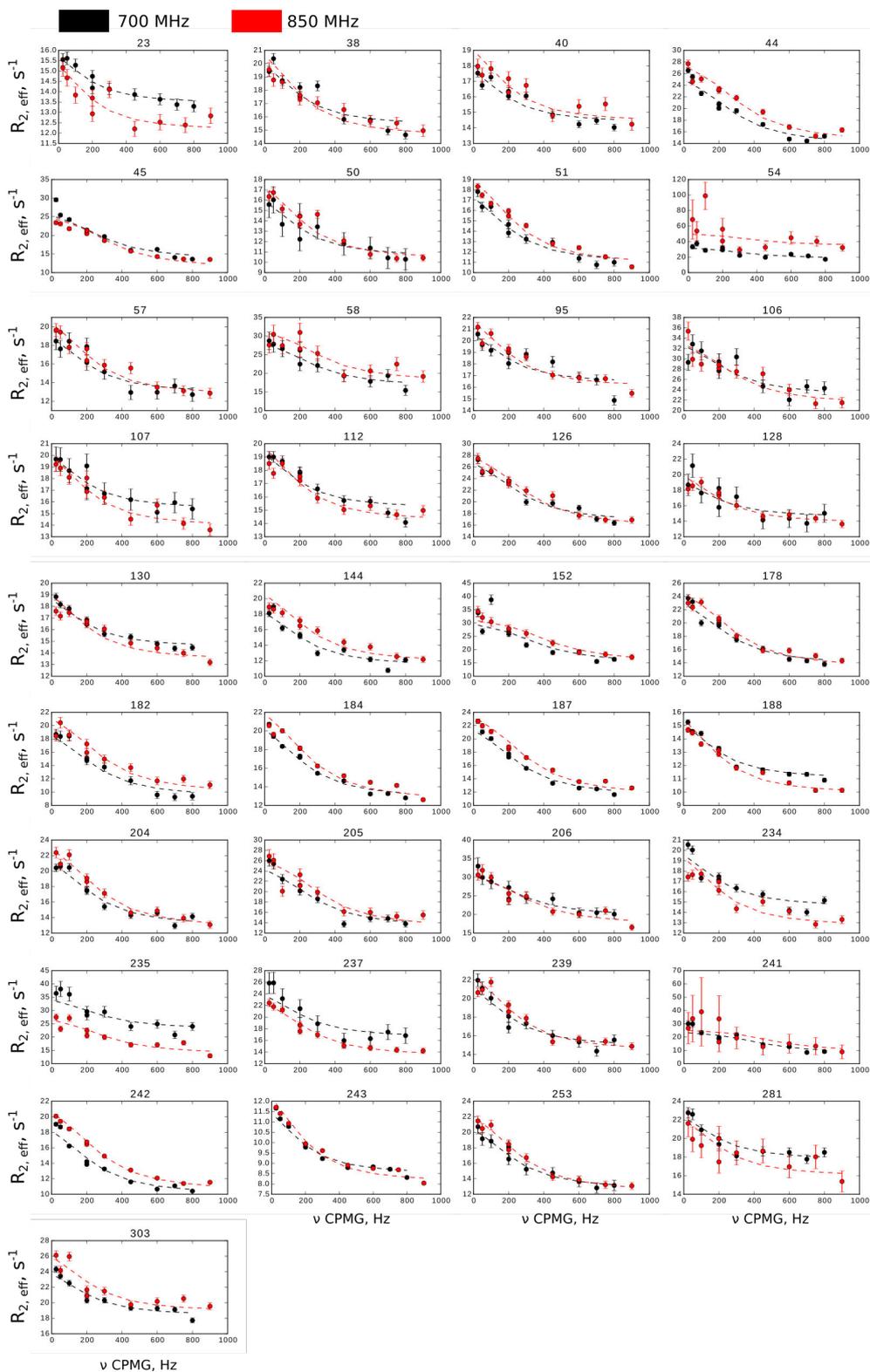


Figure 4.S5:  $^{15}\text{N}$  CPMG relaxation dispersion data for apo GGC1 recorded on 700 (black dots) and 850 (red dots) MHz NMR spectrometers and the corresponding fits (dashed lines) to a global two-state model. Numbers above each panel indicate residue numbers of GGC1.

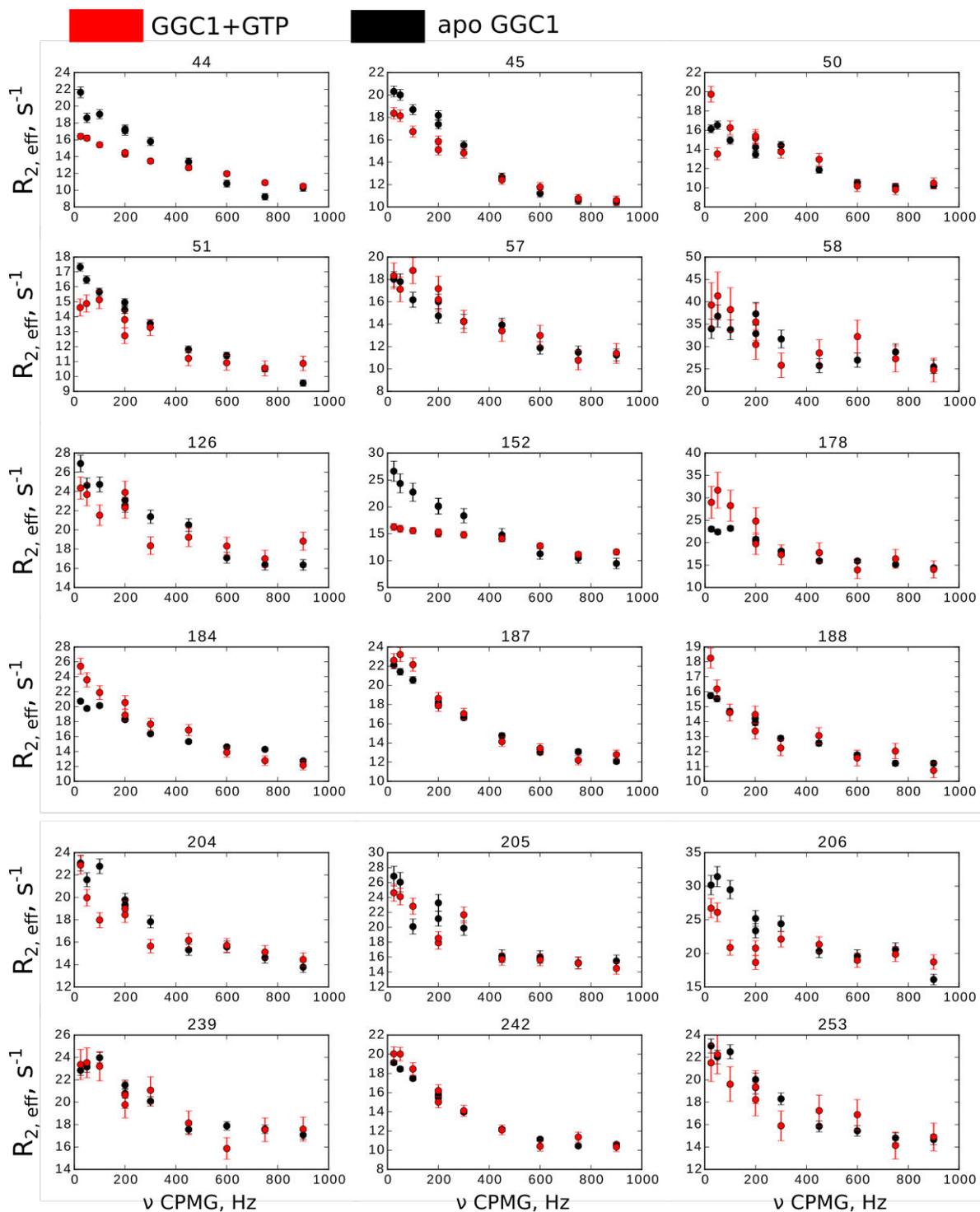


Figure 4.S6: Comparison of dispersion profiles observed for apo GGC1 (black dots) and GGC1 in the presence of 20 mM GTP (red dots) on 850 MHz spectrometer. Numbers above each panel correspond to the residue number in GGC1. Largest differences observed between the two samples correspond to the mitochondrial loops. The dispersions corresponding to the residues in TM and mitochondrial helices are hardly affected by the presence of GTP. To facilitate the comparison of the CPMG dispersion curves, an additive offset of  $R_{2,\text{eff}}$  (i.e. plateau  $R_{2,\text{eff}}$ ) was added to some residues. Such an offset does not have any impact on the fit results, because the plateau is not attributed any physical meaning in terms of exchange.

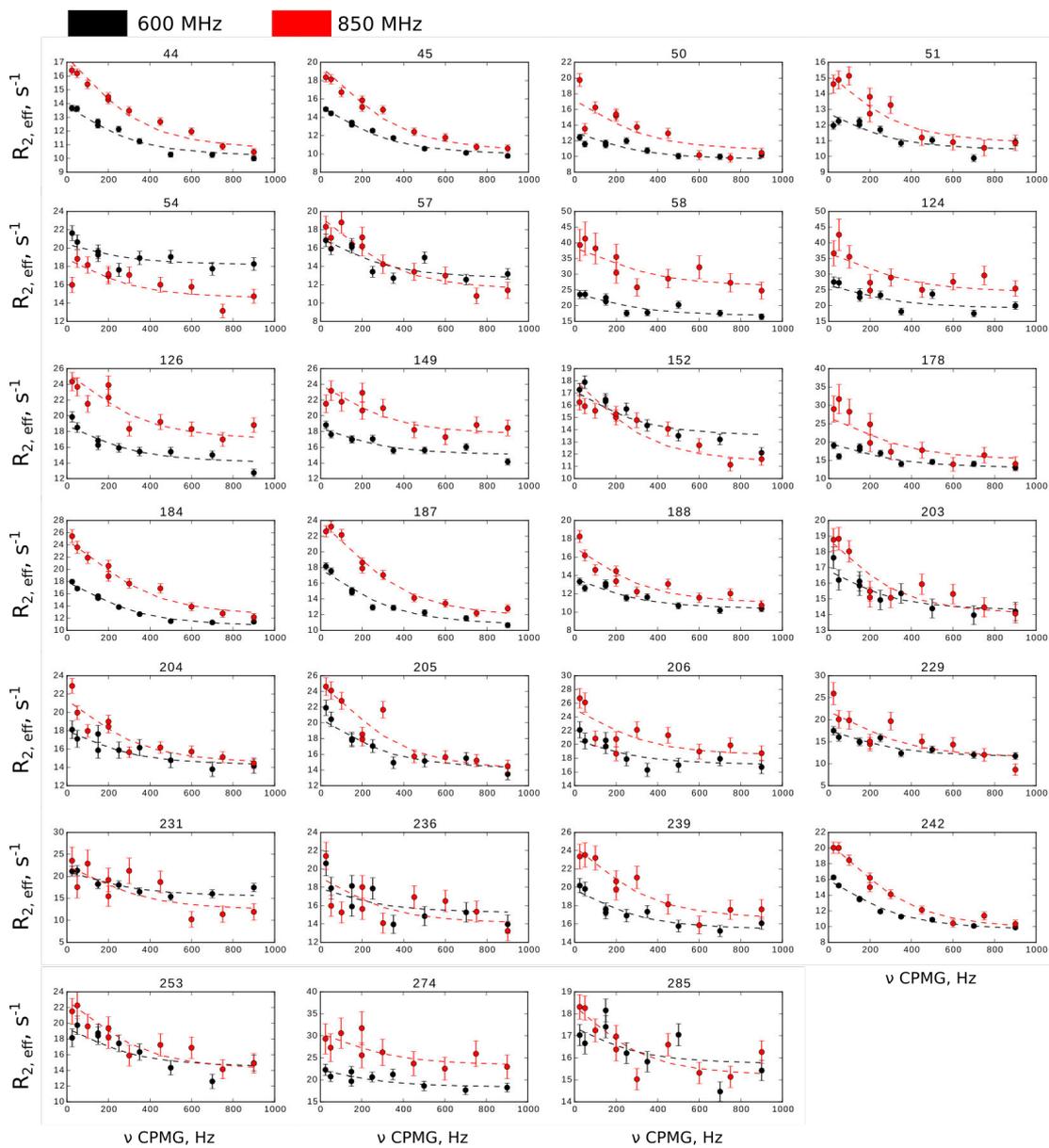


Figure 4.S7: Relaxation dispersion data for GGC1 in the presence of 20 mM GTP recorded on 600 (black dots) and 850 (red dots) MHz NMR spectrometers and the corresponding fits (dashed lines) to a global two-state model. Numbers above each plot indicate residue numbers of GGC1.

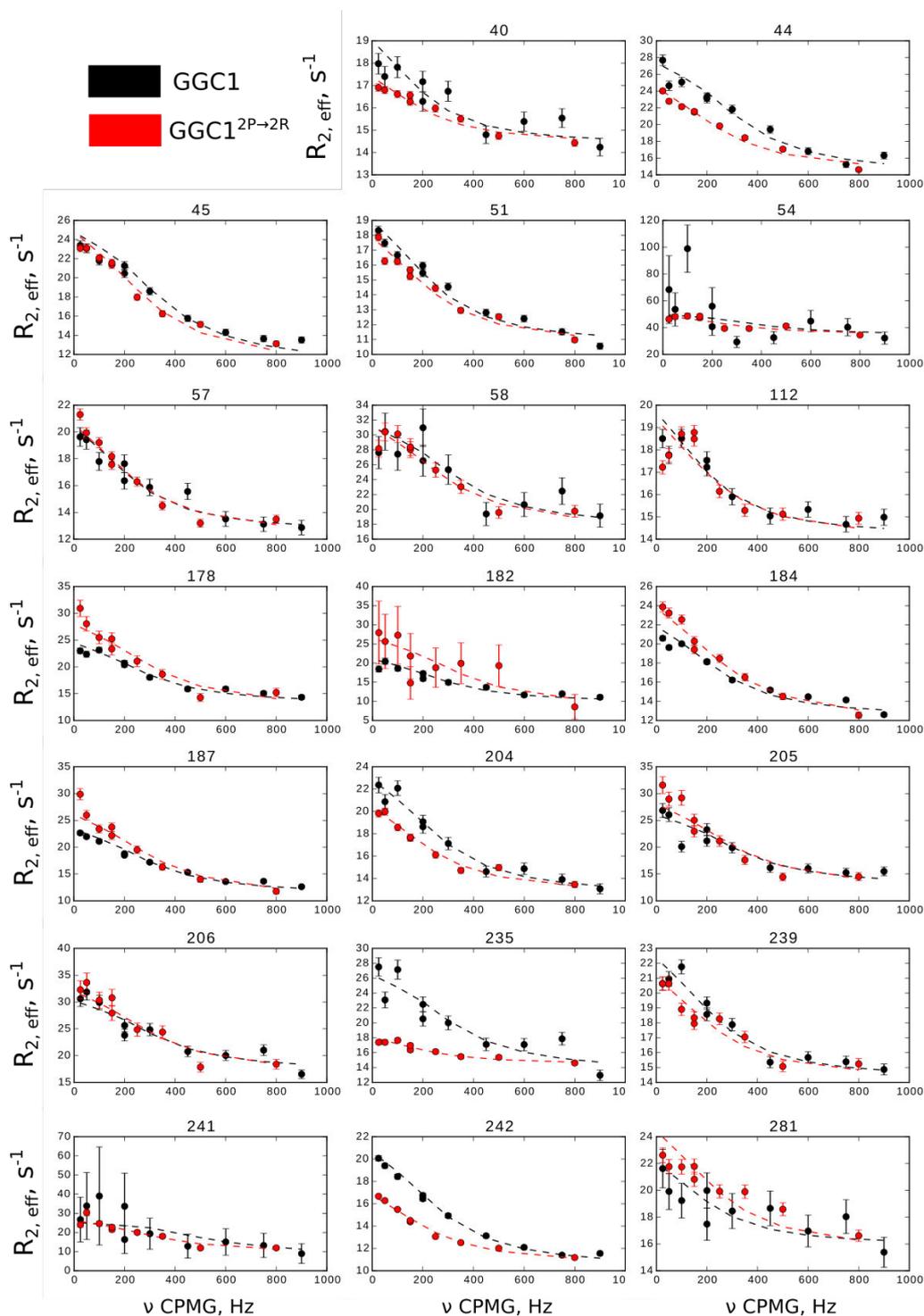
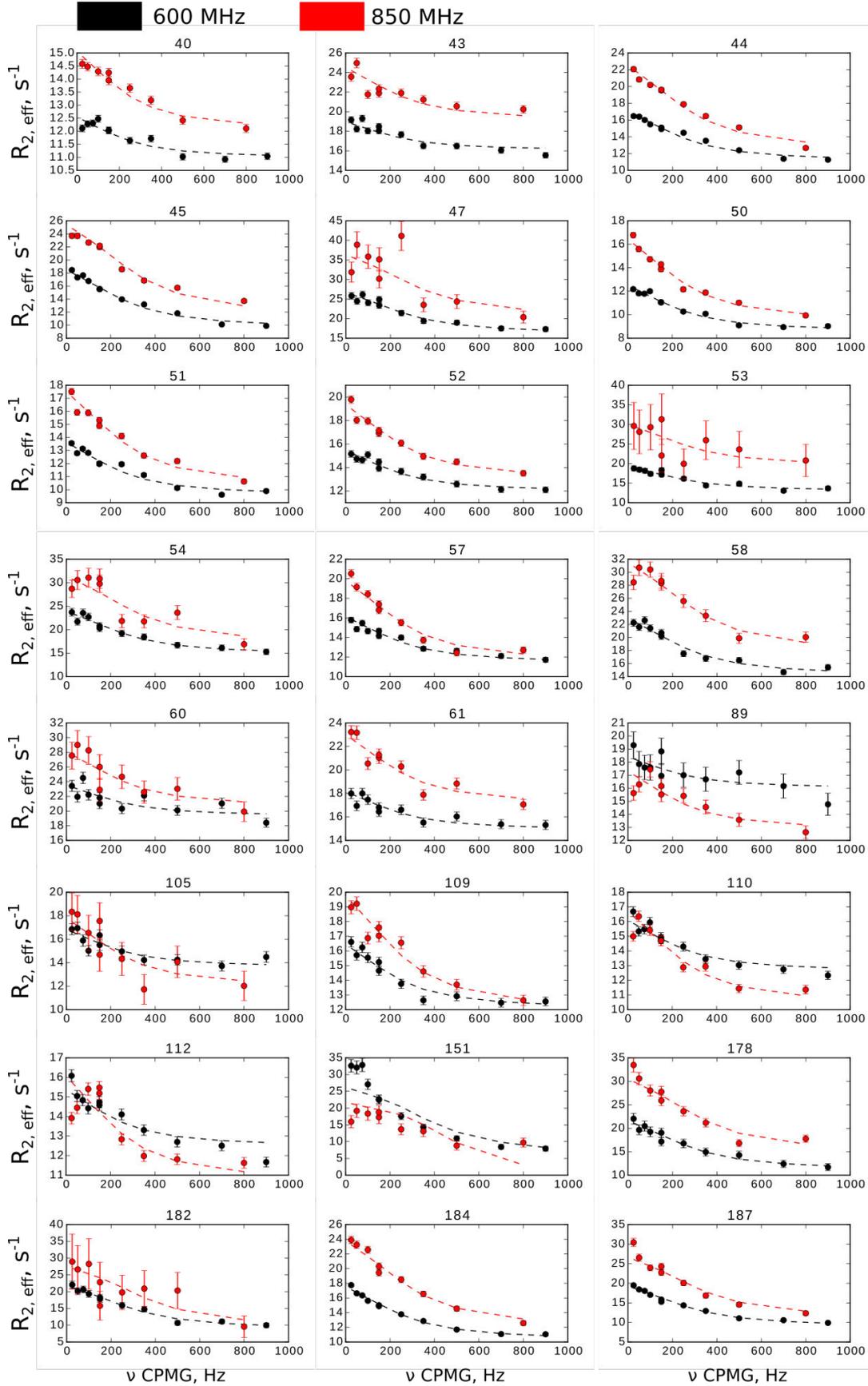


Figure 4.S8: Comparison of the dispersion profiles observed for wild-type GGC1 (black dots) and GGC1 Pro138,230Arg mutant (red dots) obtained on 850 MHz spectrometer. Numbers above each panel correspond to residue numbers in GGC1. To facilitate the comparison of the CPMG dispersion curves, an additive offset of  $R_{2,\text{eff}}$  (i.e. plateau  $R_{2,\text{eff}}$ ) was added to some residues. Such an offset does not have any impact on the fit results, because the plateau is not attributed any physical meaning in terms of exchange.



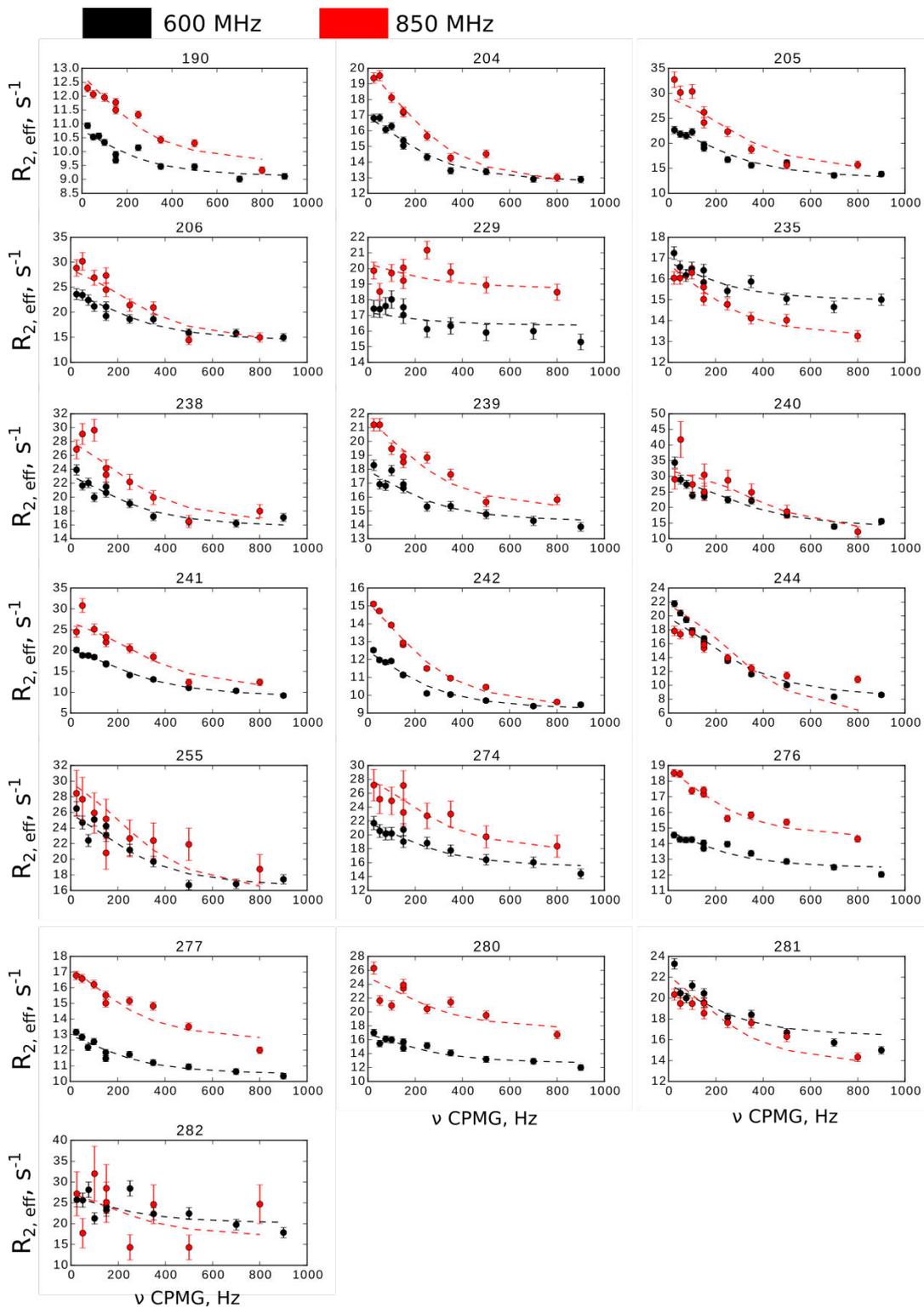


Figure 4.S9: Relaxation dispersion data for GGC1 Pro138,230Arg mutant on 600 (black dots) and 850 (red dots) MHz NMR spectrometers and the corresponding fits (dashed lines) to a global two-state model. Numbers above each panel indicate residue numbers of GGC1.

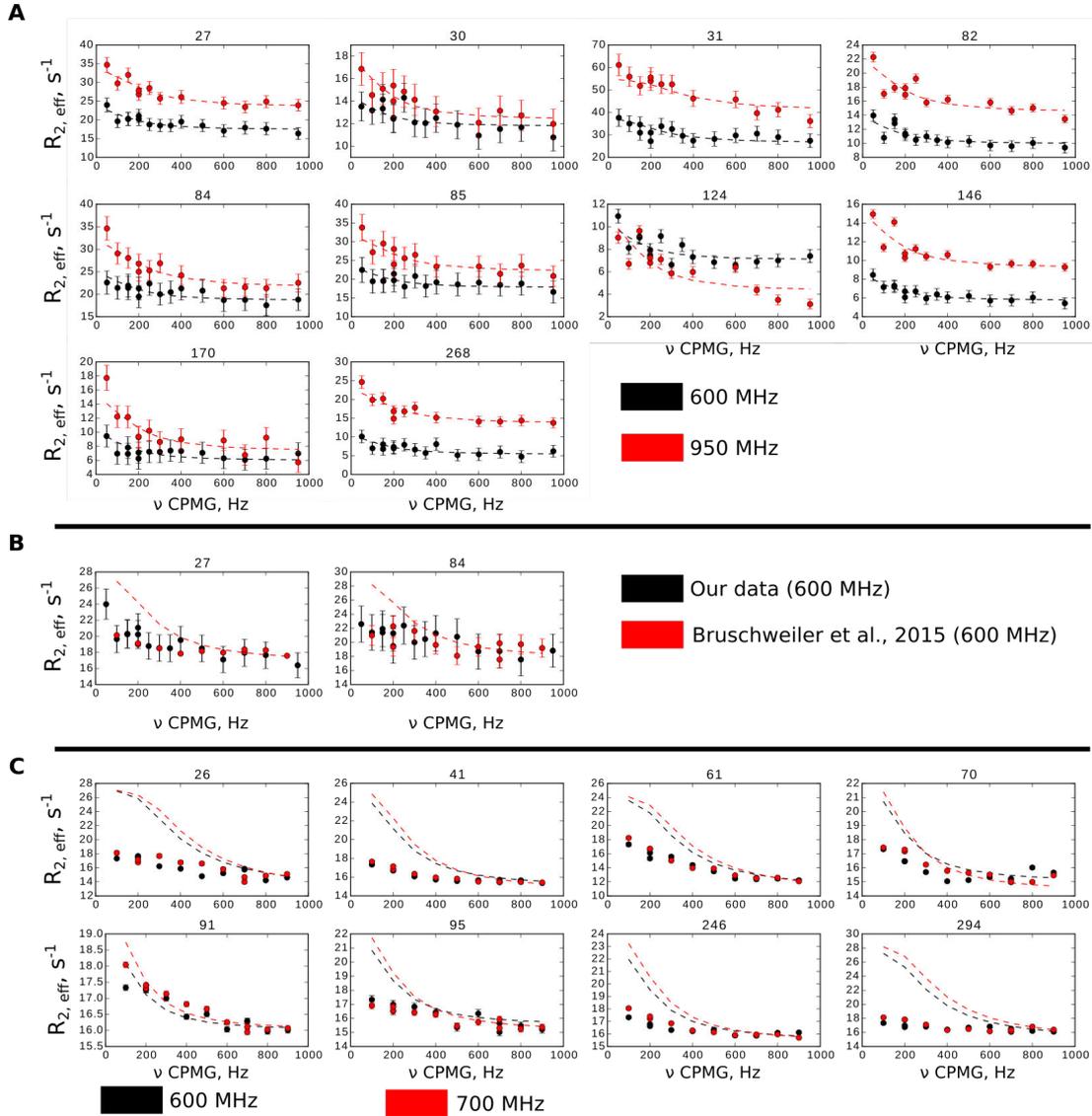


Figure 4.S10: Two field  $^{15}\text{N}$  CPMG relaxation dispersion data for AAC3<sup>c-salt-bridge</sup> mutant and comparison of our AAC3<sup>c-salt-bridge</sup> mutant CPMG data with the data of AAC3\* obtained by Bruschweiler et al, 2015. A, CPMG dispersion profiles of AAC3<sup>c-salt-bridge</sup> mutant on 600 (black dots) and 950 (red dots) MHz spectrometers. Dashed lines indicate the fits. B, comparison of our AAC3<sup>c-salt-bridge</sup> mutant data (black dots) with the AAC3\* data from Bruschweiler et al, 2015 (red dots) on 600 MHz spectrometer. The error bars for AAC3\* are arbitrary (as they were not to us with the data). Dashed curve indicates the expected dispersion magnitudes for published  $\Delta\omega$  values of 3.5 ppm (residue 27) and 3.7 ppm (residue 84). See supporting information of ref. (200). C, The data provided to us by S. Bruschweiler measured of AAC3\* on 600 (black dots) and 700 (red dots) MHz spectrometers. The dashed lines indicate  $R_{2,eff}$  back-calculated from their published fit data ( $k_{ex}=870\text{ s}^{-1}$  for TM residues (26, 41, 91, 95, 246 and 294) and  $k_{ex}=940\text{ s}^{-1}$  for residues in the matrix helices (61, 70),  $p_b=2\%$  and the  $\Delta\omega$  of 5.7 (residue 26), 3.1 (residue 41), 4.7 (residue 61), 2.2 (residue 70), 1.2 (residue 91), 2.1 (residue 95), 2.4 (residue 246) and 4.5 (residue 294) ppm. The disagreement of the measured data and 2-state model fit parameters published for that data is obvious.

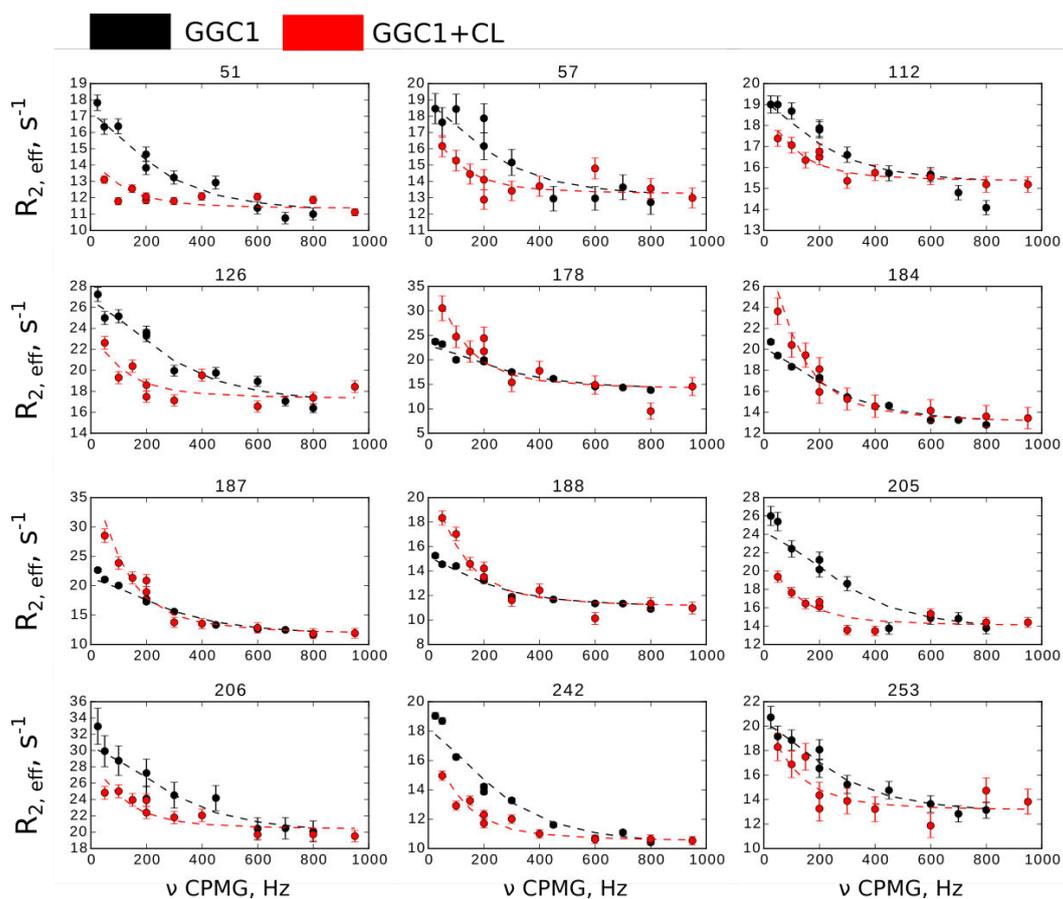


Figure 4.S11: Comparison of non-flat  $^{15}\text{N}$  CPMG relaxation dispersion profiles of apo GGC1 (red dots) and GGC1 in the presence of 6 mM cardiolipin (black dots) recorded on 700 MHz spectrometer. Numbers above each panel indicated residue number of GGC1. Closer inspection reveals different magnitudes and shapes of dispersion profiles, which likely suggests the reduced exchange rate of GGC1 in the presence of cardiolipin. To facilitate the comparison of the CPMG dispersion curves, an additive offset of  $R_{2,\text{eff}}$  (i.e. plateau  $R_{2,\text{eff}}$ ) was added to some residues. Such an offset does not have any impact on the fit results, because the plateau is not attributed any physical meaning in terms of exchange.

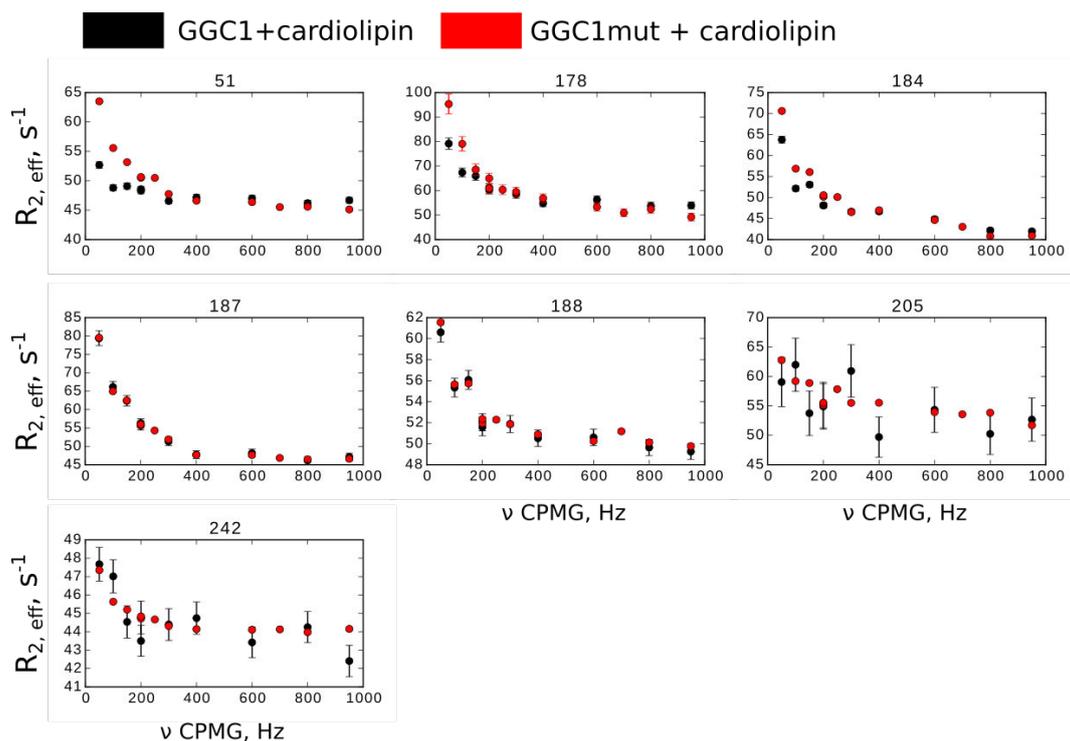


Figure 4.S12: Comparison of non-flat  $^{15}\text{N}$  CPMG relaxation dispersion profiles of wild-type GGC1 (black dots) and GGC1 $^{2\text{P}\rightarrow 2\text{R}}$  (red dots) in the presence of 6 mM cardiolipin recorded on 850 MHz spectrometer. Numbers above each panel indicated the residue number of GGC1. The dispersion profiles are nearly identical, indicating very similar millisecond dynamics experienced by the corresponding residues. Residue 51 is located in the mitochondrial helix. To facilitate the comparison of the CPMG dispersion curves, an additive offset of  $R_{2,\text{eff}}$  (i.e. plateau  $R_{2,\text{eff}}$ ) was added to some residues. Such an offset does not have any impact on the fit results, because the plateau is not attributed any physical meaning in terms of exchange.

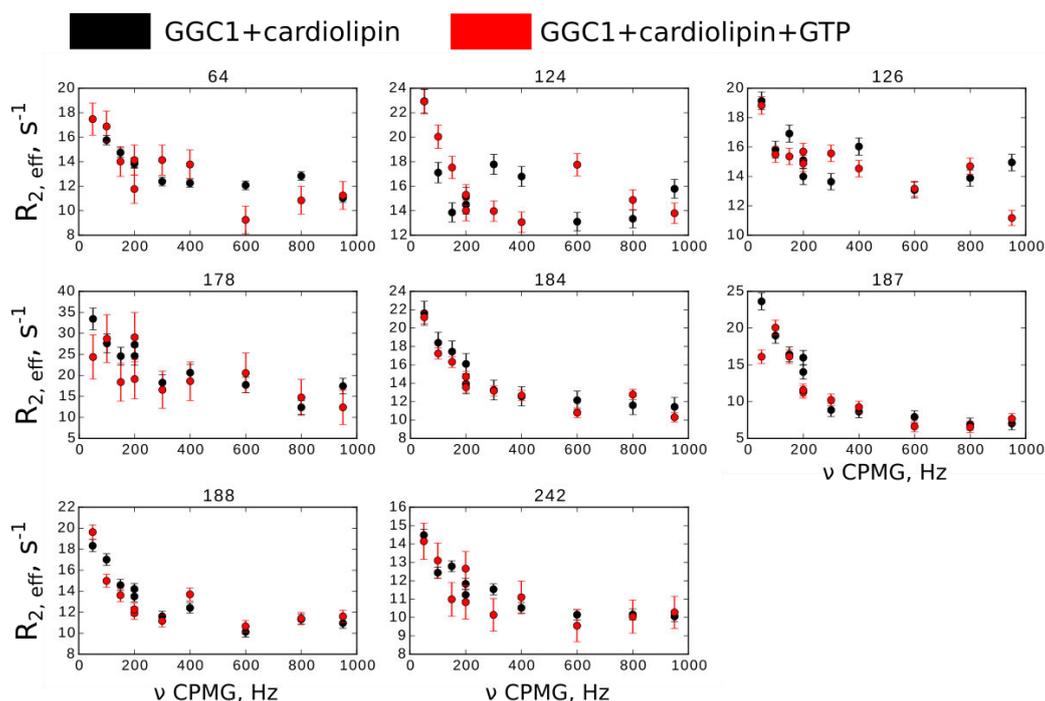


Figure 4.S13: Comparison of non-flat  $^{15}\text{N}$  CPMG relaxation dispersion profiles of apo GGC1 (black dots) and GGC1 with 20 mM GTP (red dots) in the presence of 6 mM cardiolipin recorded on 700 MHz spectrometer. The relaxation profiles are nearly identical, indicating similar millisecond dynamics. To facilitate the comparison of the CPMG dispersion curves, an additive offset of  $R_{2, \text{eff}}$  (i.e. plateau  $R_{2, \text{eff}}$ ) was added to some residues. Such an offset does not have any impact on the fit results, because the plateau is not attributed any physical meaning in terms of exchange.

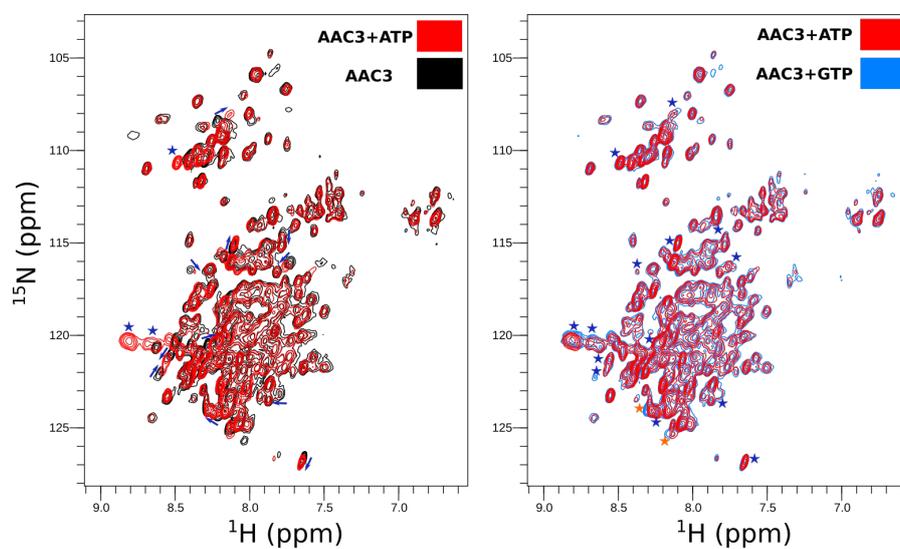


Figure 4.S14: Binding of AAC3<sup>c-salt-bridge</sup> mutant to GTP and ATP. Left - overlay of 2D BT-HSQC spectra representing the apo AAC3<sup>c-salt-bridge</sup> mutant (black) and AAC3<sup>c-salt-bridge</sup> mutant in the presence of 20 mM ATP (red). More pronounced perturbations are indicated by blue arrows. Right - Overlay of AAC3<sup>c-salt-bridge</sup> mutant spectra in the presence of 20 mM ATP (red) or 20 mM GTP (blue). Blue stars indicate peaks displaying more pronounced chemical shift perturbations (as in the spectra on the left) for both ATP and GTP. Orange stars indicate differences. Almost completely identical overlay of the spectra indicates, that the effect sensed by AAC3<sup>c-salt-bridge</sup> mutant is identical both in the presence of GTP and ATP.

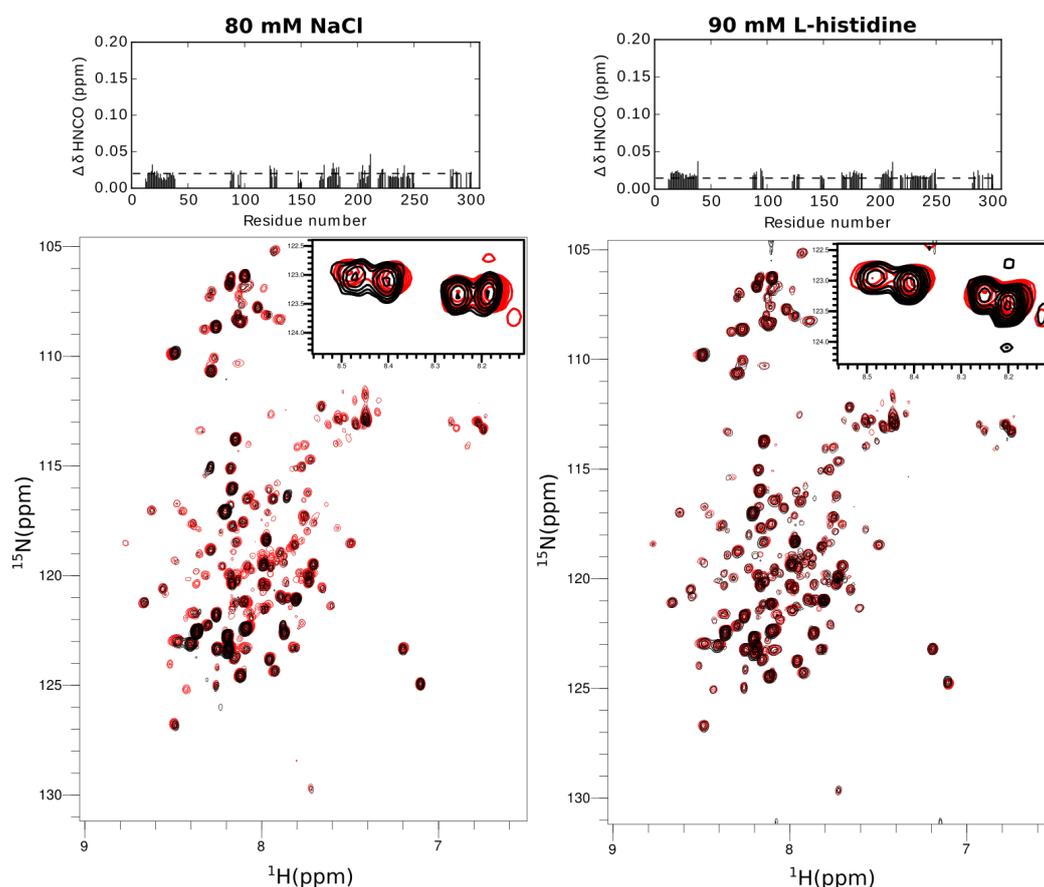


Figure 4.S15: Data of yeast ORC1 in the presence (black) and absence (red) of 80 mM NaCl (left) and 90 mM L-histidine. Top: residue-wise chemical shift perturbations extracted from the HNCO spectra below, normalized to the chemical shift perturbation of  $^1\text{H}$ . Horizontal dashed line represents 3 standard deviations extracted by analysing all chemical shift perturbations for assigned residues. The scale of the ordinate in the bar-plot is kept the same as in the previous examples for GGC1 binding, to allow easier visual comparison of the binding data for the two proteins. Bottom: overlay of BT-HNCO spectra in the presence and absence of aforementioned solutes. Insert indicates the same residues as in the figure above, which show small perturbations in the presence of charged solutes. The decrease of signal intensity after addition of NaCl might be attributed to the decreased sensitivity of cryogenic probe at higher salt concentrations.

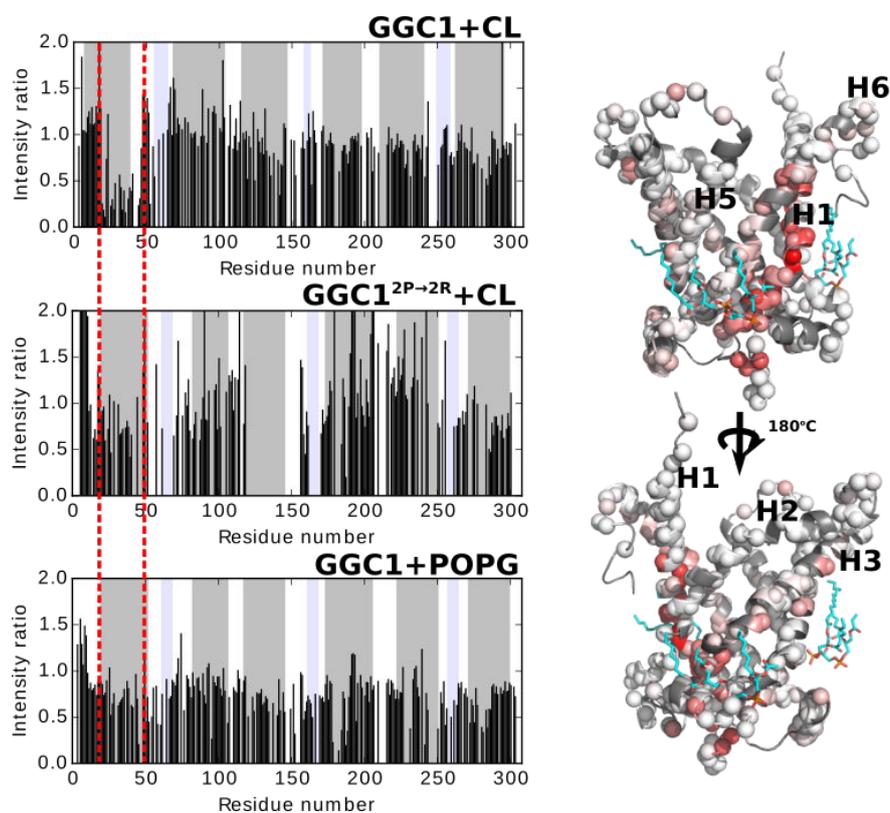


Figure 4.S16: Peak intensity changes upon addition of different lipids to GGC1. Left - residue wise-peak intensity changes for wild-type GGC1 in the presence of 6 mM cardiolipin (top), GGC1<sup>2P→2R</sup> in the presence of 6 mM cardiolipin (middle), and GGC1 in the presence of 6 mM POPG extracted from 3D HNCO spectra. Gray and blue shaded background indicates TM and matrix helices, respectively. Red dashed lines outlines the region, where a strong decrease of peak intensities are observed for wild-type GGC1 in the presence of cardiolipin, but not in other cases. Right - intensity changes of GGC1 in the presence of 6 mM cardiolipin plotted on a homology model of GGC1. Red spheres represent strong intensity changes, white the white spheres correspond to the residues which experience increased or unchanged intensities. Cyan molecules are cardiolipin molecules visible in the X-ray structure of bovine heart AAC1 (PDB:1OKC). Their positions relative to GGC1 structural model were determined after the structure-based superimposition of the GGC1 model and 1OKC.

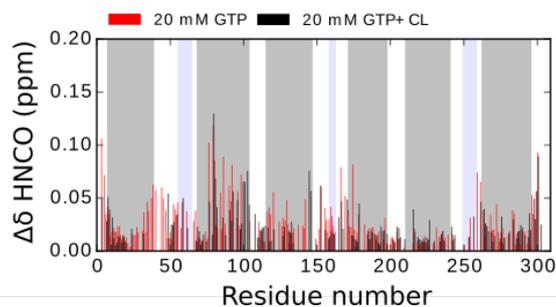


Figure 4.S17: Bar-plot indicating chemical shift perturbations upon addition of 20 mM GTP to GGC1 (red) or GGC1 with 6 mM cardiolipin. Chemical shift perturbations are scaled by the gyromagnetic ratios. Gray and blue shaded background represent TM and matrix helices, respectively. No major changes in GTP binding are observed for GGC1 in the presence of cardiolipin.

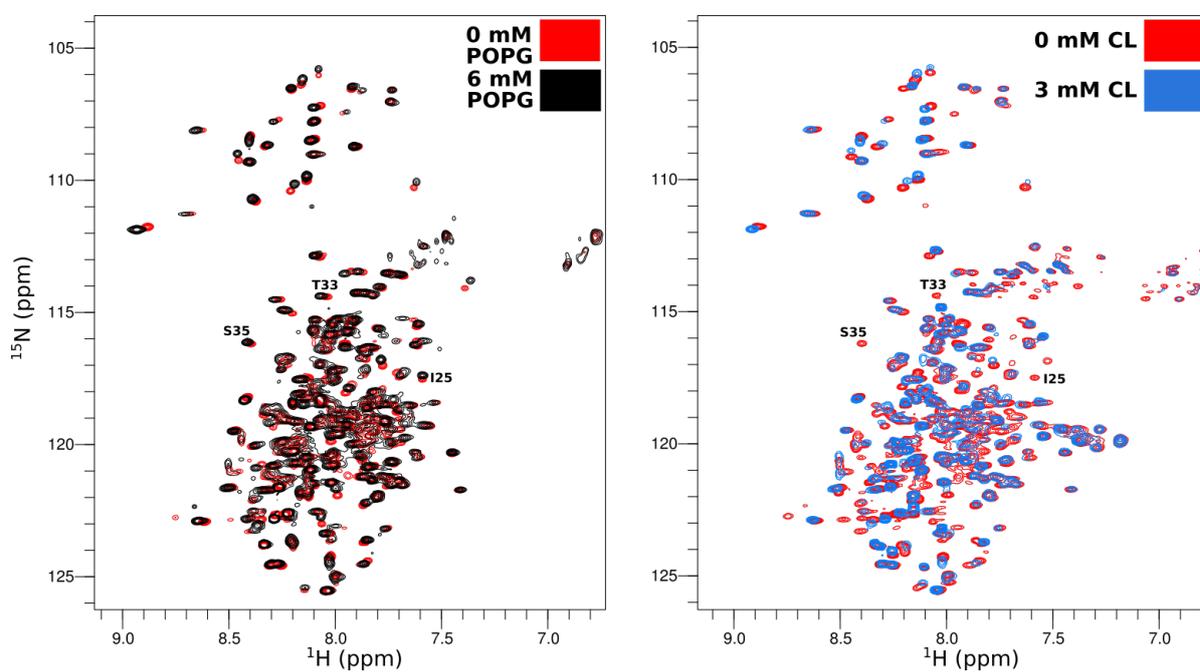


Figure 4.S18: 2D HSQC spectra of wild-type GGC1 in the absence and presence of 6 mM POPG (black) or 3 mM cardiolipin (blue). Similar directions and magnitudes of chemical shift perturbations in both cases are evident. Some residues disappearing in the presence of cardiolipin but still present in samples with POPG are indicated.

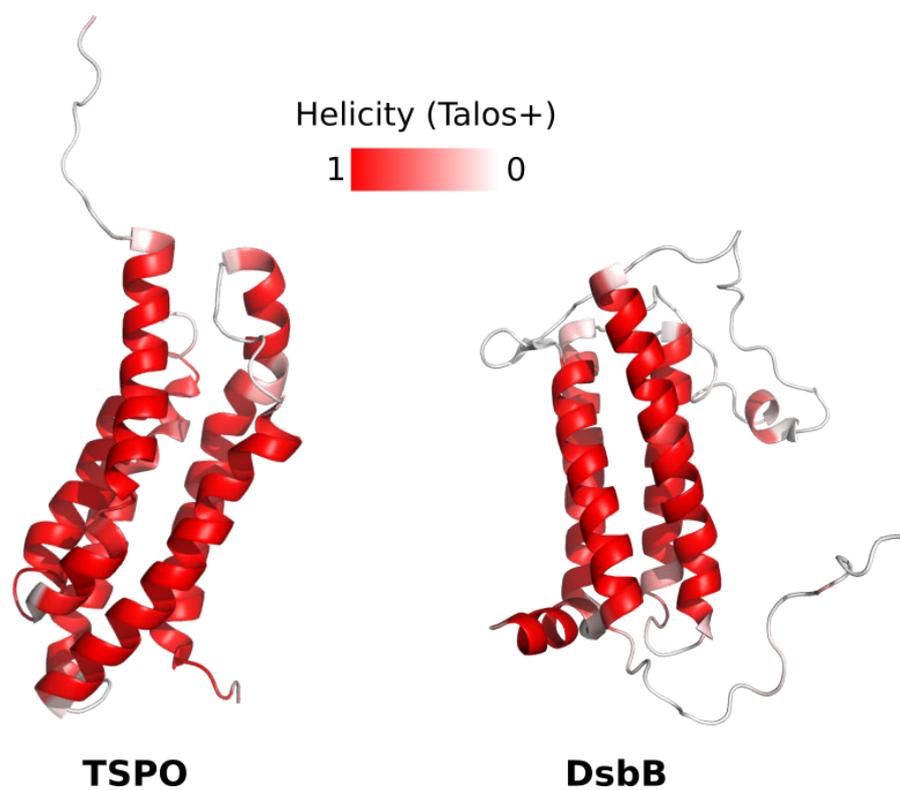


Figure 4.S19: Helicity as predicted from chemical shifts by Talos+ software plotted on the NMR structures of TSPO (left) and DsbB (right) solved in DPC. Talos+ secondary structure predictions are generally used for soluble proteins, so we wanted to establish with these calculations that Talos+ predicts secondary structures correctly for membrane proteins in DPC micelles. Here we show that there is an extremely good agreement between the predictions and observed secondary structures. Regions which are predicted to be helical are colored in red, while loops are colored in white.

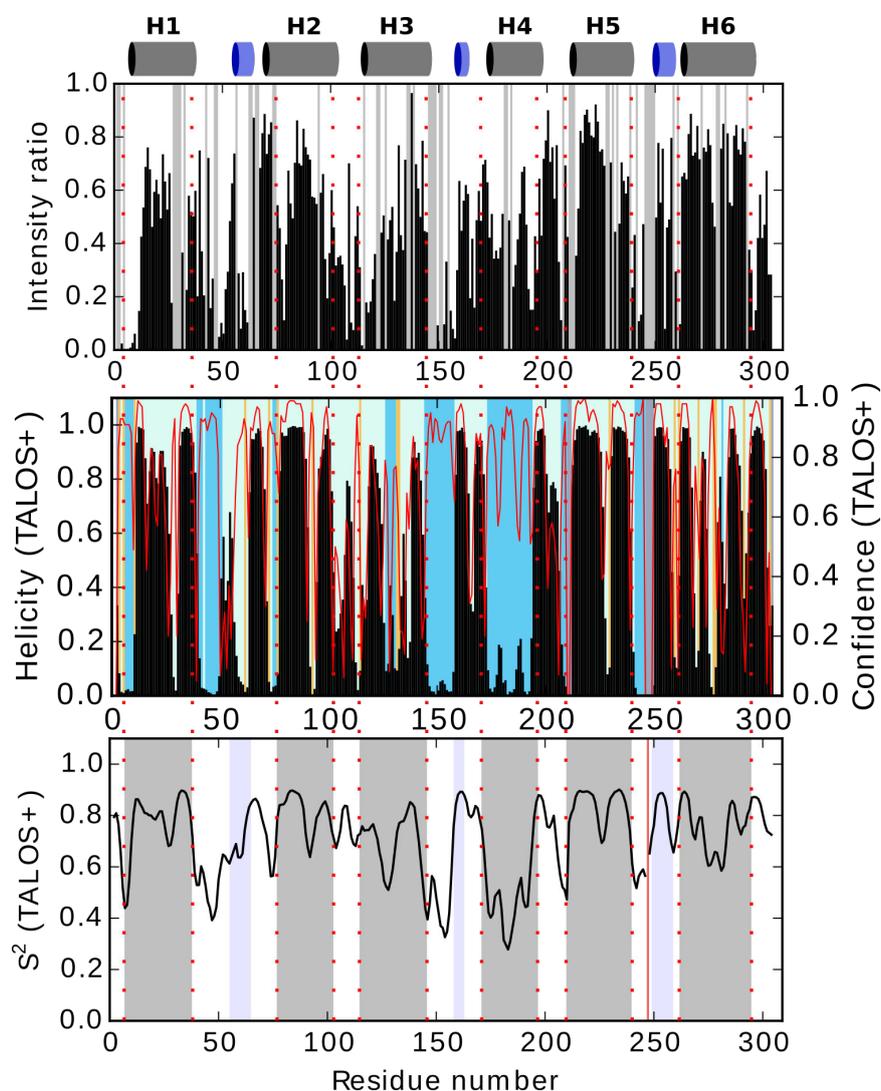


Figure 4.S20: Paramagnetic relaxation enhancement (A), Talos+ helicity prediction (B) and Talos+ order parameter prediction (C) data for GGC1. Gray and blue cylinders indicate TM and matrix helices respectively. A, ratio of intensities after and before addition of 1.5 mM PRE agent gadodiamide. Gray background indicated regions where the data is not available. B, Talos+ helicity prediction. Black bars indicate the helicity. Yellow background indicates regions where the prediction is ambiguous. Light green regions indicates unambiguous predictions. Blue regions indicate dynamic, and therefore unstructured regions. Gray background indicates regions where prediction is not available. Red curve represents the confidence of prediction, which is indicated by the ordinate on the right. C, Talos+ predicted order parameter. Gray and blue shades in the background indicates TM and matrix helices, respectively. Red vertical line indicates the absence of prediction. Red dots along the plots delimit the regions corresponding to the secondary structures as extracted from the homology model. As can be seen, there is a good agreement between paramagnetic relaxation enhancement data and the predicted parameters.

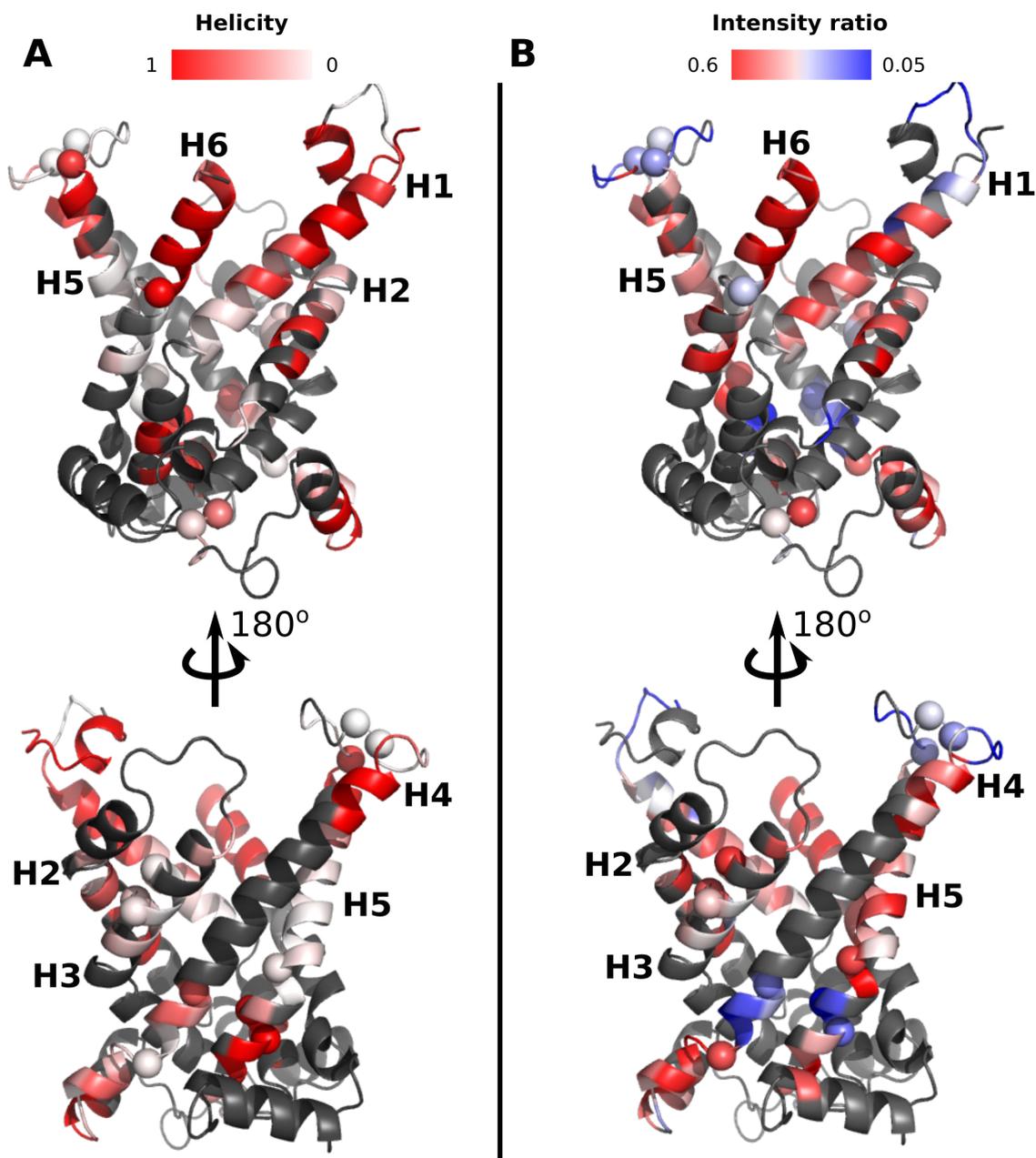


Figure 4.S21: Talos+ helicity prediction (A) and paramagnetic relaxation enhancement (B) data plotted on a structural model of yeast ORC1. A, Talos+ helicity prediction. Red regions indicate high helicity prediction, while the white regions indicate low prediction scores. For gray regions the helicity prediction was not available due to the lacking assignment. Two opposite orientations of the model are shown at the top and the bottom. B, PRE data calculated as a ratio of peak intensities after and before the addition of 3 mM gadodiamide. The horizontal color-bar indicates regions strongly affected by PRE agent (blue) and more weakly affected regions (red). Spheres represent the residues which display significantly non-flat CPMG relaxation dispersion profiles. It can be seen that some TM regions in ORC1 are affected by the paramagnetic agent, however the agreement between secondary structure predictions and PRE data is poor.

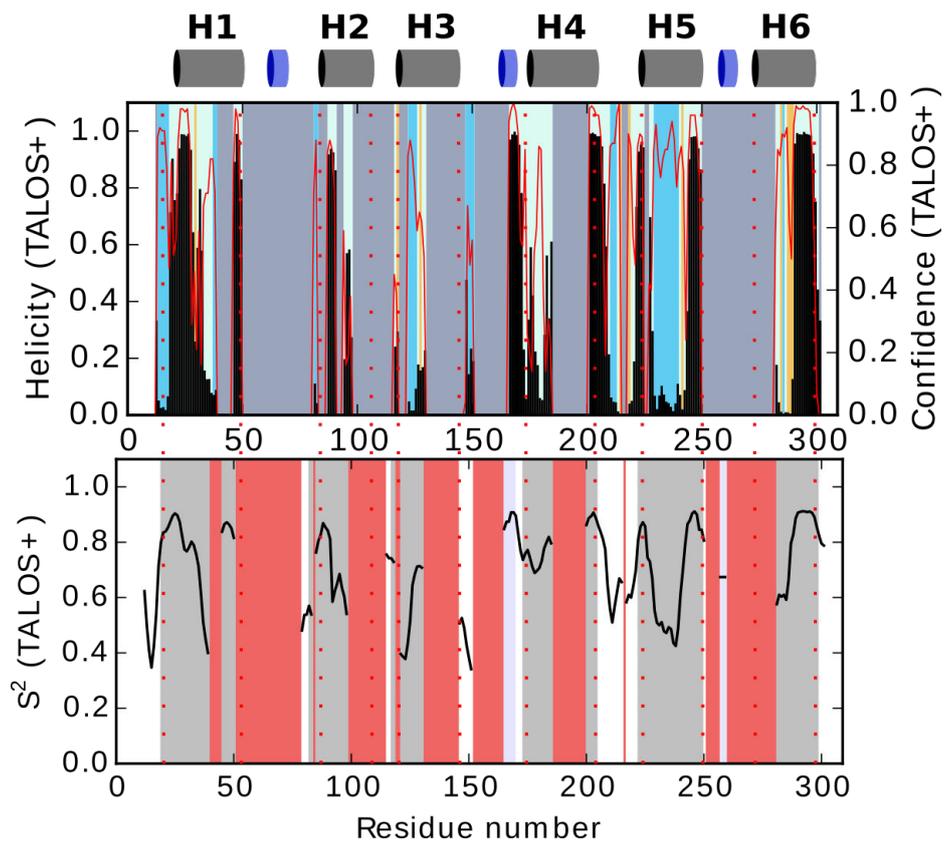


Figure 4.S22: Talos+ helicity prediction data and order parameter prediction data for yeast ORC1 (bottom). Helicity prediction - yellow background indicates ambiguous prediction, light green background - good prediction, blue background - dynamic and thus unstructured regions. Gray regions correspond to the lack of prediction due to the missing assignment. Red curve represents the confidence of prediction. Order parameter prediction - gray and blue regions correspond to TM and matrix helices, respectively. Red regions correspond to regions where order parameter was not predicted due to the missing assignment. Red dots across two bar-plots and the helix symbols on top of the plot delineate the limits of helices estimated from the ORC1 model.

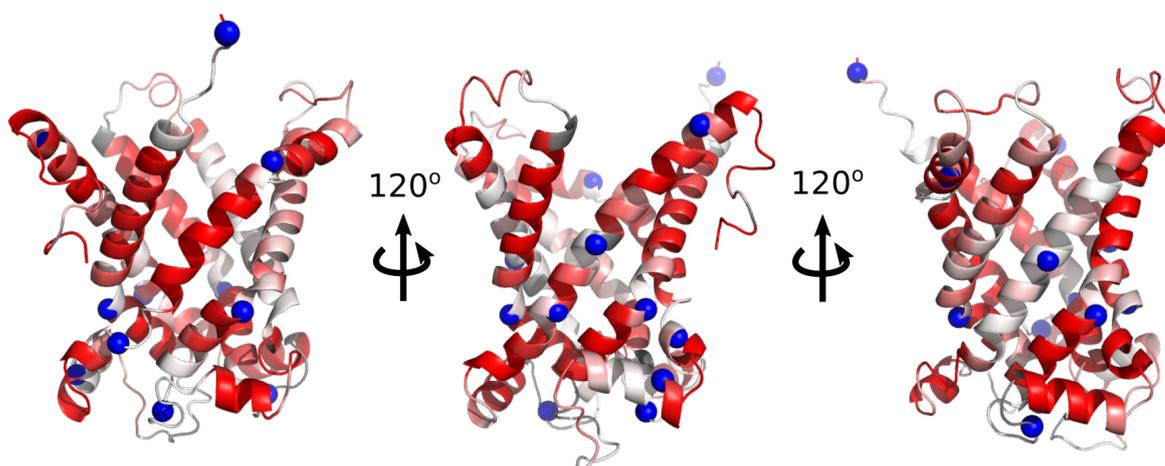


Figure 4.S23: The same GGC1 helicity plots as in Figure 4.25, but GGC1 is shown here in three orientations. Blue spheres represent the locations of Pro residues. Note, that in TM segments low helicity values are associated with the presence of Pro residue

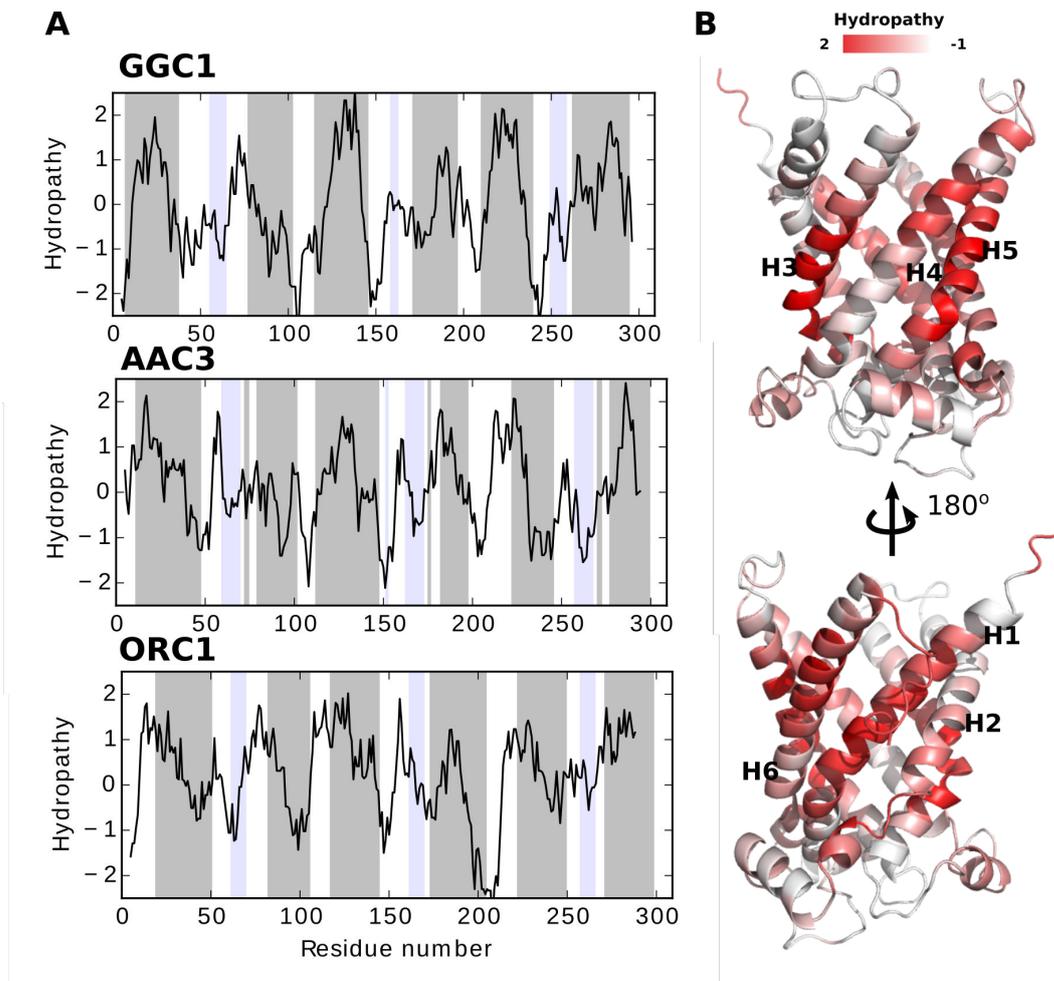


Figure 4.S24: A, hydropathy profiles of the three mitochondrial carriers. Gray shaded background represents TM helices, blue background corresponds to matrix helices. B, hydropathy plot on a homology model of GGC1. H2 and H4 are less hydrophobic as compared to the other TM helices in GGC1.

Table 4.S1:  $|\Delta\omega|$  values for the AAC3 (and AAC3<sup>c-salt-bridge</sup>) obtained from Chemex

Residue number	AAC3		AAC3 <sup>c-salt-bridge</sup>	
	$ \Delta\omega $	Error	$ \Delta\omega $	Error
26	2.25	0.37	n.a	n.a
27	1.72	0.28	2.6	0.38
41	1.65	0.25	n.a	n.a
61	3.05	0.62	n.a	n.a.
70	1.73	0.27	n.a	n.a.
84	2.69	0.67	2.63	0.51
91	1.51	0.23	n.a	n.a.
95	1.50	0.29	n.a	n.a.
246	1.53	0.23	n.a	n.a.
294	1.48	0.22	n.a	n.a.

Table 4.S2: Comparison of CPMG fit parameters obtained from jackknife analysis ( $\approx 20$  residues) and a global single fit of  $\approx 40$  residues

Protein	Jackknife		Global	
	$k_{ex}, s^{-1}$	Population, %	$k_{ex}, s^{-1}$	Population, %
GGC1	1157 $\pm$ 104	1.91 $\pm$ 0.12	1280 $\pm$ 51.4	1.67 $\pm$ 0.064
GGC1+GTP	1387 $\pm$ 118	3.33 $\pm$ 0.94	1560 $\pm$ 90.7	2.60 $\pm$ 0.55
GGC1 <sup>2P<math>\rightarrow</math>2R</sup>	1486 $\pm$ 84	3.4 $\pm$ 0.9	1492 $\pm$ 48.5	2.18 $\pm$ 0.2



## 5 Discussion

We have found in this study that mitochondrial carriers in DPC micelles display many features which are not compatible with their biological function. To summarize our findings:

- three relatively distinct mitochondrial carriers display similar millisecond dynamics. These dynamics occur on a similar time-scale as the expected interconversion of the carrier between the m- and the c-states. Mutations designed to abolish the interconversion (which indeed inactivate the carriers in liposome transport assays), however, do not affect the dynamics. This suggests that these two processes are not related.
- the three carriers we have studied are either unable to bind their dedicated substrates (ORC1) or do not show preference for their natural substrates (AAC3, GGC1) or ligands like cardiolipin, nor very strong and selective inhibitors (CATR).
- the carriers do not have a compact tertiary fold.
- the carriers are perturbed on a secondary structure level. These perturbations likely occur due to the solvent exposure of TM segments.

It is therefore likely that most, if not all mitochondrial carriers are prone to structural perturbations in DPC micelles. However, several NMR studies have been published on these carriers in DPC in high impact factor journals, linking the behaviour of the carriers in DPC to some of the biologically relevant phenomena. We would like to look at these studies here and reinterpret their findings in the light of our observations.

Two of the published NMR studies deal with AAC3 (*200*) and GGC1 (*228*) carriers. We have already extensively discussed the biological relevance of AAC3 and GGC1 in DPC micelles, and suggested that our observations are not in agreement with functional state of these two proteins in DPC.

We will here analyze the literature published on UCP2 and SCaMC1 and will try to put the published data into the light of our findings.

### 5.1 Structure of UCP2 in DPC micelles

A structure of the uncoupling protein 2 in 150 mM DPC in the presence of 5 mM GDP was solved by NMR using a molecular fragment replacement approach guided by chemical shifts and RDC together with constraints derived from PRE (*141*). Initial observation of the structure reveals similarity to the crystal structures of AAC: 6 transmembrane  $\alpha$ -helices surround an internal cavity, and three matrix helices were observed to connect two TM helices in the same domain. Pro-kinks in odd helices are also present in UCP2 structure, although these kinks are much smaller in helices 3 and 5 as compared to the crystal structures of AAC. A closer examination of the NMR structure, however, reveals several features which are not compatible with the function of the carrier. These observations prompted a more detailed investigations

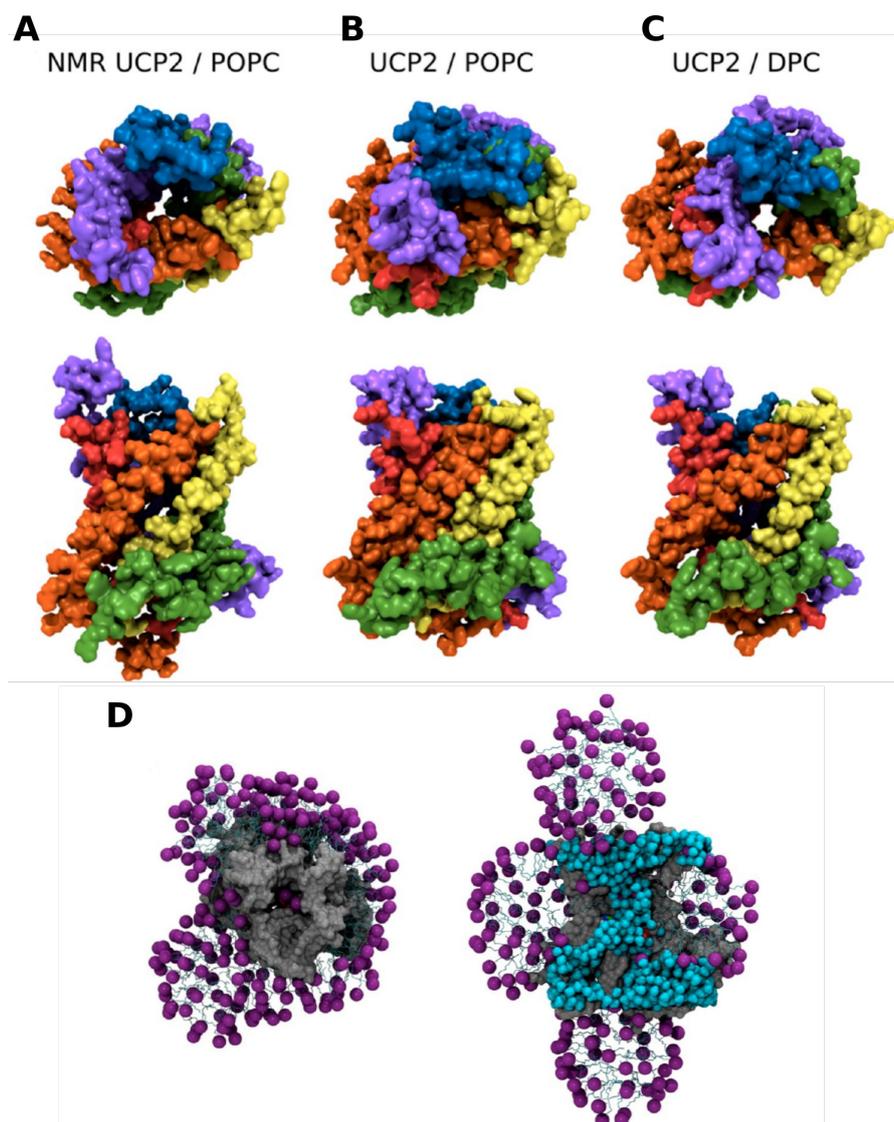


Figure 5.1: A,B and C: Adaptations of UCP2 NMR structure in DPC micelles to different environments, as determined by molecular dynamics simulations. Bottom and side views. A, an original structure of UCP2, displaying large packing defects in the TM region. B, UCP2 structure after relaxing inside POPC bilayer. It can be seen that the packing defects have largely disappeared. C, UCP2 structure after relaxing in DPC micelles. The packing defects persist and the structure is more similar to the original NMR structure. D, detergent binding to UCP2 structure, bottom and side views. DPC head-groups are illustrated in purple. Asymmetric distribution of detergent molecules is observed. In addition, small micelles are also formed on the cytosolic and the matrix sides of the carriers. Blue spheres illustrate water molecules. Figure taken from (172).

by molecular dynamics simulations (172).

Among the most striking features observed in the UCP2 structure is an  $\approx 6\text{\AA}$  wide channel, continuing from the cytoplasmic to the matrix side of the protein (Figure 5.1D, right). This feature of UCP2 is in a stark disagreement with the CATR inhibited crystal structures of AAC, where the protein cavity is fully closed on the matrix

side. Such a large channel in a mitochondrial carrier is unlikely, as it would allow uncontrolled flow of protons (and dissipation of the proton gradient across the inner mitochondrial membrane). In fact, the width of the channel observed in UCP2 is comparable to the pore-forming proteins, such as  $\alpha$ -hemolysin.

Another striking feature observed in the NMR structure of UCP2 is a very poor packing of the TM helices (Figure 5.1A), with large cavities present between TM helices. We have already discussed in the introduction (Section 1.3.3) the importance of the packing and van der Waals interactions in determining the stability of membrane proteins.

MD simulations were used by Zoonens et al to investigate the behaviour of UCP2 in several different scenarios: 1) UCP2 structure obtained from the NMR experiments was placed inside POPC bilayer, a mimic of natural environment of the protein. A biologically relevant structure should correspond to an energetically most stable conformation in a native environment. 2) UCP2 structure was also investigated in the presence of DPC micelles. In addition, the effect of GDP<sup>3-</sup> was investigated in both cases.

MD simulations have revealed that when the NMR structure of UCP2 is allowed to relax to the most energetically favourable conformation in POPC bilayer, the protein undergoes rapid contraction and becomes much more compact, deviating largely from the original NMR structure (RSMD $\approx$ 6Å for the backbone atoms) (see Figure 5.1B). Given the presence of large holes between TM helices, this behaviour is expected.

In contrast, the poorly packed state of UCP2 appears to be stabilized in the DPC environment. In the course of MD simulation, detergent molecules intercalate between the TM helices and compete for the protein-protein interactions (see Figure 5.1C). In addition, MD simulations revealed that the binding of DPC molecules to UCP2 structure is far from an expected detergent belt around the hydrophobic parts of the protein: both head-groups and alkyl chains were shown to interact with UCP2. Even more interestingly, small micelles were observed to form on both the cytoplasmic and matrix side of the protein (Figure 5.1D).

In addition, in both DPC and POPC environments, GDP<sup>3-</sup> binding could be detected in the simulations. However, this inhibitor of the UCP2 did not appear to have a strong impact on the protein in the two environments investigated.

It therefore appears that the structure of UCP2 in detergent micelles does not represent a biologically relevant form of the protein. We do not know, however, if the observed perturbations are introduced by the detergent, or result from limited amount of data available for the structure calculations. However, in the light of our observations on mitochondrial carriers, it is very likely that these proteins do present large packing defects, which in turn allow them to be flexible. Using relaxation dispersion experiments we identified that several TM helices of these proteins displayed millisecond dynamics. Furthermore, at least some TM helices of mitochondrial carriers displayed nanosecond dynamics, as could be observed from R<sub>2</sub> measurements. However, it is possible that relatively short time-scales of the simulations (<200 ns)

did not allow to detect these motions.

In addition to the molecular dynamics simulations, Zoonens et al also investigated the activity of UCP2. They found UCP2 to be inactive when reconstituted from a DPC solubilized state. This was interpreted such that DPC irreversibly inactivates mitochondrial carriers.

Dr. Loredana Capobianco and her team have performed similar experiments on AAC3. In these assays the activity of AAC3 was still present, albeit reduced by  $\approx 15\%$  as compared to AAC3 reconstituted from sarkosyl. However, our collaborators have noted that detergent Triton X-114 was necessary during the proteoliposome reconstitution, where DPC solubilized AAC3 would be inserted into liposomes. Furthermore, Audrey Hessel from our group has discovered that DPC is a very poor detergent to destabilize liposomes<sup>35</sup>. The lack of activity of UCP2 solubilized in DPC might have therefore resulted from the fact that DPC is not compatible with protein reconstitution in liposomes, and not because it inactivates mitochondrial carriers irreversibly.

## 5.2 Mechanism of proton transport by UCP2

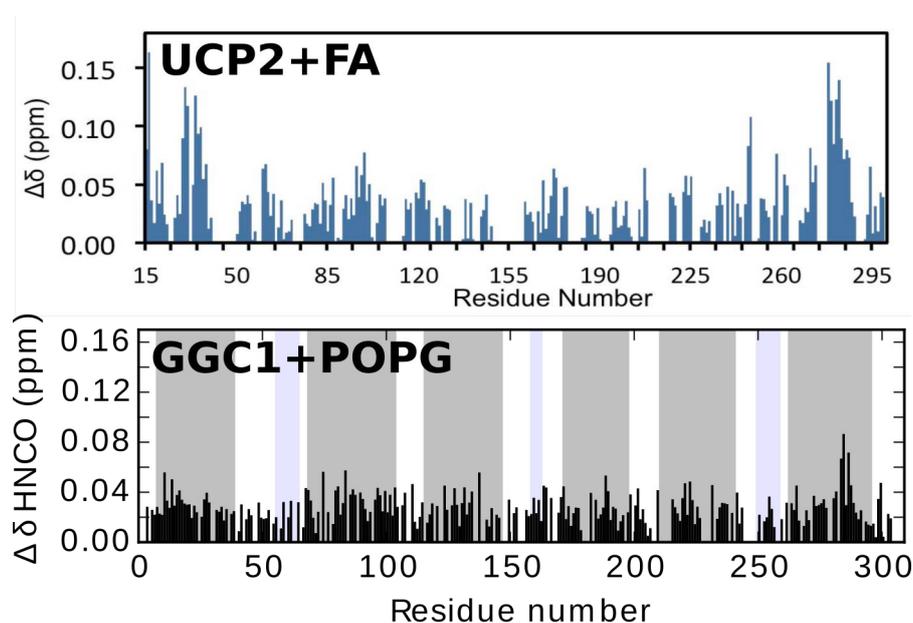


Figure 5.2: Top, chemical shift perturbations induced to UCP2 upon addition of 20 mM fatty acid. Wide-spread perturbations are visible, with regions corresponding to helix 1 and helix 6 showing a stronger effect. Figure taken from Berardi and Chou. Bottom, chemical shift perturbations induced to GGC1 upon addition of 6 mM POPG. Similar profile can be seen: wide-spread chemical shift perturbations, with helix 6 showing a stronger effect.

Berardi and Chou have further published a study on UCP2, from which they concluded on the possible proton transport mechanism by this protein. The aforementioned NMR structure was used to get functional insights.

<sup>35</sup>Proteoliposomes need to be destabilized prior to protein insertion in these protocols.

The large part of the paper is devoted to the investigation of UCP2 binding to palmitic acid molecules. Upon addition of 20 mM of palmitic acid, widespread HNCO CSPs of moderate magnitudes ( $\approx 0.3$ - $0.5$  ppm) were reported, which can be seen in Figure 5.2. These disperse perturbations could be caused by the change in the UCP2 environment when hydrophobic alkyl chains of poorly soluble palmitic acid partitions from the aqueous phase into DPC micelles. The peaks corresponding to the residues in helices 1 and 6 displayed stronger CSPs, some of them above 0.1 ppm.

The saturation of the binding was also investigated when palmitic acid was titrated by incrementing concentrations from 50  $\mu$ M to 20 mM. The authors claim that they observe an obvious binding saturation. This, however, does not become clear from a rigorous inspection of the experimental data presented (FIGURE).

The authors have also investigated the binding of another negatively charged ligand, ASO<sup>-</sup> (alkylsulfonic acid, an analog of fatty acids). This molecule induced nearly identical chemical shift perturbations as compared to palmitic acid. On one hand it could thus suggest that UCP2 is not able to distinguish between these two molecules that differ in their polar group (carboxyl versus sulfonyl)<sup>36</sup>. On the other hand, we do not know if UCP2 is able to distinguish between these two molecules also in its native environment, or only when reconstituted in DPC micelles.

These observations are interesting, because we have noticed a similar behaviour of GGC1 while titrating it with the negatively charged lipids, cardiolipin or POPG: 1) addition of lipids introduced widespread CSPs in GGC1 spectra; 2) we could not detect a clear saturation of the binding at higher concentrations of the lipids; 3) we could observe larger CSPs for the residues in helix 6 of GGC1 in the presence of POPG (see Figure 5.2).

When these observations are taken into consideration, we cannot dismiss the possibility that UCP2 binding to the fatty acids observed by NMR in DPC micelles does not reflect the biologically relevant binding. The binding pattern could be explained taking the amphiphatic nature of a fatty acid into consideration: the alkyl chain is expected to display a preference to partition inside the lipophilic phase of the detergent. The negatively charged group, however, would have a preference to bind inside the positively charged cavity. As the helices 1 and 6 are not connected to each other, the fatty acid molecule could insert into the opening.

The biological relevance of the UCP2 binding to the fatty acids was investigated by introducing mutations into the binding site derived from NMR experiments and performing proteoliposome transport assays. As fatty acids possess a negatively charged head-group, it was reasoned that positively charged residues might be involved in binding. Therefore, several positively charged residues in the vicinity of the proposed fatty acid binding site were mutated to Ser or Thr.

Very interestingly, some of these mutations abolished the ASO<sup>-</sup> transport by UCP2. More importantly, the inhibition of ASO<sup>-</sup> transport was strongly correlated with the inhibition of H<sup>+</sup> transport. It was therefore concluded that H<sup>+</sup> transport by UCP2 depend on the transport of fatty acids.

The substitutions displaying the strongest inhibition of transport (50%) included

---

<sup>36</sup>The pK<sub>a</sub> of ASO<sup>-</sup> is 1, while the pK<sub>a</sub> of palmitic acid is 4.75. The ionization state of both molecules, therefore, should be similar at pH 6.5, at which the measurements were performed.

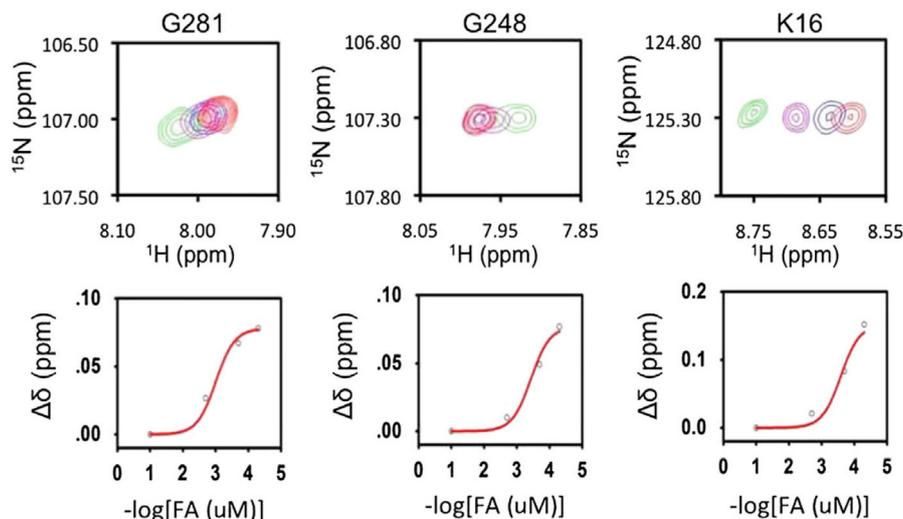


Figure 5.3: The fatty acid binding data provided by Berardi et al ((229)). Panels on the top display CSPs observed UPC2 spectra at increasing concentrations (50 $\mu\text{M}$  to 20 mM) of fatty acid. The correspondance between the color and the concentration is not indicated in the paper, however. Panels on the bottom illustrate plots of the data on top, where CSP is plotted against the concentration of the fatty acid. Upon a closer inspection it can be noticed that the concentration range mentioned in the paper does not quite correspond to the provided plots.

R60S (K51 in btAAC1), K271S (K271 in btAAC1), K141T (R137 in btAAC1) and R279T (R279 in btAAC1). All these residues are strongly conserved in all the mitochondrial carriers and not only in uncoupling proteins. More importantly, residues R60 and K271 are a part of the cardiolipin binding motifs [WYF][KR]G. It is well known that cardiolipin is essential for the function of mitochondrial carriers, and it is thus not surprising that such mutations have a strong effect on the UCP2 transport. Residue K141 belongs to the PX[DE]XX[KR] motif and participates in the formation of the matrix salt bridge network. A strong inhibitory effect upon substitution of this residue may also be anticipated. Residue K279 also appears to be conserved between UCP2 and AAC1, its functional role, however, is not clear.

Altogether there is strong evidence, that the reduced transport of  $\text{ASO}^-$  by UCP2 mutants is linked to the general disruption of the carrier's function, and not a specific disruption of the fatty acid binding site. It follows, therefore, that the results obtained by the NMR studies of fatty acid binding to UCP2 in DPC micelles are not corroborated by the functional studies of UCP2 mutants in liposomes. It appears, on the contrary, that the fatty acid binding to UCP2 may be non-specific, as we have established for the binding of cardiolipin or POPG to GGC1.

### 5.3 Studies on ATP/Mg- $\text{P}_i$ transporter inhibition by the cytoplasmic $\text{Ca}^{2+}$ domain

Another study suggesting the relevance of DPC-reconstituted mitochondrial carriers was presented by Yang et al. (81) In this study, the mechanism of short  $\text{Ca}^{2+}$ -

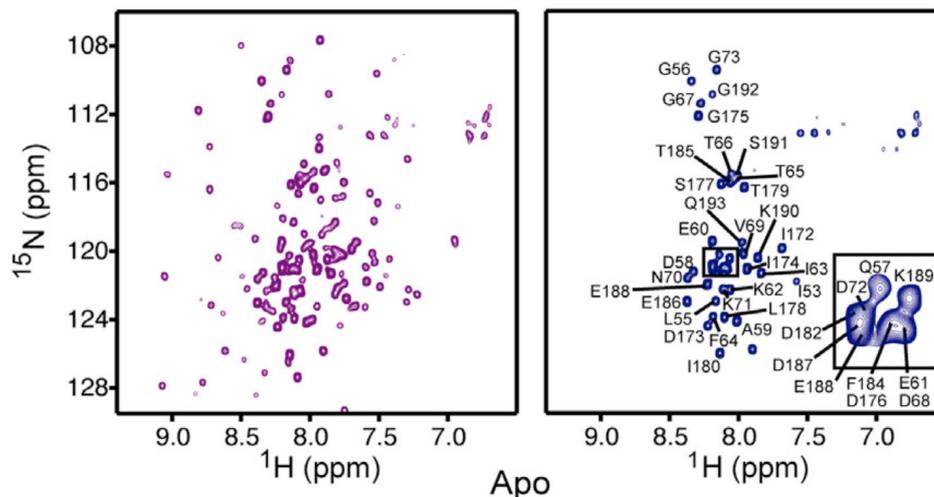


Figure 5.4: Figure illustrating the data, which was used to suggest the mechanism of SCaMC1 NTD binding to the TM domain. On the left is the spectrum of SCaMC1 NTD in the presence of empty liposomes. On the right the spectrum obtained after mixing 100  $\mu\text{M}$  of SCaMC1 NTD with 100  $\mu\text{M}$  of SCaMC1 TM domain. No peaks of the structured regions are visible. The intense peaks likely corresponding to the loop regions can be perfectly superposed in both of the spectra. Figure taken from Yang et al.

binding mitochondrial carrier (SCaMC1) inhibition by  $\text{Ca}^{2+}$  ions was investigated. This carrier, as opposed to many mitochondrial carriers, contains an N-terminal domain, which is involved in  $\text{Ca}^{2+}$  binding. Upon removal of the  $\text{Ca}^{2+}$  ions the carrier is inhibited by an unknown mechanism.

The N-terminal domain was characterized in the study in the presence and absence of  $\text{Ca}^{2+}$  ions. More disorder was observed when  $\text{Ca}^{2+}$  was removed from the domain, and it was suggested that these conformational changes might be important for the inhibition of the carrier.

To investigate the inhibitory effect, the interaction between N-terminal domain and TM domain was studied by NMR. To do so, the TM domain of SCaMC1 was reconstituted in the proteoliposomes. The N-terminal domain was then mixed together with the TM domain and the changes in the NMR spectra of N-terminal domain were observed. If the presence of  $\text{Ca}^{2+}$  ions were affecting the binding between the two domains, changes in spectra were anticipated, as the NTD bound to the proteoliposome-embedded TM domain would virtually become invisible to NMR, due to its decreased tumbling rate.

100  $\mu\text{M}$  of N-terminal domain in the presence of  $\text{Ca}^{2+}$  displayed virtually identical NMR spectra regardless if 100  $\mu\text{M}$  of TM domain was present in the liposomes. However, in the absence of  $\text{Ca}^{2+}$  ions, the peaks corresponding to the structured (therefore rigid) parts of NTD completely disappeared and only the loop regions remained visible when the TM domain was present in the liposomes (Figure 5.4). These observations were interpreted, therefore, as the binding of the whole NTD inside of the cavity of TM domain.

However, the complete disappearance of the structured regions from the spectra is somehow surprising. First of all, the  $K_D$  of the NTD binding to TM domain was determined to be  $\approx 250 \mu\text{M}$  by SPR. Given this binding constant, and the reported experimental concentrations of  $100 \mu\text{M}$  for both NTD and TM domain in the sample, only  $\approx 50\%$  of the NTD should be associated with the TM domain, and the cross-peaks corresponding to the unbound fraction should still be visible. Moreover, additional considerations should be taken into account. First of all, the reconstitution yield is assumed as  $100\%$  in their study. However, this usually is not the case for the mitochondrial carriers<sup>37</sup>. Furthermore, the orientation of proteins which do not have pronounced extramembranous domains is generally assumed to be random or close to random (206). We thus could expect that roughly half of the protein would expose a binding incompetent matrix side to the exterior of the liposome. Lastly, not all the TM domains might be open in the c-state, which is able to bind the NTD. With all these considerations in mind, the NTD binding competent fraction of TM domain may be much less than  $50\%$ . It cannot be assumed, therefore, that the spectrum of apo NTD in the presence of TM domain corresponds to the binding between these two domains.

#### 5.4 Molecular basis of MgATP Selectivity of the Mitochondrial SCaMC carrier

Run et al (225) have investigated the mechanism defining the selectivity of SCaMC1 carrier towards  $\text{ATP} \cdot \text{Mg}^{2+}$  ions as compared to ATP. The NMR experiments were performed on SCaMC1-TMD (the transmembrane domain of the carrier) reconstituted in  $60 \text{ mM}$  DPC micelles.

Binding of SCaMC1-TMD to ATP and  $\text{ATP} \cdot \text{Mg}^{2+}$  was investigated by following the CSPs. The binding data taken from the supplementary materials of the paper is presented in the Figure 5.5. The plots show CSP information for all non-overlapping residues in the protein, in the presence of  $\approx 17 \text{ mM}$  of ATP or  $\text{ATP} \cdot \text{Mg}^{2+}$ . According to the paper, residues I220, Q223, V224, K313, A317 and A457 all located in the cavity of the carrier display strong chemical shift perturbations. However, upon closer inspection of the figure 5.5 we can only find CSP information for residues A317 and A457. The chemical shift perturbations for these residues are in fact small if compared to the region 333-340, which encompasses residues in one of the matrix helices. This is not mentioned in the paper and the reasons for large CSPs are not discussed.

We can also see in the figure 5.5 that the binding of ATP induced somewhat larger CSPs in SCaMC1-TMD as compared to  $\text{ATP} \cdot \text{Mg}^{2+}$ . This observation is surprising, as  $\text{ATP} \cdot \text{Mg}^{2+}$  is a preferred substrate for SCaMC1, and the  $K_M$  is smaller for  $\text{ATP} \cdot \text{Mg}^{2+}$  ( $0.2 \text{ mM}$ ) as compared to ATP ( $0.33 \text{ mM}$ ) (226).

Strong perturbations in the matrix helices, low  $K_D$  values for the substrates and no preference for  $\text{ATP} \cdot \text{Mg}^{2+}$  are the properties of SCaMC1 carrier in DPC micelles which requires complicated explanations, if the carrier is assumed to be in its native

<sup>37</sup>Typical yields of correctly refolded protein is generally below  $5\%$  (L. Capobianco, personal communication).

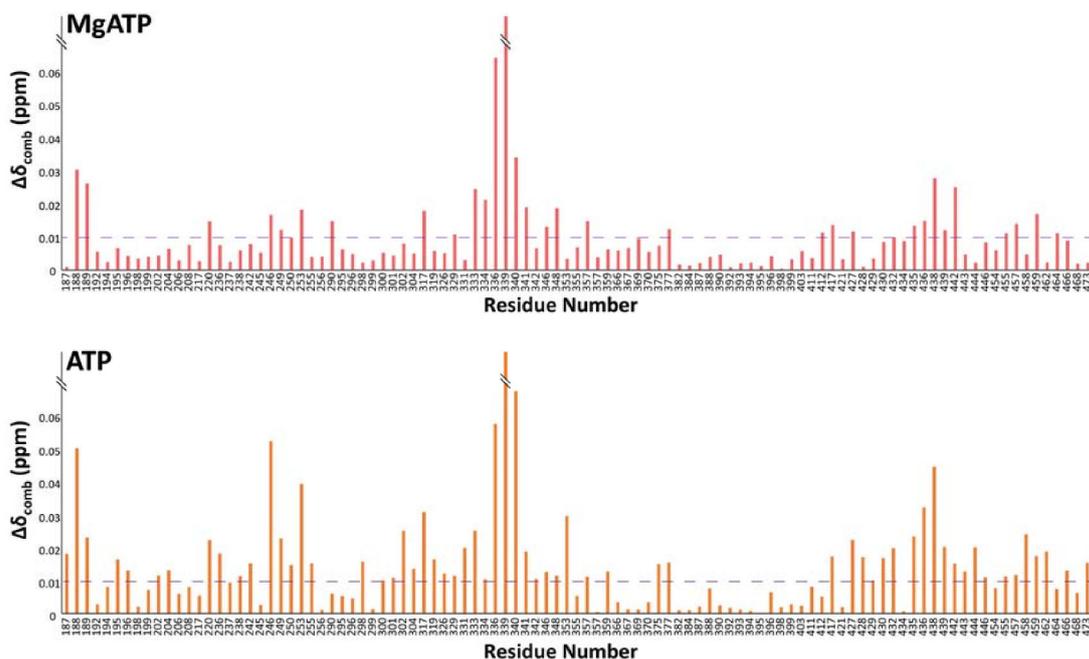


Figure 5.5: Figure illustrating the chemical shift perturbations induced in SCaMC1 upon addition of 17 mM of ATP · Mg<sup>2+</sup> (top) or 17 mM ATP (bottom). Similar binding profile is observed, while slightly smaller CSPs are induced by a preferred substrate, ATP · Mg<sup>2+</sup>. Note also, that the region with strongest chemical shift perturbations correspond to the matrix helices. Figure taken from the supplementary material of ref. (225)

state. However, these observations are in a very good agreement to what we have observed for denatured GGC1 and AAC3 carriers: mitochondrial helices display an abundance of positively charged residues, where the ligands can bind non-specifically. Furthermore, if we consider non-specific substrate binding inside the charged cavity of the denatured transporter, stronger ATP binding could be anticipated due to the lower negative charge of ATP · Mg<sup>2+</sup> complex.

This interpretation is further substantiated when the reported secondary structures are considered. If the Talos+ predictions for SCaMC1-TMD reported in the paper (Figure 1, ref. (225) are plotted on a homology model of the protein, multiple TM regions display the lack of  $\alpha$ -helical structures, see Figure 5.6). Structural perturbations are most evident in the helices 3 and 4. Helix 2, on the other hand, was not assigned. This could be due to the presence of millisecond dynamics leading to increased relaxation rates, as we have observed in the case of DPC-reconstituted ORC1.

Additionally, substitutions introduced in the protein did not result in a visible change in the spectra, as reported in the supplementary material of the paper: an observation which is in agreement with our data on GGC1 and AAC3. However, the regions where these substitutions were introduced do not appear to be in well packed regions in structural model, and thus it is not clear whether large perturbations could be expected.

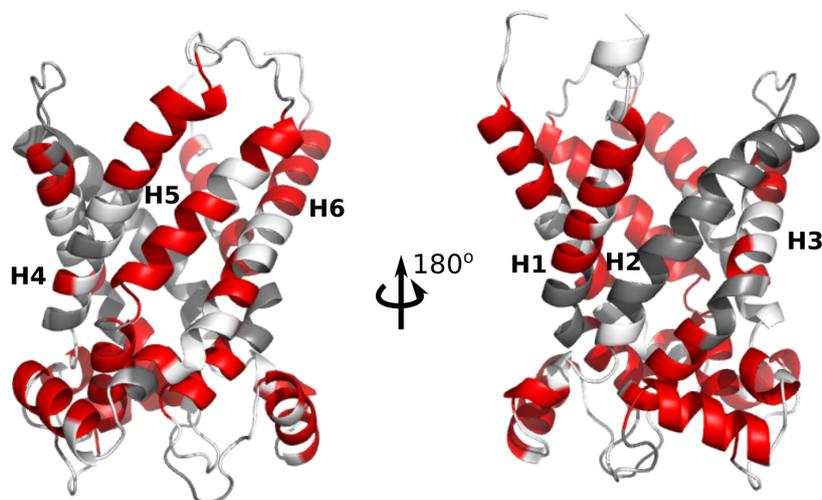


Figure 5.6: Talos+ helicity prediction data taken from Run et al., plotted on the homology model of SCaMC1. Multiple TM regions can be seen to lack properly folded  $\alpha$ -helices, similarly to what we have observed on AAC3, GGC1 and ORC1.

It therefore appears that the SCaMC1-TMD data are in very good agreement with our observations on DPC denatured mitochondrial carriers.

It has to be noted here, however, that among the residues identified to be involved in the binding of  $\text{ATP} \cdot \text{Mg}^{2+}$ , one residue (D361), which resides inside the cavity, appeared to be important for determining SCaMC1 substrate specificity in liposome transport assays. However, as already discussed for UCP2 transport of lipids, the agreement between observations from solution state NMR and functional assays might be purely accidental. Moreover, replacement of this residue resulted in barely equalization of transport rates, while AAC is actually inhibited in the presence of  $\text{Mg}^{2+}$ .

## 5.5 Summary of the observations on the mitochondrial carriers in DPC

After a closer look of published literature on mitochondrial carriers in DPC micelles it appears, that other carriers, such as UCP2 and SCaMC1 display similar behaviour as we have observed on AAC3, ORC1 and GGC1. It thus appears that DPC is not compatible with the biological function of mitochondrial carriers<sup>38</sup>. We have also gained some insight into the structural perturbations induced by DPC. In contrast to a static structure, our observations suggest that mitochondrial carriers are extremely flexible in detergent micelles. Not only the helices are not associated into well-defined tertiary structures, the secondary structures experience major perturbations as well.

<sup>38</sup>We have already shown that the refolding of the carrier is likely not an issue, as carriers refolded from the inclusion bodies of *E. coli* or expressed in a cell-free system display very similar NMR properties.

We do not know, however, if mitochondrial carriers form a particular class of membrane proteins which are more prone to denaturation by DPC, or the detergent itself is more deleterious than other detergents to the structures of membrane proteins.

## 5.6 Studies of other membrane proteins in DPC

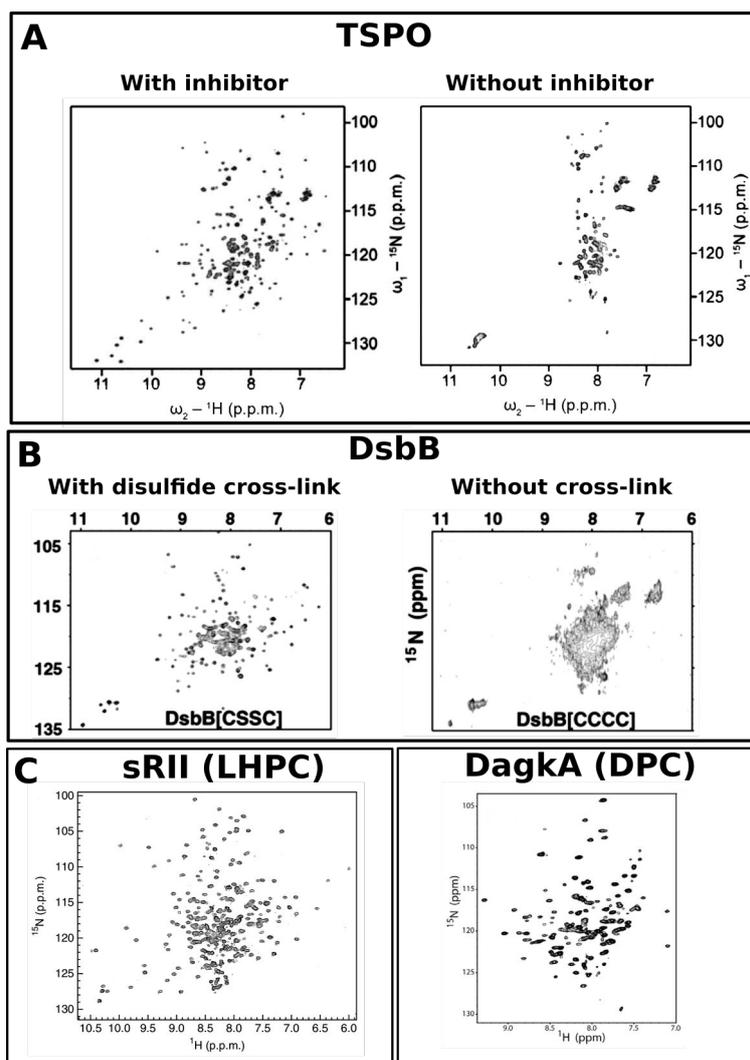


Figure 5.7: Helical transmembrane proteins not necessarily show a narrow peak dispersion. A, TSPO with (left) and without (right) PK11195 in DPC micelles; middle - DsbB with (left) and without (right) a stabilizing disulfide bridge in DPC micelles; C, spectrum of sensory rhodopsin in LHPC micelles (left) and a spectrum of DgkA in DPC micelles (right). Both proteins are of similar size.

DPC is by far the most common detergent used for the NMR studies of membrane proteins, due to its strong solubilizing capacity and small size of the micelles. It is, however, very rarely used in the field of X-ray crystallography. To our knowledge, only four crystal structures of membrane proteins were solved in DPC micelles.

Among these proteins, there is an outer membrane protein OmpF (177) ( $\beta$ -barrel proteins are generally more resistant to the denaturation by detergents), a membrane anchored monoamide oxidase (178) (with only a single TM helix) and a dimeric metal sensor NccX (244), containing two TM helices. Only a single protein, zinc transporter YiiP, for which a crystal structure was solved in DPC (176), is a polytopic membrane protein. It has to be mentioned, however, that diffracting crystals of an inhibited GPCR  $\beta$ 1-adrenergic receptor were obtained in DPC (179).

Given such a low amount of crystal structures, it would be tempting to speculate that DPC is indeed inducing similar flexible states as we have observed for mitochondrial carriers in many membrane proteins, and thus they are not able to crystallize any more. We do not know, however, if the low number of crystal structures solved in DPC is due to the inability of proteins to form crystals or whether this detergent is generally excluded from crystallography screens due to its bad reputation among the community of crystallographers.

Considering the NMR studies of membrane proteins in DPC, one interesting feature is a very low dispersion of  $\alpha$ -helical membrane protein amide spectra ( $\approx 2$  ppm). We know that helix packing likely provides an important contribution to the membrane protein stability. Well packed helices would be expected to display rather broad dispersion of NMR cross-peaks, as for example observed for the sensory rhodopsin spectra in DHPC (126). The narrow dispersion could therefore be attributed to an inherent flexibility of these membrane proteins in DPC.

From the structural studies we know, that low spectral dispersion can often be improved by addition of strongly binding ligands or restricting the flexibility of helices by introducing constraints, such as disulfide bonds. Some proteins, such as TSPO and DsbB display a very narrow dispersion of  $^1\text{H}$  frequencies (1-2 ppm), indicating flexible nature of these  $\alpha$ -helical membrane proteins. Upon addition of ligands, such as PK11195, ubiquinone or rimantidine, a large increase in spectral dispersion is observed: TSPO dispersion increases from  $\approx 1$  ppm to  $\approx 4$  ppm and DsbB displays an increase in dispersion from  $\approx 1$ -1.5 ppm to 4 ppm (Figure 5.7A, B). The increase in dispersion may be attributed to the aromatic nature of the ligands used. However, as PK11195 possesses an aromatic rings, which might induce very large effects on  $^1\text{H}$  chemical shifts due to the ring-currents, the increase in peak dispersion of DsbB results from a disulfide bond cross-link of two sequentially non-adjacent helices.

Other proteins, such as diacylglycerol kinase (DgkA), might display a very low spectral dispersion (see Figure 5.7C) in DPC micelles, although in DM detergent it only unfolds (in a cooperative transition) at  $90^\circ\text{C}$  (29), suggesting a well-packed structure.

Although inherent flexibility of membrane proteins was suggested to result in the poor resonance-cross peak dispersion, it has to be verified for each individual study whether this is not only due to effects of the detergent used, as we have observed for mitochondrial carriers.

Albeit few, there are other reports suggesting that DPC can denature membrane proteins. For example it was observed that a very stable protein, bacteriorhodopsin

starts to loose bound retinal cofactor (and thereby is denatured) at DPC concentrations above 80 mM (186). Detergent screens for X-ray crystallography have also revealed a strong denaturing ability of fos-choline detergents as observed in the loss of activity of Dengue fever protease NS2BFL-NS3pro (180), multidrug resistance protein 3 (MDR3) and bile salt export pump (BSEP) (181). The crystal structure of NccX, solved in DPC, also displayed a conformation of the protein, which was not in agreement with the native state (however, only 3 mM of DPC were used, and it is possible that the observed collapse of the structure was due to insufficient amount of detergent).

### 5.6.1 Perspectives

The use of detergents is very often unavoidable in the functional and structural studies of membrane proteins. The hydrophobic environment provided by detergents can be very different from the native bilayer where these proteins reside. Although it is well known that different detergents may inactivate membrane proteins, how these detergents affect membrane protein structure, and inturn their function, is very poorly understood. We have presented here a detailed study of such effects of DPC detergent on the mitochondrial carriers, as observed by NMR spectroscopy.

Although DPC appeared to perturb the tertiary and secondary structures as well as substrate specificity of all the mitochondrial carriers we have studied, this study provides a valuable insight into the details regarding membrane protein denaturation in the presence of detergents. Instead of populating one or several misfolded conformations, mitochondrial carriers appear to lose their function due to increased flexibility. A highly flexible detergent-denatured state of a membrane protein is in agreement with the fact that many membrane proteins can be refolded in bilayers after being solubilized in strong detergents, such as SDS. This highly flexible state of membrane proteins can be reasoned to occur mainly due to two factors: 1) the lack of stabilizing bilayer forces in detergent micelle and 2) strong competition of detergent molecules with protein-protein van der Waals contacts.

Membrane protein folding is relatively poorly understood as compared to soluble proteins. More high resolution structural studies on different proteins and detergents are necessary to obtain a better understanding of the folding and stability of this extremely important class of proteins.



## References

- (1) Denning, E. J., and Beckstein, O. (2013). Influence of lipids on protein-mediated transmembrane transport. *Chemistry and Physics of Lipids* 169, 57–71.
- (2) Nick Pace, C., Martin Scholtz, J., and Grimsley, G. R. (2014). Forces stabilizing proteins. *FEBS Letters* 588, 2177–2184.
- (3) Wimley, W. C. & W. S. H. (1999). Membrane protein folding and stability: physical principles. *Annual Review of Biophysics & Biomolecular Structure* 28, 319–365.
- (4) Popot, J. L. (2014). Folding membrane proteins in vitro: A table and some comments. *Archives of Biochemistry and Biophysics* 564, 314–326.
- (5) Nymeyer, H., and Zhou, H.-X. (2008). A method to determine dielectric constants in non-homogeneous systems: application to biological membranes. *Biophysical journal* 94, 1185–1193.
- (6) Sanders, C. R., and Mittagdorff, K. F. (2011). Tolerance to changes in membrane lipid composition as a selected trait of membrane proteins. *Biochemistry* 50, 7858–7867.
- (7) Popot, J.-l., and Engelman, D. M. (2000). HELICAL MEMBRANE PROTEIN FOLDING, STABILITY, AND EVOLUTION., 881–922.
- (8) Bowie, J. U. (2005). Solving the membrane protein folding problem. *Nature* 438, 581–589.
- (9) Zhou, H. X., and Cross, T. A. (2013). Modeling the membrane environment has implications for membrane protein structure and function: Influenza A M2 protein. *Protein Science* 22, 381–394.
- (10) Pogozheva, I. D., Mosberg, H. I., and Lomize, A. L. (2014). Life at the border: Adaptation of proteins to anisotropic membrane environment. *Protein Science* 23, 1165–1196.
- (11) von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *The EMBO Journal* 5, 3021–3027.
- (12) Stevens, T. J., and Arkin, I. T. (1999). Are membrane proteins 'inside-out' proteins? *Proteins: Structure, Function and Genetics* 36, 135–143.
- (13) Engelman, D. M., and Zaccari, G. (1980). Bacteriorhodopsin is an inside-out protein. *Proceedings of the National Academy of Sciences of the United States of America* 77, 5894–8.
- (14) De Marothy, M. T., and Elofsson, A. (2015). Marginally hydrophobic transmembrane  $\alpha$ -helices shaping membrane protein folding. *Protein Science* 24, 1057–1074.
- (15) Cymer, F., Von Heijne, G., and White, S. H. (2015). Mechanisms of integral membrane protein insertion and folding. *Journal of Molecular Biology* 427, 999–1022.
- (16) Gao, J., Bosco, D. a., Powers, E. T., and Kelly, J. W. (2009). Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nature structural & molecular biology* 16, 684–690.
- (17) Killian, J. A., and Von Heijne, G. (2000). How proteins adapt to a membrane-water interface. *Trends in Biochemical Sciences* 25, 429–434.
- (18) MacCallum, J. L., Bennett, W. F. D., and Tieleman, D. P. (2008). Distribution of amino acids in a lipid bilayer from computer simulations. *Biophysical journal* 94, 3393–3404.
- (19) Adamian, L., Nanda, V., DeGrado, W. F., and Liang, J. (2005). Empirical lipid propensities of amino acid residues in multispans alpha helical membrane proteins. *Proteins: Structure, Function and Genetics* 59, 496–509.
- (20) Ulmschneider, M. B., and Sansom, M. S. P. (2001). Amino acid distributions in integral membrane protein structures. *Biochim Biophys Acta* 1512, 1–14.

- (21) Choma, C., Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2000). Asparagine-mediated self-association of a model transmembrane helix. *Nature structural biology* 7, 161–166.
- (22) Chen, G. Q., and Gouaux, E. (1999). Probing the folding and unfolding of wild-type and mutant forms of bacteriorhodopsin in micellar solutions: Evaluation of reversible unfolding conditions. *Biochemistry* 38, 15380–15387.
- (23) Placone, J., and Hristova, K. (2012). Direct Assessment of the Effect of the Gly380Arg Achondroplasia Mutation on FGFR3 Dimerization Using Quantitative Imaging FRET. *PLoS ONE* 7, 1–7.
- (24) Manni, S., Mineev, K. S., Usmanova, D., Lyukmanova, E. N., Shulepko, M. A., Kirpichnikov, M. P., Winter, J., Matkovic, M., Deupi, X., Arseniev, A. S., and Ballmer-Hofer, K. (2014). Structural and functional characterization of alternative transmembrane domain conformations in VEGF receptor 2 activation. *Structure* 22, 1077–1089.
- (25) Sarabipour, S., and Hristova, K. (2016). Mechanism of FGF receptor dimerization and activation. *Nature communications* 7, 10262.
- (26) Nemethy, G., and Nberg, I. Z. S. (1963). Structure and of Hydrophobic Interactions. *Biopolymers* 1, 43–69.
- (27) Joh, N. H., Min, A., Faham, S., Whitelegge, J. P., Yang, D., Jr, V. L. W., and Bowie, J. U. (2008). Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *453*, DOI: 10.1038/nature06977.
- (28) Baker, R. P., and Urban, S. (2012). Architectural and thermodynamic principles underlying intramembrane protease function. *Nature chemical biology* 8, 759–68.
- (29) Lau, F. W., and Bowie, J. U. (1997). A method for assessing the stability of a membrane protein. *Biochemistry* 36, 5884–5892.
- (30) Wen, J., Chen, X., and Bowie, J. U. (1996). Exploring the allowed sequence space of a membrane protein. *Nat Struct Biol* 3, 141–148.
- (31) Cao, Z., and Bowie, J. U. (2012). Shifting hydrogen bonds may produce flexible transmembrane helices. *109*, 8121–8126.
- (32) Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001). The Calpha —H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proceedings of the National Academy of Sciences of the United States of America* 98, 9056.
- (33) Mueller, B. K., Subramaniam, S., and Senes, A. (2014). A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical Ca-H hydrogen bonds. *Proceedings of the National Academy of Sciences of the United States of America* 111, E888–95.
- (34) Khadria, A. S., Mueller, B. K., Stefely, J. A., Tan, C. H., Pagliarini, D. J., and Senes, A. (2014). A gly-zipper motif mediates homodimerization of the transmembrane domain of the mitochondrial kinase ADCK3. *Journal of the American Chemical Society* 136, 14068–14077.
- (35) Yohannan, S., Faham, S., Yang, D., Grosfeld, D., Chamberlain, A. K., and Bowie, J. U. (2004). A C alpha-H...O hydrogen bond in a membrane protein is not stabilizing. *Journal of the American Chemical Society* 126, 2284–2285.
- (36) Senes, A., Engel, D. E., and DeGrado, W. F. (2004). Folding of helical membrane proteins: The role of polar, GxxxG-like and proline motifs. *Current Opinion in Structural Biology* 14, 465–479.
- (37) Robinson, A. J., Overy, C., and Kunji, E. R. S. (2008). The mechanism of transport by mitochondrial carriers based on analysis of symmetry (Supplementary information). *Proceedings of the National Academy of Sciences of the United States of America* 105, 17766–17771.

- (38) Solcan, N., Kwok, J., Fowler, P. W., Cameron, A. D., Drew, D., Iwata, S., and Newstead, S. (2012). Alternating access mechanism in the POT family of oligopeptide transporters. *The EMBO journal* *31*, 3411–3421.
- (39) Stelzl, L. S., Fowler, P. W., Sansom, M. S. P., and Beckstein, O. (2014). Flexible gates generate occluded intermediates in the transport cycle of LacY. *Journal of Molecular Biology* *426*, 735–751.
- (40) Cui, G., Freeman, C. S., Knotts, T., Prince, C. Z., Kuang, C., and McCarty, N. A. (2013). Two salt bridges differentially contribute to the maintenance of cystic fibrosis transmembrane conductance regulator (CFTR) channel function. *Journal of Biological Chemistry* *288*, 20758–20767.
- (41) Kim, J.-M., Altenbach, C., Kono, M., Oprian, D. D., Hubbell, W. L., and Khorana, H. G. (2004). Structural origins of constitutive activation in rhodopsin: Role of the K296/E113 salt bridge. *Proc Natl Acad Sci U S A* *101*, 12508–12513.
- (42) Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y., and Yan, N. (2012). Crystal structure of a bacterial homologue of glucose transporters GLUT1–4. *Nature* *490*, 361–366.
- (43) Hong, H. (2014). Toward understanding driving forces in membrane protein folding. *Archives of Biochemistry and Biophysics* *564*, 297–313.
- (44) Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trézéguet, V., Lauquin, G. J.-M., and Brandolin, G. (2003). Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* *426*, 39–44.
- (45) Jaremko, L. (2014). Structure of the Mitochondrial translocator protein in complex with a diagnostic ligand. *343*, 1363–1367.
- (46) Yau, W. M., Wimley, W. C., Gawrisch, K., and White, S. H. (1998). The preference of tryptophan for membrane interfaces. *Biochemistry* *37*, 14713–14718.
- (47) Domene, C., Bond, P. J., Deol, S. S., and Sansom, M. S. P. (2003). Lipid / Protein Interactions and the Membrane / Water Interfacial Region Lipid / Protein Interactions and the Membrane / Water Interfacial Region. *125*, 14966–14967.
- (48) Hong, H., and Tamm, L. K. (2004). Elastic coupling of integral membrane protein stability to lipid bilayer forces. *Proc Natl Acad Sci U S A* *101*, 4065–4070.
- (49) Holt, A., Rougier, L., R?at, V., Jolibois, F., Saurel, O., Czaplicki, J., Antoinette Killian, J., and Milon, A. (2010). Order parameters of a transmembrane helix in a fluid bilayer: Case study of a WALP peptide. *Biophysical Journal* *98*, 1864–1872.
- (50) Soubias, O., Teague, W. E., Hines, K. G., and Gawrisch, K. (2015). Rhodopsin/Lipid Hydrophobic Matching—Rhodopsin Oligomerization and Function. *Biophysical Journal* *108*, 1125–1132.
- (51) Iversen, L., Mathiasen, S., Larsen, J. B., and Stamou, D. (2015). Membrane curvature bends the laws of physics and chemistry. *Nature Chemical Biology* *11*, 822–825.
- (52) Killian, J. A. (1998). Hydrophobic mismatch between proteins and lipids in membranes. *Biochimica et Biophysica Acta - Reviews on Biomembranes* *1376*, 401–415.
- (53) Brown, M. F. (2012). Curvature Forces in Membrane Lipid Protein Interactions. *Biochemistry* *51*, 9782–9795.
- (54) Yasuda, S., Kajiwara, Y., Takamuku, Y., Suzuki, N., Murata, T., and Kinoshita, M. (2016). Identification of Thermostabilizing Mutations for Membrane Proteins: Rapid Method Based on Statistical Thermodynamics. *The Journal of Physical Chemistry B*, acs.jpcc.6b01405.
- (55) Allen, S. J., Curran, A. R., Templer, R. H., Meijberg, W., and Booth, P. J. (2004). Controlling the folding efficiency of an integral membrane protein. *Journal of Molecular Biology* *342*, 1293–1304.

- (56) Miller, D., Charalambous, K., Rotem, D., Schuldiner, S., Curnow, P., and Booth, P. J. (2009). In vitro Unfolding and Refolding of the Small Multidrug Transporter EmrE. *Journal of Molecular Biology* 393, 815–832.
- (57) MacKenzie, K. R., and Fleming, K. G. (2008). Association energetics of membrane spanning  $\alpha$ -helices. *Current Opinion in Structural Biology* 18, 412–419.
- (58) Chen, L., Novicky, L., Merzlyakov, M., Hristov, T., and Hristova, K. (2010). Measuring the energetics of membrane protein dimerization in mammalian membranes. *Journal of the American Chemical Society* 132, 3628–3635.
- (59) Dong, H., Sharma, M., Zhou, H.-X., and Cross, T. a. (2012). Glycines: Role in alpha-Helical Membrane Protein Structures and a Potential Indicator of Native Conformation. *Biochemistry* 51, 4779–4789.
- (60) Yang, L., Adam, C., Nichol, G. S., and Cockroft, S. L. (2013). How much do van der Waals dispersion forces contribute to molecular recognition in solution? *Nature Chemistry* 5, 1006–1010.
- (61) Fiedler, S., Broecker, J., and Keller, S. (2010). Protein folding in membranes. *Cellular and molecular life sciences : CMLS* 67, 1779–1798.
- (62) Joh, N. H., Oberai, A., Yang, D., Whitelegge, J. P., and Bowie, J. U. (2009). Similar energetic contributions of packing in the core of membrane and water-soluble proteins. *Journal of the American Chemical Society* 131, 10846–10847.
- (63) Gall, A., Ellervee, A., Sturgis, J., Fraser, N., Cogdell, R., Freiberg, A., and Robert, B. (2003). Membrane Protein Stability: High Pressure Effects on the Structure and Chromophore-Binding Properties of the Light-Harvesting Complex LH2. *Biochemistry* 42, 13019–13026.
- (64) Bugge, K., Lindorff-Larsen, K., and Kragelund, B. B. (2016). Understanding single-pass transmembrane receptor signaling from a structural viewpoint - what are we missing? *The FEBS Journal*, 1–28.
- (65) Zhang, Y., Kulp, D. W., Lear, J. D., and DeGrado, W. F. (2009). Experimental and computational evaluation of forces directing the association of transmembrane helices. *Journal of the American Chemical Society* 131, 11341–11343.
- (66) Feng, X., and Barth, P. (2016). A topological and conformational stability alphabet for multipass membrane proteins. *Nature Chemical Biology* 12, 167–173.
- (67) Min, D., Jefferson, R. E., Bowie, J. U., and Yoon, T.-Y. (2015). Mapping the energy landscape for second-stage folding of a single membrane protein. *Nature chemical biology* 11, 981–7.
- (68) Robertson, J. L., Kolmakova-Partensky, L., and Miller, C. (2010). Design, function and structure of a monomeric ClC transporter. *Nature* 468, 844–7.
- (69) Hong, H., Blois, T. M., Cao, Z., and Bowie, J. U. (2010). Method to measure strong protein-protein interactions in lipid bilayers using a steric trap. *Proceedings of the National Academy of Sciences of the United States of America* 107, 19802–7.
- (70) Meruelo, A. D., Han, S. K., Kim, S., and Bowie, J. U. (2012). Structural differences between thermophilic and mesophilic membrane proteins. *Protein Science* 21, 1746–1753.
- (71) Popot, J. L., and Engelman, D. M. (2016). Membranes Do Not Tell Proteins How to Fold. *Biochemistry* 55, 5–18.
- (72) Hessa, T., Meindl-Beinker, N. M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S. H., and von Heijne, G. (2007). Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* 450, 1026–1030.

- (73) Bosshart, P. D., Iordanov, I., Garzon-Coral, C., Demange, P., Engel, A., Milon, A., and Müller, D. J. (2012). The transmembrane protein KpOmpA anchoring the outer membrane of *Klebsiella pneumoniae* unfolds and refolds in response to tensile load. *Structure* 20, 121–127.
- (74) Kessler, M., Gottschalk, K. E., Janovjak, H., Muller, D. J., and Gaub, H. E. (2006). Bacteriorhodopsin folds into the membrane against an external force. *Journal of Molecular Biology* 357, 644–654.
- (75) Thomas, J. A., and Tate, C. G. (2014). Quality Control in Eukaryotic Membrane Protein Overproduction. *Journal of Molecular Biology* 426, 4139–4154.
- (76) Indiveri, C., Tonazzi, A., Stipani, I., and Palmieri, F. (1997). The purified and reconstituted ornithine/citrulline carrier from rat liver mitochondria: electrical nature and coupling of the exchange reaction with H<sup>+</sup> translocation. *The Biochemical journal* 327 ( Pt 2, 349–355.
- (77) Palmieri, L., De Marco, V., Iacobazzi, V., Palmieri, F., Runswick, M. J., and Walker, J. E. (June 1997). Identification of the yeast ARG-11 gene as a mitochondrial ornithine carrier involved in arginine biosynthesis. *FEBS Letters* 410, 447–451.
- (78) Vozza, A., Blanco, E., Palmieri, L., and Palmieri, F. (2004). Identification of the mitochondrial GTP/GDP transporter in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 279, 20850–20857.
- (79) Todisco, S., Agrimi, G., Castegna, A., and Palmieri, F. (2006). Identification of the mitochondrial NAD<sup>+</sup> transporter in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 281, 1524–1531.
- (80) Dolce, V., Scarcia, P., Iacopetta, D., and Palmieri, F. (2005). A fourth ADP/ATP carrier isoform in man: Identification, bacterial expression, functional characterization and tissue distribution. *FEBS Letters* 579, 633–637.
- (81) Yang, Q., Br??schweiler, S., and Chou, J. J. (2014). A self-sequestered calmodulin-like ca<sup>2+</sup> sensor of mitochondrial ScaMC carrier and its implication to Ca<sup>2+</sup>-Dependent ATP-Mg/P<sub>i</sub> Transport. *Structure* 22, 209–217.
- (82) Etzkorn, M., Raschle, T., Hagn, F., Gelev, V., Rice, A. J., Walz, T., and Wagner, G. (2013). Cell-free expressed bacteriorhodopsin in different soluble membrane mimetics: Biophysical properties and NMR accessibility. *Structure* 21, 394–401.
- (83) Rawlings, A. E. (2016). Membrane proteins: always an insoluble problem? *Biochemical Society Transactions* 44, 790–795.
- (84) Cui, T., Mowrey, D., Bondarenko, V., Tillman, T., Ma, D., Landrum, E., Perez-Aguilar, J. M., He, J., Wang, W., Saven, J. G., Eckenhoff, R. G., Tang, P., and Xu, Y. (2012). NMR structure and dynamics of a designed water-soluble transmembrane domain of nicotinic acetylcholine receptor. *Biochimica et biophysica acta* 1818, 617–26.
- (85) Perez-Aguilar, J. M., Xi, J., Matsunaga, F., Cui, X., Selling, B., Saven, J. G., and Liu, R. (2013). A Computationally Designed Water-Soluble Variant of a G-Protein-Coupled Receptor: The Human Mu Opioid Receptor. *PLoS ONE* 8, DOI: 10.1371/journal.pone.0066009.
- (86) Li, H., Cocco, M. J., Steitz, T. A., and Engelman, D. M. (2001). Conversion of phospholamban into a soluble pentameric helical bundle. *Biochemistry* 40, 6636–6645.
- (87) Shibuya, I. (1992). Metabolic regulations and biological functions of phospholipids in *Escherichia coli*. *Progress in lipid research* 31, 245–99.
- (88) Hänsel, R., Luh, L. M., Corbeski, I., Trantirek, L., and Dötsch, V. (2014). In-cell NMR and EPR spectroscopy of biomacromolecules. *Angewandte Chemie (International ed. in English)* 53, 10300–10314.

- (89) Renault, M., Tommassen-van Boxtel, R., Bos, M. P., Post, J. A., Tommassen, J., and Baldus, M. (2012). Cellular solid-state nuclear magnetic resonance spectroscopy. *Proceedings of the National Academy of Sciences* 109, 4863–4868.
- (90) Dörr, J. M., Koorengevel, M. C., Schäfer, M., Prokofyev, A. V., Scheidelaar, S., van der Cruijssen, E. A. W., Dafforn, T. R., Baldus, M., and Killian, J. A. (2014). Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K<sup>+</sup> channel: the power of native nanodiscs. *Proceedings of the National Academy of Sciences of the United States of America* 111, 18607–12.
- (91) Chun, B. J., Choi, J. I., and Jang, S. S. (2015). Molecular dynamics simulation study of sodium dodecyl sulfate micelle: Water penetration and sodium dodecyl sulfate dissociation. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 474, 36–43.
- (92) Abel, S., Dupradeau, F. Y., and Marchi, M. (2012). Molecular dynamics simulations of a characteristic DPC micelle in water. *Journal of Chemical Theory and Computation* 8, 4610–4623.
- (93) Magid, L. J., and Li, Z. (2000). Flexibility of Elongated Sodium Dodecyl Sulfate Micelles in Aqueous Sodium Chloride : A Small-Angle Neutron Scattering Study. *Langmuir*, 10028–10036.
- (94) Rath, A., and Deber, C. M. (2012). Protein Structure in Membrane Domains. *Annual Review of Biophysics* 41, 135–155.
- (95) Newstead, S. S. F. S. I. (2007). Rationalizing  $\alpha$ -helical membrane protein crystalization. *Protein Science* 1, 455–472.
- (96) Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G., and Deber, C. M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* 106, 1760–5.
- (97) Otzen, D. E. (2011). Mapping the folding pathway of the transmembrane protein DsbB by protein engineering. *Protein Engineering, Design and Selection* 24, 139–149.
- (98) Curnow, P., and Booth, P. J. (2009). The transition state for integral membrane protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 106, 773–778.
- (99) Renthal, R. (2006). An unfolding story of helical transmembrane proteins. *Biochemistry* 45, 14559–14566.
- (100) Dutta, A., Kim, T. Y., Moeller, M., Wu, J., Alexiev, U., and Klein-Seetharaman, J. (2010). Characterization of membrane protein non-native states. 2. the SDS-unfolded states of rhodopsin. *Biochemistry* 49, 6329–6340.
- (101) Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J., and Engelman, D. M. (1992). Sequence specificity in the dimerization of transmembrane  $\alpha$ -helices. *Biochemistry* 31, 12719–12725.
- (102) Prodöhl, A., Weber, M., Dreher, C., and Schneider, D. (2007). A mutational study of transmembrane helix-helix interactions. *Biochimie* 89, 1433–1437.
- (103) Nielsen, M. M., Andersen, K. K., Westh, P., and Otzen, D. E. (2007). Unfolding of beta-sheet proteins in SDS. *Biophysical journal* 92, 3674–3685.
- (104) Nadeau, V. G., Rath, A., and Deber, C. M. (2012). Sequence hydrophathy dominates membrane protein response to detergent solubilization. *Biochemistry* 51, 6228–6237.
- (105) Pervushin, K. V., Orekhov VYu, Popov, a. I., Musina LYu, and Arseniev, a. S. (1994). Three-dimensional structure of (1-71)bacterioopsin solubilized in methanol/chloroform and SDS micelles determined by <sup>15</sup>N-1H heteronuclear NMR spectroscopy. *European journal of biochemistry / FEBS* 219, 571–583.

- (106) Howell, S. C., Mesleh, M. F., and Opella, S. J. (2005). NMR structure determination of a membrane protein with two transmembrane helices in micelles: MerF of the bacterial mercury detoxification system. *Biochemistry* 44, 5196–5206.
- (107) Lu, G. J., Tian, Y., Vora, N., Marassi, F. M., and Opella, S. J. (2013). The structure of the mercury transporter MerF in phospholipid bilayers: A large conformational rearrangement results from N-terminal truncation. *Journal of the American Chemical Society* 135, 9299–9302.
- (108) Krishnamani, V., Hegde, B. G., Langen, R., and Lanyi, J. K. (2012). Denatured State.
- (109) Dutta, A., Altenbach, C., Mangahas, S., Yanamala, N., Gardner, E., Hubbell, W. L., and Klein-Seetharaman, J. (2014). Differential dynamics of extracellular and cytoplasmic domains in denatured states of rhodopsin. *Biochemistry* 53, 7160–7169.
- (110) Krishnamani, V., Hegde, B. G., Langen, R., and Lanyi, J. K. (2012). Secondary and tertiary structure of bacteriorhodopsin in the SDS denatured state. *Biochemistry* 51, 1051–1060.
- (111) Tulumello, D. V., and Deber, C. M. (2011). Positions of polar amino acids alter interactions between transmembrane segments and detergents. *Biochemistry* 50, 3928–3935.
- (112) Tulumello, D. V., and Deber, C. M. (2009). SDS micelles as a membrane-mimetic environment for transmembrane segments. *Biochemistry* 48, 12096–12103.
- (113) Maslennikov, I., Klammt, C., Hwang, E., Kefala, G., Okamura, M., Esquivies, L., Mörs, K., Glaubitz, C., Kwiatkowski, W., Jeon, Y. H., and Choe, S. (2010). Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis. *Proceedings of the National Academy of Sciences of the United States of America* 107, 10902–7.
- (114) Sheu, S. Y., Schlag, E. W., Selzle, H. L., and Yang, D. Y. (2009). Hydrogen bonds in membrane proteins. *J Phys Chem B* 113, 5318–5326.
- (115) Krueger-Koplin, R. D., Sorgen, P. L., Krueger-Koplin, S. T., Rivera-Torres, a. O., Cahill, S. M., Hicks, D. B., Grinius, L., Krulwich, T. a., and Girvin, M. E. (2004). An evaluation of detergents for NMR structural studies of membrane proteins. *Journal of Biomolecular Nmr* 28, 43–57.
- (116) Page, R. C., Moore, J. D., Nguyen, H. B., Sharma, M., Chase, R., Gao, F. P., Mobley, C. K., Sanders, C. R., Ma, L., S?nnichsen, F. D., Lee, S., Howell, S. C., Opella, S. J., and Cross, T. A. (2006). Comprehensive evaluation of solution nuclear magnetic resonance spectroscopy sample preparation for helical integral membrane proteins. *Journal of Structural and Functional Genomics* 7, 51–64.
- (117) Eichmann, C., Tzitzilonis, C., Bordignon, E., Maslennikov, I., Choe, S., and Riek, R. (2014). Solution NMR structure and functional analysis of the integral membrane protein YgaP from Escherichia coli. *Journal of Biological Chemistry* 289, 23482–23503.
- (118) Butterwick, J. A., and MacKinnon, R. (2010). Solution Structure and Phospholipid Interactions of the Isolated Voltage-Sensor Domain from KvAP. *Journal of Molecular Biology* 403, 591–606.
- (119) Kucharska, I., Seelheim, P., Edrington, T., Liang, B., and Tamm, L. K. (2015). OprG Harnesses the Dynamics of its Extracellular Loops to Transport Small Amino Acids across the Outer Membrane of Pseudomonas aeruginosa. *Structure* 23, 2234–2245.
- (120) Edrington, T. C., Kintz, E., Goldberg, J. B., and Tamm, L. K. (2011). Structural basis for the interaction of lipopolysaccharide with outer membrane protein H (OprH) from Pseudomonas aeruginosa. *Journal of Biological Chemistry* 286, 39211–39223.
- (121) Reckel, S., Gottstein, D., Stehle, J., Löhr, F., Verhoeven, M. K., Takeda, M., Silvers, R., Kainosho, M., Glaubitz, C., Wachtveitl, J., Bernhard, F., Schwalbe, H., Güntert, P., and Dötsch, V. (2011). Solution NMR structure of proteorhodopsin. *Angewandte Chemie - International Edition* 50, 11942–11946.

- (122) Johansson, M. U., Alioth, S., Hu, K., Walser, R., Koebnik, R., and Pervushin, K. (2007). A minimal transmembrane  $\beta$ -barrel platform protein studied by nuclear magnetic resonance. *Biochemistry* 46, 1128–1140.
- (123) Arora, A., Abildgaard, F., Bushweller, J. H., and Tamm, L. K. (2001). Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nature structural biology* 8, 334–338.
- (124) Fernández, C., Adeishvili, K., and Wüthrich, K. (2001). Transverse relaxation-optimized NMR spectroscopy with the outer membrane protein OmpX in dihexanoyl phosphatidylcholine micelles. *Proceedings of the National Academy of Sciences of the United States of America* 98, 2358–2363.
- (125) Fernández, C., Hilty, C., Wider, G., Güntert, P., and Wüthrich, K. (2004). NMR Structure of the Integral Membrane Protein OmpX. *Journal of Molecular Biology* 336, 1211–1221.
- (126) Gautier, A., Mott, H. R., Bostock, M. J., Kirkpatrick, J. P., and Nietlispach, D. (2010). Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nature structural & molecular biology* 17, 768–774.
- (127) Schnell, J. R., and Chou, J. J. (2008). Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* 451, 591–595.
- (128) Wang, J., Pielak, R. M., McClintock, M. a., and Chou, J. J. (2009). Solution structure and functional analysis of the influenza B proton channel. *Nature structural & molecular biology* 16, 1267–71.
- (129) Pielak, R. M., Oxenoid, K., and Chou, J. J. (2011). Structural investigation of rimantadine inhibition of the AM2-BM2 chimera channel of influenza viruses. *Structure* 19, 1655–1663.
- (130) Fox, D. A., Larsson, P., Lo, R. H., Kroncke, B. M., Kasson, P. M., and Columbus, L. (2014). Structure of the neisserial outer membrane protein Opa60: Loop flexibility essential to receptor recognition and bacterial engulfment. *Journal of the American Chemical Society* 136, 9938–9946.
- (131) Langelaan, D. N., Reddy, T., Banks, A. W., Dellaire, G., Dupré, D. J., and Rainey, J. K. (2013). Structural features of the apelin receptor N-terminal tail and first transmembrane segment implicated in ligand binding and receptor trafficking. *Biochimica et Biophysica Acta - Biomembranes* 1828, 1471–1483.
- (132) Mineev, K. S., Bocharov, E. V., Volynsky, P. E., Goncharuk, M. V., Tkach, E. N., Ermolyuk, Y. S., Schulga, A. A., Chupin, V. V., Maslennikov, I. V., Efremov, R. G., and Arseniev, A. S. (2011). Dimeric structure of the transmembrane domain of glycophorin a in lipidic and detergent environments. *Acta naturae* 3, 90–8.
- (133) MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997). A transmembrane helix dimer: structure and implications. *Science* 276, 131–133.
- (134) Li, Q., Wong, Y. L., and Kang, C. (2014). Solution structure of the transmembrane domain of the insulin receptor in detergent micelles. *Biochimica et Biophysica Acta - Biomembranes* 1838, 1313–1321.
- (135) Mineev, K. S., Khabibullina, N. F., Lyukmanova, E. N., Dolgikh, D. A., Kirpichnikov, M. P., and Arseniev, A. S. (2011). Spatial structure and dimer-monomer equilibrium of the ErbB3 transmembrane domain in DPC micelles. *Biochimica et Biophysica Acta - Biomembranes* 1808, 2081–2088.
- (136) Bocharov, E. V., Lesovoy, D. M., Goncharuk, S. A., Goncharuk, M. V., Hristova, K., and Arseniev, A. S. (2013). Structure of FGFR3 transmembrane domain dimer: Implications for signaling and human pathologies. *Structure* 21, 2087–2093.

- (137) Liang, B., Kiessling, V., and Tamm, L. K. (2013). Prefusion structure of syntaxin-1A suggests pathway for folding into neuronal trans-SNARE complex fusion intermediate. *Proceedings of the National Academy of Sciences of the United States of America* 110, 19384–9.
- (138) Wang, J., Wu, Y., Ma, C., Fiorin, G., Wang, J., Pinto, L. H., Lamb, R. A., Klein, M. L., and DeGrado, W. F. (2012). Structure and inhibition of the drug-resistant S31N mutant of the M2 ion channel of influenza A virus. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1315–1320.
- (139) OuYang, B., Xie, S., Berardi, M. J., Zhao, X., Dev, J., Yu, W., Sun, B., and Chou, J. J. (2013). Unusual architecture of the p7 channel from hepatitis C virus. *Nature* 498, 521–525.
- (140) Rodriguez, F., Rouse, S. L., Tait, C. E., Harmer, J., De Riso, A., Timmel, C. R., Sansom, M. S. P., Berks, B. C., and Schnell, J. R. (2013). Structural model for the protein-translocating element of the twin-arginine transport system. *Proceedings of the National Academy of Sciences of the United States of America* 110, E1092–101.
- (141) Berardi, M. J., Shih, W. M., Harrison, S. C., and Chou, J. J. (2011). Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature* 476, 109–113.
- (142) Shi, L., Traaseth, N. J., Verardi, R., Gustavsson, M., Gao, J., and Veglia, G. (2011). Paramagnetic-based NMR restraints lift residual dipolar coupling degeneracy in multidomain detergent-solubilized membrane proteins. *Journal of the American Chemical Society* 133, 2232–2241.
- (143) Oxenoid, K., and Chou, J. J. (2005). The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proceedings of the National Academy of Sciences of the United States of America* 102, 10870–10875.
- (144) Vostrikov, V. V., Mote, K. R., Verardi, R., and Veglia, G. (2013). Structural dynamics and topology of phosphorylated phospholamban homopentamer reveal its role in the regulation of calcium transport. *Structure* 21, 2119–2130.
- (145) Williamson, J. a., Cho, S.-H., Ye, J., Collet, J.-F., Beckwith, J. R., and Chou, J. J. (2015). Structure and multistate function of the transmembrane electron transporter CcdA. *Nature structural & molecular biology* 22, 809–814.
- (146) Liang, B., and Tamm, L. K. (2007). Structure of outer membrane protein G by solution NMR spectroscopy. *Proc Natl Acad Sci U S A* 104, 16140–16145.
- (147) Hwang, P. M., Choy, W.-Y., Lo, E. I., Chen, L., Forman-Kay, J. D., Raetz, C. R. H., Privé, G. G., Bishop, R. E., and Kay, L. E. (2002). Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proceedings of the National Academy of Sciences of the United States of America* 99, 13560–5.
- (148) Papavoine, C. H., Christiaans, B. E., Folmer, R. H., Konings, R. N., and Hilbers, C. W. (1998). Solution structure of the M13 major coat protein in detergent micelles: a basis for a model of phage assembly involving specific residues. *Journal of Molecular Biology* 282, 401–19.
- (149) Bragin, P. E., Mineev, K. S., Bocharova, O. V., Volynsky, P. E., Bocharov, E. V., and Arseniev, A. S. (2016). HER2 Transmembrane Domain Dimerization Coupled with Self-Association of Membrane-Embedded Cytoplasmic Juxtamembrane Regions. *Journal of Molecular Biology* 428, 52–61.
- (150) Bocharov, E. V., Mineev, K. S., Volynsky, P. E., Ermolyuk, Y. S., Tkach, E. N., Sobol, A. G., Chupin, V. V., Kirpichnikov, M. P., Efremov, R. G., and Arseniev, A. S. (2008). Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *Journal of Biological Chemistry* 283, 6950–6956.

- (151) Call, M. E., Schnell, J. R., Xu, C., Lutz, R. A., Chou, J. J., and Wucherpennig, K. W. (2006). The Structure of the  $\alpha$ 5 $\beta$ 1 Transmembrane Dimer Reveals Features Essential for Its Assembly with the T Cell Receptor. *Cell* 127, 355–368.
- (152) Call, M. E., Wucherpennig, K. W., and Chou, J. J. (2010). The structural basis for intramembrane assembly of an activating immunoreceptor complex. *Nature immunology* 11, 1023–9.
- (153) Ellena, J. F., Liang, B., Wiktor, M., Stein, A., Cafiso, D. S., Jahn, R., and Tamm, L. K. (2009). Dynamic structure of lipid-bound synaptobrevin suggests a nucleation-propagation mechanism for trans-SNARE complex formation. *Proc Natl Acad Sci U S A* 106, 20306–20311.
- (154) Zhou, Y., Cierpicki, T., Jimenez, R. H. F., Lukasik, S. M., Ellena, J. F., Cafiso, D. S., Kadokura, H., Beckwith, J., and Bushweller, J. H. (2008). NMR Solution Structure of the Integral Membrane Enzyme DsbB : Functional Insights into DsbB-Catalyzed Disulfide Bond Formation. *Molecular Cell* 31, 896–908.
- (155) Van Horn, W. D., Kim, H. J., Ellis, C. D., Hadziselimovic, A., Sulistijo, E. S., Karra, M. D., Tian, C., Sonnichsen, F. D., and Sanders, C. R. (2009). Solution nuclear magnetic resonance structure of membrane integral diacylglycerol kinase. *Science* 324, 1726–1728.
- (156) Ling, S., Zhang, C., Wang, W., Cai, X., Yu, L., Wu, F., Zhang, L., and Tian, C. (2016). Combined approaches of EPR and NMR illustrate only one transmembrane helix in the human IFITM3. *Scientific reports* 6, 24029.
- (157) Nadezhdin, K. D., García-Carpio, I., Goncharuk, S. A., Mineev, K. S., Arseniev, A. S., and Vilar, M. (2016). Structural basis of p75 transmembrane domain dimerization. *Journal of Biological Chemistry* 291, 12346–12357.
- (158) Oxenoid, K., Dong, Y., Cao, C., Cui, T., Sancak, Y., Markhard, A. L., Grabarek, Z., Kong, L., Liu, Z., Ouyang, B., Cong, Y., Mootha, V. K., and Chou, J. J. (2016). Architecture of the mitochondrial calcium uniporter. *Nature* 533, 1–17.
- (159) Roosild, T. P., Greenwald, J., Vega, M., Castronovo, S., Riek, R., and Choe, S. (2005). NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* 307, 1317–1321.
- (160) Bondarenko, V., Mowrey, D. D., Tillman, T. S., Seyoum, E., Xu, Y., and Tang, P. (2014). NMR structures of the human  $\alpha$ 7 nAChR transmembrane domain and associated anesthetic binding sites. *Biochimica et Biophysica Acta - Biomembranes* 1838, 1389–1395.
- (161) Hiller, S., Garces, R. G., Malia, T. J., Orekhov, V. Y., Colombini, M., and Wagner, G. (2008). in Detergent Micelles. *Science* 321, 1206–1210.
- (162) Li, Y., Liew, L. S. Y., Li, Q., and Kang, C. (2016). Structure of the transmembrane domain of human nicastrin—a component of  $\gamma$ -secretase. *Scientific reports* 6, 19522.
- (163) Teriete, P., Franzin, C. M., Choi, J., and Marassi, F. M. (2007). Structure of the Na,K-ATPase regulatory protein FXYD1 in micelles. *Biochemistry* 46, 6774–6783.
- (164) Dev, J., Park, D., Fu, Q., Chen, J., Ha, H. J., Ghantous, F., Herrmann, T., Chang, W., Liu, Z., Frey, G., Seaman, M. S., Chen, B., and Chou, J. J. (2016). Structural basis for membrane anchoring of HIV-1 envelope spike. *Science (New York, N.Y.)* 353, 22341–22345.
- (165) Bocharov, E. V., Mineev, K. S., Goncharuk, M. V., and Arseniev, A. S. (2012). Structural and thermodynamic insight into the process of "weak" dimerization of the ErbB4 transmembrane domain by solution NMR. *Biochimica et Biophysica Acta - Biomembranes* 1818, 2158–2170.
- (166) Mineev, K. S., Bocharov, E. V., Pustovalova, Y. E., Bocharova, O. V., Chupin, V. V., and Arseniev, A. S. (2010). Spatial Structure of the Transmembrane Domain Heterodimer of ErbB1 and ErbB2 Receptor Tyrosine Kinases. *Journal of Molecular Biology* 400, 231–243.

- (167) Bocharov, E. V., Mayzel, M. L., Volynsky, P. E., Goncharuk, M. V., Ermolyuk, Y. S., Schulga, A. A., Artemenko, E. O., Efremov, R. G., and Arseniev, A. S. (2008). Spatial structure and pH-dependent conformational diversity of dimeric transmembrane domain of the receptor tyrosine kinase EphA1. *Journal of Biological Chemistry* 283, 29385–29395.
- (168) Yu, L. et al. (2005). Nuclear Magnetic Resonance Structural Studies of a Potassium Channel - Charybdotoxin Complex ‡. *41*, 15834–15841.
- (169) Verardi, R., Shi, L., Traaseth, N. J., Walsh, N., and Veglia, G. (2011). Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. *Proceedings of the National Academy of Sciences of the United States of America* 108, 9101–9106.
- (170) Cross, T. A., Murray, D. T., and Watts, A. (2013). Helical membrane protein conformations and their environment. *European Biophysics Journal* 42, 731–755.
- (171) Ghasriani, H., Kwok, J. K. C., Sherratt, A. R., Foo, A. C. Y., Qureshi, T., and Goto, N. K. (2014). Micelle-Catalyzed Domain Swapping in the GlpG Rhomboid Protease Cytoplasmic Domain. *Biochemistry* 53, 5907–5915.
- (172) Zoonens, M., Masscheleyn, S., Pebay-peyroula, E., Chipot, C., and Miroux, B. (2013). Dangerous Liaisons between Detergents and Membrane Proteins. The Case of Mitochondrial Uncoupling Protein 2.
- (173) Guo, Y., Kalathur, R. C., Liu, Q., Kloss, B., Bruni, R., Ginter, C., Kloppmann, E., Rost, B., and Hendrickson, W. A. (2015). Protein structure. Structure and activity of tryptophan-rich TSPO proteins. *Science (New York, N.Y.)* 347, 551–5.
- (174) Jaremko, M., Jaremko, ?, Giller, K., Becker, S., and Zweckstetter, M. (2015). Structural Integrity of the A147T Polymorph of Mammalian TSPO. *ChemBioChem* 16, 1483–1489.
- (175) Li, F., Liu, J., Zheng, Y., Garavito, R. M., and Ferguson-Miller, S. (2015). Crystal structures of translocator protein (TSPO) and mutant mimic of a human polymorphism. *Science* 347, 555–558.
- (176) Lu, M., and Fu, D. (2007). Structure of the zinc transporter YiiP. *Science (New York, N.Y.)* 317, 1746–8.
- (177) Kefala, G., Ahn, C., Krupa, M., Esquivies, L., Maslennikov, I., Kwiatkowski, W., and Choe, S. (2010). Structures of the OmpF porin crystallized in the presence of foscholine-12. *Protein Science* 19, 1117–1125.
- (178) Son, S.-Y., Ma, J., Kondou, Y., Yoshimura, M., Yamashita, E., and Tsukihara, T. (2008). Structure of human monoamine oxidase A at 2.2-Å resolution: the control of opening the entry for substrates/inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 105, 5739–5744.
- (179) Warne, T., Serrano-Vega, M. J., Tate, C. G., and Schertler, G. F. X. (2009). Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expression and Purification* 65, 204–213.
- (180) Liew, L. S. Y., Lee, M. Y., Wong, Y. L., Cheng, J., Li, Q., and Kang, C. (2016). Selection of suitable detergents for obtaining an active dengue protease in its natural form from *E. coli*. *Protein Expression and Purification* 121, 141–148.
- (181) Huang, Q., Li, Q., Joy, J., Chen, A. S., Ruiz-Carrillo, D., Hill, J., Lescar, J., and Kang, C. (2013). Lyso-myristoyl phosphatidylcholine micelles sustain the activity of Dengue non-structural (NS) protein 3 protease domain fused with the full-length NS2B. *Protein Expression and Purification* 92, 156–162.
- (182) Ellinger, P., Kluth, M., Stindt, J., Smits, S. H. J., and Schmitt, L. (2013). Detergent Screening and Purification of the Human Liver ABC Transporters BSEP (ABCB11) and MDR3 (ABCB4) Expressed in the Yeast *Pichia pastoris*. *PLoS ONE* 8, 1–12.

- (183) Constantinescu, S. N., Keren, T., Socolovsky, M., Nam, H., Henis, Y. I., and Lodish, H. F. (2001). Ligand-independent oligomerization of cell-surface erythropoietin receptor is mediated by the transmembrane domain. *Proceedings of the National Academy of Sciences of the United States of America* 98, 4379–4384.
- (184) Li, Q., Wong, Y. L., Huang, Q., and Kang, C. (2014). Structural insight into the transmembrane domain and the juxtamembrane region of the erythropoietin receptor in micelles. *Biophysical Journal* 107, 2325–2336.
- (185) Herrmann, M., Danielczak, B., Textor, M., Klement, J., and Keller, S. (2015). Modulating bilayer mechanical properties to promote the coupled folding and insertion of an integral membrane protein. *European Biophysics Journal* 44, 503–512.
- (186) Wang, X., Huang, H., Sun, C., and Huang, F. (2015). Structural analysis of bacteriorhodopsin solubilized by lipid-like phosphocholine biosurfactants with varying micelle concentrations. *Journal of Colloid and Interface Science* 437, 170–180.
- (187) Kjaergaard, M. (2015). Can proteins be intrinsically disordered inside a membrane? *Intrinsically Disordered Proteins* 3, 1–7.
- (188) Palmieri, F. (2013). The mitochondrial transporter family SLC25: Identification, properties and physiopathology. *Molecular Aspects of Medicine* 34, 465–484.
- (189) Mifsud, J., Ravaud, S., Krammer, E.-M., Chipot, C., Kunji, E. R. S., Pebay-Peyroula, E., and Dehez, F. (2013). The substrate specificity of the human ADP/ATP carrier AAC1. *Molecular membrane biology* 30, 160–8.
- (190) Ruprecht, J. J., Hellowell, A. M., Harding, M., Crichton, P. G., McCoy, A. J., Kunji, Edmund R S, and Kunji, E. R. S. (2014). Structures of yeast mitochondrial ADP/ATP carriers support a domain-based alternating-access transport mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 111, E426–34.
- (191) Weidemann, M. J., Erdelt, H., and Klingenberg, M. (1970). Adenine nucleotide translocation of mitochondria. Identification of carrier sites. *Eur.J.Biochem.* 16, 313–335.
- (192) Nury, H., Dahout-Gonzalez, C., Trézéguet, V., Lauquin, G. J. M., Brandolin, G., and Pebay-Peyroula, E. (2006). Relations between structure and function of the mitochondrial ADP/ATP carrier. *Annual review of biochemistry* 75, 713–741.
- (193) Lauquin, G. J., Duplaa, a. M., Klein, G., Rousseau, a., and Vignais, P. V. (1976). Isobongkrekic acid, a new inhibitor of mitochondrial ADP-ATP transport: radioactive labeling and chemical and biological properties. *Biochemistry* 15, 2323–7.
- (194) King, M. S., Kerr, M., Crichton, P. G., Springett, R., and Kunji, E. R. S. (2016). Formation of a cytoplasmic salt bridge network in the matrix state is a fundamental step in the transport mechanism of the mitochondrial ADP/ATP carrier. *Biochimica et Biophysica Acta - Bioenergetics* 1857, 14–22.
- (195) Block, M. R., and Vignais, P. V. (1986). Dependence of the conformational state of the isolated adenine nucleotide carrier protein on the detergent used for solubilization. *Biochemistry* 25, 374–9.
- (196) Scherer, B., and Klingenberg, M. (1974). Demonstration of the relationship between the adenine nucleotide carrier and the structural changes of mitochondria as induced by adenosine 5'-diphosphate. *Biochemistry* 13, 161–170.
- (197) Vignais, P. V., Vignais, P. M., and Defaye, G. (1973). Adenosine diphosphate translocation in mitochondria. Nature of the receptor site for carboxyatractyloside (gummiferin). *Biochemistry* 12, 1508–1519.
- (198) Klingenberg, M. (2008). The ADP and ATP transport in mitochondria and its carrier. *Biochimica et Biophysica Acta - Biomembranes* 1778, 1978–2021.

- (199) Riccio, P., Burkhard, S., and Klingenberg, M. (1973). ISOLATION OF A NEW ATRACTYLOSIDE TYPE COMPOUND. *FEBS Letters* 31, 11–14.
- (200) Brüscheweiler, S., Yang, Q., Run, C., and Chou, J. J. (2015). Substrate-modulated ADP / ATP-transporter dynamics revealed by NMR relaxation dispersion. *Nature Publishing Group* 22, 636–641.
- (201) Bojanovski, D., Schlimme, E., Wang, C.-S., and Alaupovic, P. (1976). Studies on the Adenine Nucleotide Translocase from Rat Liver Mitochondria. *Eur.J.Biochem.* 71, 539–548.
- (202) Kramer, R., and Klingenberg, M. (1977). Reconstitution of Inhibitor Binding Properties of the Isolated Adenosine 5'-Diphosphate, Adenosine 5'-Triphosphate Carrier-Linked Binding Protein. *Biochemistry* 16, 4954–4961.
- (203) Robinson, A. J., and Kunji, E. R. S. (2006). Mitochondrial carriers in the cytoplasmic state have a common substrate binding site. *Proceedings of the National Academy of Sciences of the United States of America* 103, 2617–22.
- (204) Gropp, T., Brustovetsky, N., Klingenberg, M., Muller, V., Fendler, K., and Bamberg, E. (1999). Kinetics of electrogenic transport by the ADP/ATP carrier. *Biophys J* 77, 714–726.
- (205) Kraemer, R., and Klingenberg, M. (1980). Modulation of the reconstituted adenine nucleotide exchange by membrane potential. *Biochemistry* 19, 556–560.
- (206) Brandolin, G., Doussiere, J., Gulik, A., Gulik-Krzywicki, T., Lauquin, G., and Vignais, P. (1980). Kinetic, binding and ultrastructural properties of the beef heart adenine nucleotide carrier protein after incorporation into phospholipid vesicles. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 592, 592–614.
- (207) Beyer, K., and Klingenberg, M. (1985). ADP/ATP carrier protein from beef heart mitochondria has high amounts of tightly bound cardiolipin, as revealed by <sup>31</sup>P nuclear magnetic resonance. *Biochemistry* 24, 3821–3826.
- (208) Schlame, M., Beyer, K., Hayer-Hartl, M., and Klingenberg, M. (1991). Molecular species of cardiolipin in relation to other mitochondrial phospholipids. Is there an acyl specificity of the interaction between cardiolipin and the ADP/ATP carrier? *European journal of biochemistry / FEBS* 199, 459–466.
- (209) Jiang, F., Ryan, M. T., Schlame, M., Zhao, M., Gu, Z., Klingenberg, M., Pfanner, N., and Greenberg, M. L. (2000). Absence of cardiolipin in the *crd1* null mutant results in decreased mitochondrial membrane potential and reduced mitochondrial function. *Journal of Biological Chemistry* 275, 22387–22394.
- (210) Heimpel, S., Basset, G., Odoj, S., and Klingenberg, M. (2001). Expression of the mitochondrial ADP/ATP carrier in *Escherichia coli*. Renaturation, reconstitution, and the effect of mutations on 10 positive residues. *Journal of Biological Chemistry* 276, 11499–11506.
- (211) Kramer, R., Klingenberg, M., and Biochemie, P. (1980). ENHANCEMENT OF RECONSTITUTED ADP,ATP EXCHANGE ACTIVITY BY PHOSPHATIDYLETHANOLINE AND BY ANIONIC PHOSPHOLIPIDS R. KRAMER and M. KLINGENBERG. 119, 257–260.
- (212) Palmieri, F., Pierri, C. L., De Grassi, A., Nunes-Nesi, A., and Fernie, A. R. (Apr. 2011). Evolution, structure and function of mitochondrial carriers: a review with new insights. *The Plant journal : for cell and molecular biology* 66, 161–81.
- (213) Nury, H., Dahout-Gonzalez, C., Trézéguet, V., Lauquin, G., Brandolin, G., and Pebay-Peyroula, E. (2005). Structural basis for lipid-mediated interactions between mitochondrial ADP/ATP carrier monomers. *FEBS Letters* 579, 6031–6036.
- (214) Kunji, E. R. S., Aleksandrova, A., King, M. S., Majd, H., Ashton, V. L., Cerson, E., Springett, R., Kibalchenko, M., Tavoulari, S., Crichton, P. G., and Ruprecht, J. J. (2016). The transport mechanism of the mitochondrial ADP/ATP carrier. *Biochimica et biophysica acta* 1863, 2379–2393.

- (215) Faham, S., Yang, D., Bare, E., Yohannan, S., Whitelegge, J. P., and Bowie, J. U. (2004). Side-chain contributions to membrane protein structure and stability. *Journal of Molecular Biology* 335, 297–305.
- (216) Kunji, E. R. S., and Robinson, A. J. (2006). The conserved substrate binding site of mitochondrial carriers. *Biochimica et biophysica acta* 1757, 1237–48.
- (217) Britto, P. J., Knipling, L., and Wolff, J. (2002). The local electrostatic environment determines cysteine reactivity of tubulin. *Journal of Biological Chemistry* 277, 29018–29027.
- (218) Dehez, F., Pebay-Peyroula, E., and Chipot, C. (2008). Binding of ADP in the mitochondrial ADP/ATP carrier is driven by an electrostatic funnel. *Journal of the American Chemical Society* 130, 12725–12733.
- (219) Krammer, E. M., Ravaud, S., Dehez, F., Frelet-Barrand, A., Pebay-Peyroula, E., and Chipot, C. (2009). High-chloride concentrations abolish the binding of adenine nucleotides in the mitochondrial ADP/ATP carrier family. *Biophysical Journal* 97, 25–27.
- (220) Kunji, E. R. S., and Crichton, P. G. (2010). Mitochondrial carriers function as monomers. *Biochimica et Biophysica Acta - Bioenergetics* 1797, 817–831.
- (221) Palmieri, F., and Pierri, C. L. (2010). Structure and function of mitochondrial carriers - Role of the transmembrane helix P and G residues in the gating and transport mechanism. *FEBS Letters* 584, 1931–1939.
- (222) Fiermonte, G., Dolce, V., Santorelli, F. M., Palmieri, F., John, E., David, L., Dionisi-vici, C., and Walker, J. E. (2003). Metabolism and Bioenergetics : The Mitochondrial Ornithine Transporter : BACTERIAL EXPRESSION , CHARACTERIZATION , AND TISSUE DISTRIBUTION OF TWO HUMAN., DOI: 10.1074/jbc.M302317200.
- (223) Indiveri, C., and Palmieri, L. (1994). Biochimica et Biophysica Acta. *Transport* 1188, 293–301.
- (224) Monné, M., Miniero, D. V., Daddabbo, L., Palmieri, L., Porcelli, V., and Palmieri, F. (2015). Mitochondrial transporters for ornithine and related amino acids: A review. *Amino Acids* 47, 1763–1777.
- (225) Run, C., Yang, Q., Liu, Z., Ouyang, B., and Chou, J. J. (2015). Molecular Basis of MgATP Selectivity of the Mitochondrial SCaMC Carrier. *Structure* 23, 1394–1403.
- (226) Fiermonte, G., De Leonadis, F., Todisco, S., Palmieri, L., Lasorsa, F. M., and Palmieri, F. (2004). Identification of the mitochondrial ATP-Mg/Pi transporter: Bacterial expression, reconstitution, functional characterization, and tissue distribution. *Journal of Biological Chemistry* 279, 30722–30730.
- (227) Harborne, S. P., Ruprecht, J. J., and Kunji, E. R. S. (2015). Calcium-induced conformational changes of the regulatory domain of human mitochondrial aspartate/glutamate carriers. *Biochimica et Biophysica Acta (BBA)* 1847, 1245–1253.
- (228) Sounier, R., Bellot, G., and Chou, J. J. (2015). Mapping Conformational Heterogeneity of Mitochondrial Nucleotide Transporter in Uninhibited States \*\*, DOI: 10.1002/ange.201408417.
- (229) Berardi, M. J., and Chou, J. J. (2014). Fatty acid flippase activity of UCP2 is essential for its proton transport in mitochondria. *Cell Metabolism* 20, 541–552.
- (230) Orekhov, V. Y., Ibraghimov, I., and Billeter, M. (2003). Optimizing resolution in multidimensional NMR by three-way decomposition. *Journal of Biomolecular NMR* 27, 165–173.
- (231) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005). The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins: Structure, Function and Genetics* 59, 687–696.

- (232) Favier, A., and Brutscher, B. (2011). Recovering lost magnetization: Polarization enhancement in biomolecular NMR. *Journal of Biomolecular NMR* 49, 9–15.
- (233) Sekhar, A., Rumfeldt, J. A. O., Broom, H. R., Doyle, C. M., Bouvignies, G., Meiering, E. M., and Kay, L. E. (2015). Thermal fluctuations of immature SOD1 lead to separate folding and misfolding pathways. *eLife* 4, 1–33.
- (234) Korzhnev, D. M., Kloiber, K., and Kay, L. E. (2004). Multiple-Quantum Relaxation Dispersion NMR Spectroscopy Probing Millisecond Time-Scale Dynamics in Proteins : Theory and Application Multiple-Quantum Relaxation Dispersion NMR Spectroscopy Probing Millisecond Time-Scale Dynamics in Proteins : Theory and A., 7320–7329.
- (235) Lee, Y., Willers, C., Kunji, E. R. S., and Crichton, P. G. (2015). Uncoupling protein 1 binds one nucleotide per monomer and is stabilized by tightly bound cardiolipin. *112*, DOI: 10.1073/pnas.1503833112.
- (236) Crichton, P. G., Lee, Y., Ruprecht, J. J., Cerson, E., Thangaratnarajah, C., King, M. S., and Kunji, E. R. S. (2015). Trends in Thermostability Provide Information on the Nature of Substrate , Inhibitor , and Lipid Interactions with Mitochondrial Carriers \*. *290*, 8206–8217.
- (237) Tribolet, R., and Sigel, H. (1988). Influence of the protonation degree on the self-association properties of adenosine 5'-triphosphate (ATP). *European Journal of Biochemistry / FEBS* 170, 617–26.
- (238) Olofsson, G., and Sparr, E. (2013). Ionization Constants pKa of Cardiolipin. *PLoS ONE* 8, DOI: 10.1371/journal.pone.0073040.
- (239) Luciani, S., and Varotto, R. (1975). Difference between atractyloside and carboxyatractyloside on the binding to the mitochondrial membrane. *FEBS Letters* 56, 194–197.
- (240) Matsuura, H., Shimotakahara, S., Sakuma, C., Tashiro, M., Shindo, H., Mochizuki, K., Yamagishi, A., Kojima, M., and Takahashi, K. (2004). Thermal unfolding of ribonuclease T1 studied by multidimensional NMR spectroscopy. *Biological Chemistry* 385, 1157–1164.
- (241) Berjanskii, M. V., and Wishart, D. S. (2005). A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. *Journal of the American Chemical Society* 127, 14970–14971.
- (242) Lakomek, N. A., Ying, J., and Bax, A. (2012). Measurement of <sup>15</sup>N relaxation rates in perdeuterated proteins by TROSY-based methods. *Journal of Biomolecular NMR* 53, 209–221.
- (243) Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology* 157, 105–132.
- (244) Ziani, W., Maillard, A. P., Petit-Hrtlein, I., Garnier, N., Crouzy, S., Girard, E., and Cov?s, J. (2014). The X-ray structure of NccX from *Cupriavidus metallidurans* 31A illustrates potential dangers of detergent solubilization when generating and interpreting crystal structures of membrane proteins. *Journal of Biological Chemistry* 289, 31160–31172.



## 6 Manuscript: Visualizing the impact of the crystalline environment on slow protein dynamics

### 6.1 Preface

In this chapter we investigate how crystalline proteins move. In particular, we are interested in slow conformational exchange processes, and study whether such processes are retained in the crystal, and possibly how the crystalline environment impacts such motions. The methods we use here are related to the methods used in the mitochondrial carriers chapter: we use relaxation dispersion methods as a primary experimental tool. However, all experimental results presented here were obtained from solid-state NMR spectroscopy, rather than solution-state NMR. This is the method of choice for studying dynamics in crystals and other non-solubilized proteins. Furthermore, instead of using CPMG relaxation dispersion, which consists of a series of refocusing pulses, we are using here R1rho relaxation dispersion, i.e. we monitor the relaxation of  $^{15}\text{N}$  spins under a continuous RF field. This approach, which has been introduced in solid-state NMR only recently by the host group, provides access to microsecond motions, as outlined below.

As a complementary approach, we use long molecular dynamics simulations of the explicit crystal lattice. These MD simulations have been performed by our collaborators, Nikolai Skrynnikov, Sergei A. Izmailov and Olga N. Rogacheva.

This chapter represents a paper draft which is being finalized for submission currently.

This manuscript was written by Paul Schanda and Nikolai Skrynnikov.

Vilius Kurauskas<sup>1,2,3,\*</sup>, Sergei A. Izmailov<sup>5,\*</sup>, Olga N. Rogacheva<sup>5,\*</sup>, Audrey Hessel<sup>1,2,3</sup>, Isabel Ayala<sup>1,2,3</sup>, Joyce Woodhouse<sup>1,2,3</sup>, Yi Xue<sup>4</sup>, Tairan Yuwen<sup>5</sup>, Oleg Mikhailovskii<sup>5</sup>, Nicolas Coquelle<sup>1,2,3</sup>, Jacques-Philippe Colletier<sup>1,2,3</sup>, Nikolai R. Skrynnikov<sup>4,5,#</sup>, Paul Schanda<sup>1,2,3,#</sup>

1 Université Grenoble Alpes, IBS, F-38044 Grenoble, France.

2 CEA, Institut de Biologie Structurale, F-38044 Grenoble, France.

3 CNRS, Institut de Biologie Structurale, F-38044 Grenoble, France.

4 Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA.

5 Laboratory of Biomolecular NMR, St. Petersburg State University, St. Petersburg 199034, Russia.

\* These authors contributed equally to this work.

Correspondence and requests for materials should be addressed to N.R.S. (email: nikolai@purdue.edu) or P.S. (email: paul.schanda@ibs.fr).

## 6.2 Abstract

Protein function is rooted in the ability of proteins to exchange between different conformational states, that may be involved in different stages of a functional process. Spectroscopic methods, such as solution-state NMR, are able to provide information about dynamics, but they do not provide a direct structural picture. X-ray crystallography can provide atomic coordinates of the conformations that co-exist in the crystal, providing thus powerful views into protein dynamics, but of course it is bound to studying proteins in densely packed 3D lattices. If and how the crystalline packing may distort motions remains poorly understood. Here we combine high-resolution solid-state NMR relaxation-dispersion measurements, including NEar-Rotary-resonance Relaxation Dispersion (NERRD) that sense angular motion, with multi-microsecond-long molecular dynamics simulations to probe internal and overall dynamics on long ( $\mu$ s) time scales. These time scales are often thought to be of particular functional relevance. We study different crystal forms of the same protein, ubiquitin and find that the packing in different crystalline arrangements impacts the equilibrium of exchanging states, i.e. the thermodynamics, as well as the interconversion rate of the states. MD simulations provide mechanistic insights and allow rationalizing how the conformational dynamics are altered by the lattice contacts. We find evidence for overall motion in one of the crystal forms, and provide a time scale of this motion; strikingly, this overall motion appears to be coupled to local dynamics, via the alternating formation of inter-molecular interactions. Our study highlights the importance of considering the packing when analyzing dynamics of crystalline proteins.

### 6.3 Introduction

Proteins perform their functions in aqueous solution at ambient temperatures by sampling a multitude of conformational states involved, for example, in enzymatic turnover or binding to different partners. Understanding protein function requires characterizing the three-dimensional structures, ideally of all these thermally accessible conformations, and decipher how the different conformations interconvert. The structural bases for understanding protein function have been laid primarily through X-ray diffraction (XRD) of proteins embedded in crystals, with over one-hundred thousand structures determined over the last 30 years. Moreover, recent methodological advances in XRD aim at characterizing also the heterogeneity of states that a protein can adopt, by fitting multiple conformers – rather than a single structure – into the electron density obtained from Bragg diffraction peaks, or by additionally making use of diffuse scattering intensity [1–6]. The direct structural insight into protein dynamics that one may obtain from such ensemble crystallographic approaches may shed light on molecular mechanisms of functional processes; in particular, when combined with techniques that provide access to kinetics and thermodynamics, such as solution-NMR spectroscopy [7], one may obtain a fairly complete spatio-temporal insight into dynamics in terms of structures, kinetics and thermodynamics, which may ultimately lead to a detailed understanding of protein function.

The extraction of dynamic information – in the sense of conformational heterogeneity – from X-ray diffraction data relies on the assumption that the packing of molecules in the crystal lattice does not alter structure and dynamics. When considering the (lowest-energy) structure - i.e. the single deposited X-ray structure - this assumption appears fairly well justified: structures obtained from crystalline proteins are generally in excellent agreement with state-of-the-art solution-NMR structures determined, and are generally as good in predicting solution-NMR observables as the NMR structures themselves [8,9]. Whether dynamics, i.e. the fine details of co-existing structures, are also conserved in crystals is less well studied.

Evidence that the dynamics required for function are retained in crystals, at least in some cases, comes from the observation of functional activity, such as ligand binding and enzymatic turnover, in crystals [10–14]. Direct experimental investigations of the dynamics of proteins in crystals at the atomic level have turned out to be difficult. Early studies focused on comparing XRD temperature factors (B-factors) with solution-NMR order parameters ( $S^2$ ) [15–17]. Such comparisons only had limited success and remained ambiguous, because of the very different nature and information content of these observables (see ref. 18 and references therein). Recent developments of magic-angle spinning solid-state NMR (MAS ssNMR) on proteins have allowed a more direct comparison of NMR parameters in solution and crystals [19–21]. Furthermore, comparative molecular dynamics (MD) simulations of proteins embedded in an explicit crystal lattice and in solution also allowed studying differences in solution and crystals, at least over the time scales accessible to MD (typically sub-microseconds) [22–25]. The picture arising from these ssNMR and MD studies suggests that the crystal lattice has indeed only little impact on dynamics on sub-microsecond time scales: changes are primarily found in loops engaged in crystal contacts, whereas secondary structure elements and loops remote from neigh-

bors showed similar order parameters. Many biological processes, however, occur on longer time scales, often micro-to-milliseconds ( $\mu\text{s}$ - $\text{ms}$ ), and they may involve rare excursions to low-populated "minor states". In many cases, such rare conformations seem to be crucial for functional activity [26]. Whether such slow dynamic processes are sustained in crystals, and how the packing alters these motions is much less studied, because MD simulation methods only very recently reached those time scales, and experimental NMR methods were unavailable until recently.

Here we use a combination of ssNMR experimental methods with microsecond-long MD simulations of explicit crystal lattices, in order to investigate how the crystalline environment impacts slow protein dynamics. We compare different crystal forms of the 8.6 kDa regulatory protein ubiquitin, and rigorously apply the same ssNMR/MD approaches to these different crystal forms. This allows us to exclude any experimental bias that might arise when comparing results from different techniques, such as solution- and solid-state methods.

Ubiquitin dynamics in solution have been extensively characterized experimentally and computationally. In particular, microsecond motion has been shown for the loop connecting ubiquitin's  $\beta_4$  strand and the  $3_{10}$  helix, comprising residues 50 to 58, which exist in two distinct states that differ by a  $\approx$ half-turn of the D52/G53 peptide plane orientation [27–32]. These two states correspond to so-called type-I and type-II  $\beta$ -turns, respectively, herein referred to as  $\beta\text{I}$  and  $\beta\text{II}$  states. The two states differ furthermore in the orientation of E24 side chain, which is H-bonded to this loop in only one of the two states (Figure 6.1a). The dynamic exchange process has been proposed to result in small (within 1 Å) but discernible reshaping of the entire molecule [30], and perturbation of the  $\beta$ -turn equilibrium has been shown to allosterically alter the binding of proteins to a surface located at ubiquitin's opposite side [30]. The dynamic process is thus thought to be of considerable functional relevance for this highly ubiquitous protein. Different crystal structures are available that are in either of these two states, and millisecond-long MD simulations in solution have detected this microsecond dynamic process [32]. Previous ssNMR data also revealed that this  $\mu\text{s}$  dynamics is present in a crystal form. In addition to this internal motional process, we have recently provided evidence that ubiquitin molecules in the crystal undergo overall "rocking" motion of a few degrees amplitude, and that this motion differs in different crystals.

We combine ssNMR relaxation dispersion experiments, that probe chemical-shift fluctuations as well as bond-vector fluctuations, with  $\mu\text{s}$ -long MD simulations. The combined results establish that crystal contacts indeed alter the kinetics (exchange rates) and thermodynamics (relative populations) of  $\mu\text{s}$  motions, as evidenced here with the well-defined  $\beta\text{I}$ -to- $\beta\text{II}$  dynamic process. We identify the exact molecular origins of these altered dynamics, by localizing inter-molecular contacts to neighboring molecules that modulate the energy landscape. Furthermore, we provide quantification of the motional time scale of this rocking, which turns out to be similar to the one of  $\beta$ -turn motion, and provide evidence for a coupling of overall and local motion, resulting in an extended inter-/intra-molecular motional network. Our results also allow to rationalize very recently determined room-temperature X-ray structures of the different crystal forms of ubiquitin.

## 6.4 Results

### 6.4.1 Type-I $\leftrightarrow$ type-II $\beta$ -turn exchange is slowed down by crystal contacts

In this study we investigate different crystal forms of ubiquitin which are referred to as (i) MPD-ub (ii) cubic-PEG-ub, which are obtained in the presence of two different precipitation agents (MPD and PEG, respectively, see Methods). X-ray diffraction based crystal structures of these different forms, collected at cryo-temperatures (100 K) have been obtained previously [33–36]. While these X-ray data reveal that ubiquitin adopts essentially the same overall structure (backbone RMSD of 0.88 Å), they differ in the conformation of the loop region encompassing residues E51 to T55: in MPD-ub this region forms a so-called type-II  $\beta$ -turn, while it is in  $\beta$ I conformation in cubic-PEG-ub (Figure 6.1a). The predominant state in solution is  $\beta$ I [30] with an estimated >95% population, in exchange with an alternate state which is believed to be  $\beta$ II [32]. In order to probe whether these states are in dynamic exchange at room temperature, and to obtain kinetic and thermodynamic information we employed ssNMR  $^{15}\text{N}$  relaxation dispersion (RD) experiments [37]. These experiments measure the relaxation rate,  $R_{1\rho}$ , of each of the  $^{15}\text{N}$  spins across the protein in the presence of a spin-lock radio-frequency (RF) field as a function of its field-strength,  $\nu_{\text{RF}}$ . The dependence of the  $R_{1\rho}$  relaxation rate constant on  $\nu_{\text{RF}}$ , also called "relaxation dispersion", provides insight into  $\mu\text{s}$  exchange dynamics, probed at the level of each residue. In MAS ssNMR, two different physical mechanisms can give rise to (non-flat) RD profiles:

(i) If the exchanging states differ in their local electronic environment around the  $^{15}\text{N}$  nucleus, and if this process occurs on a  $\mu\text{s}$ -ms time scale, the corresponding chemical-shift fluctuation contributes to the  $R_{1\rho}$  relaxation rate constant. This enhancement of  $R_{1\rho}$  depends on the strength of the RF field [38]. This effect, known as "Bloch-McConnell RD" in solution-state NMR [39], depends on the isotropic chemical-shift difference, as well as the time scale of the exchange and the relative populations of the exchanging states. A similar type of RD has been extensively used in previous chapters in this thesis.

(ii) Fluctuation of the N-H bond orientation relative to the reference frame of the crystallite also induces enhanced  $R_{1\rho}$  relaxation, in addition to the above effect that fluctuation of the isotropic chemical shift has [40–43]. If such motions occur on  $\mu\text{s}$ -ms time scales, the  $R_{1\rho}$  rate constants depend on the RF-field:  $R_{1\rho}$  increases when approaching the so-called Rotary-Resonance condition, i.e., the RF field strength that equals the sample-spinning frequency. We refer to this type of relaxation dispersion here as NEar-Rotary-resonance Relaxation-Dispersion (NERRD). Non-flat NERRD profiles are an unambiguous signature of  $\mu\text{s}$  motions, and their precise shape depends on the time scale of motion, the relative populations of the involved states, as well as the angle by which the NH bond vector differs between the different states. NERRD thus provides insight into the geometric extent of the motion, and complements Bloch-McConnell RD data, which probe motions by their isotropic chemical-shift fluctuation. Analytical equations to describe this effect are available [44], which are

approximately correct [40,45]; instead we employ here exact numerical simulations to calculate NERRD profiles and fit data (see Methods). Importantly, in solution-state NMR this type of information is unavailable, because the rapid (ns) molecular tumbling averages the anisotropic interactions such that they are not informative of slow dynamics.

Both of the above mechanisms simultaneously lead to RF-field dependent  $R_{1\rho}$  rate profiles whenever  $\mu$ s-ms motions are present. The effects on  $R_{1\rho}$  are manifest in different ranges of applied RF field strengths. In the low RF-field range (below  $\approx 10$  kHz, which we term here the "Bloch-McConnell relaxation dispersion", BMCRD), the isotropic chemical-shift fluctuation dominates the dispersion of  $R_{1\rho}$  rate constants. In the vicinity of the rotary-resonance conditions, which in our study is at 44 kHz and thus well separated from the Bloch-McConnell regime, it is the angular fluctuation of the NH bond that dominates the RD profiles. Throughout this work we will use the term BMCRD when referring to the low-RF-field range ( $< 15$  kHz), and the term NERRD for the range approaching the rotary resonance, i.e. from  $\approx 20$  to 44 kHz, even though these ranges are not fundamentally different but rather are reflecting two mechanisms of microsecond motion. In practice, these data were measured independently and at different MAS frequencies. The two ranges can be fitted jointly or independently, and such fits can in ideal cases provide exchange rates, population levels of involved states, chemical-shift differences and jump angles. In cases where the process is fast (thousands per second rate constant), which is the case here, it is not possible to disentangle populations and chemical-shift changes, and only their product can be obtained (see below).

Figure 6.1b shows Bloch-McConnell RD data for three representative amide sites in MPD-ub and cubic-PEG-ub. The non-flat RD profiles for residues located in the  $\beta$ -turn region and the adjacent N-terminal part of the  $\alpha$ -helix unambiguously show that this part of the molecule undergoes  $\mu$ s-ms dynamics in both crystals. A joint fit of a two-state exchange model to data from the five residues located in this region is shown as solid lines (see Methods). The use of a two-state exchange model appears justified in the light of the two distinct states observed in different crystal structures and MD simulation (ref. 46 and MD data below). The exchange time constant obtained from these fits is  $\tau_{\text{ex}} = 90 \pm 25 \mu\text{s}$  for MPD-ub and  $125 \pm 45 \mu\text{s}$  for cubic-PEG-ub. In the case of cubic-PEG-ub, we detect sizable (i.e. non-flat) BMCRD also for residues outside the  $\beta$ -turn. We will discuss the potential origin of these dispersions further below.

In order to gain insight into the angular fluctuation of the NH bond vectors, we measured NERRD profiles. Figure 6.1d shows the largest NERRD effects observed in MPD-ub, which are found for the two amide sites D52 and R54, which are adjacent to the peptide plane D52/G53 that undergoes a  $\approx$ half-turn flip when exchanging between  $\beta$ I and  $\beta$ II conformations. (The amide signal of G53 is not resolved in the spectra of MPD-ub.) Other residues do not show significantly rising NERRD profiles, as exemplified by the data of three residues depicted in black in Figure 6.2. An independent experiment at a different MAS frequency (20 kHz) also exhibited the largest NERRD effect for residues D52 and R54 in MPD-ub [37], whereas all other

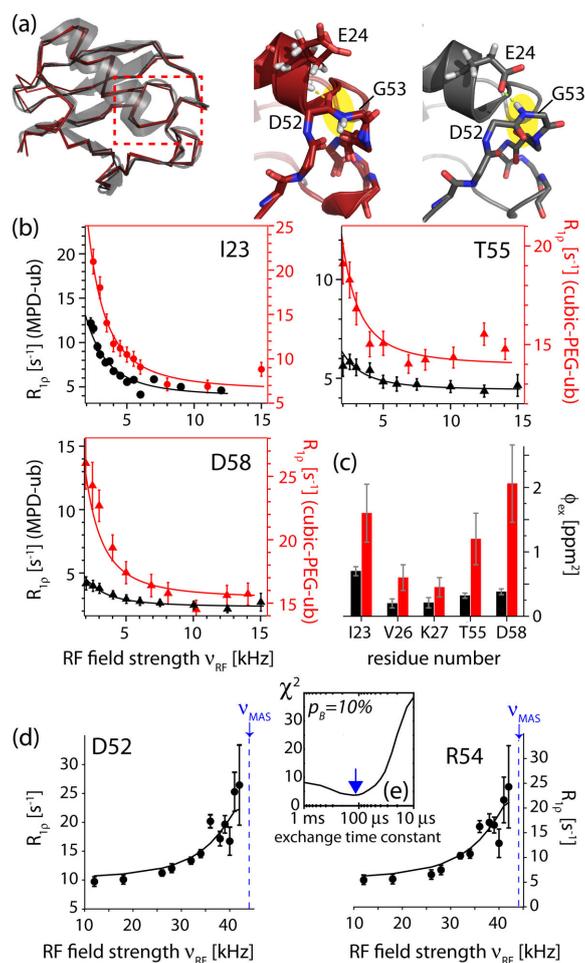


Figure 6.1: Solid-state NMR relaxation dispersion data of MPD-ub (black) and cubic-PEG-ub (red) crystals. (a) Superposition of the main chains in the two crystals, shown as  $C\alpha$  backbone traces; for cubic-PEG-ub both chains of the unit cell are shown. For clarity, the semi-transparent cartoon representation is shown only for MPD-ub. (b) Bloch-McConnell  $R_{1\rho}$  RD profiles of  $^{15}\text{N}$  sites in MPD-ub, along with a common two-site exchange model of data from residues 23, 26, 27, 55, 58. Data shown as circles and triangles were obtained at these two  $B_0$  field strengths, respectively. Note that the y-axis range is shifted for data from the two crystals, facilitating the comparison, but the span is the same, allowing absolute comparison of the size of the RDs. The offset of the cubic-PEG-ub  $R_{1\rho}$  rate constants is explained by the rocking-contribution to  $R_{1\rho}$  in cubic-PEG-ub, whereas the different "amplitudes" of the RD profiles points to differences in populations (main text). (c) Fitted  $\phi_{ex}$  values, ( $\phi_{ex} = p_A \cdot p_B \cdot \Delta\omega^2$ ) from a fit of residues I23, V26, K27, T55 and D58. The higher values of  $\phi_{ex}$  in cubic-PEG-MPD, on average by 3.3 fold, points to a higher population of the respective minor state in cubic-PEG-ub compared to MPD-ub. Estimations of the population, based on these ratios are shown in Table 6.S1. Fitted exchange time constants are reported in Table 6.S3. (d) NERRD dispersion data of the two amide sites in MPD-ub that are adjacent to the flipping peptide plane D52/G53, along with fit curves from a two-site jump model. The best-fit jump time constant is of the order of 110  $\mu$ s (indicated by an arrow in the chi-square plot of panel (e)).

Table 6.S3: Populations and time scales of conformational exchange dynamics in the region encompassing residues D52-D58 and the adjacent N-terminal part of the  $\alpha$ -helix. The population of the  $\beta$ I state is  $\beta$ I=100%-p( $\beta$ II). When fitting the MPD-ub and cubic-PEG-ub data, the same set of data from five residues was used (I23, V26, K27, T55, D58); the exact choice of residues does not significantly alter the outcome (see Figure 6.S2). In the case of cubic-PEG-ub we observe two sets of signals for several residues, corresponding presumably to the two non-equivalent chains, but for the majority of residues the signals overlap. For this reason, assigning the data to one chain or the other is not possible. † No transitions were detected over the entire 2  $\mu$ s trajectory of 24 molecules, i.e. all chains remained in  $\beta$ I. # The ratio of  $\phi_{\text{ex}}$  values from BMCRD fits were used to estimate population levels in MPD-ub, based on assumed populations of 20-50% for cubic-PEG-ub. See text and Figure 2d,e for a details. ‡ Solution-state NMR data are obtained from Smith et al [31]. Note that the exchange rate constant is largely in agreement with other studies, which, however, were only able to provide upper limit estimates, and at lower temperature [28,29].

	MD		NMR	
	$\beta$ II population	time constant ( $\tau_{\text{ex}}$ )	$\beta$ II population	time constant ( $\tau_{\text{ex}}$ )
solution	4%	2.0 $\mu$ s	3%‡	6 $\mu$ s‡
MPD-ub	90%	3.7 $\mu$ s	$\approx$ 88-97%	90 $\pm$ 25 $\mu$ s
Cubic-PEG-ub chain A	9%	0.8 $\mu$ s	2* $\approx$ 20-50%	2*125 $\pm$ 45 $\mu$ s
chain B	$\approx$ 43%	2.1 $\mu$ s		
Rod-PEG-ub chains A, B, C	2*0***	2* $\dagger$	2*n.d.	2*n.d.

residues showed either small NERRD profiles (residues I23, E51, T55) or no significant RF-field dependency of the  $R_{1\rho}$  rate constants, mirroring the present data.

We next wanted to obtain quantitative information about time scale and amplitude of motion from these NERRD measurements. To this end, we implemented a numerical fit procedure, which consists of performing a grid of spin-dynamics simulations, and comparing the resulting NERRD profiles to the experimentally observed ones. Figure 6.1d shows the best-fit curves for a joint fit of D52 and R54, and the corresponding  $\chi^2$  profile (Figure 6.1e) reveals that the time scale of this motion is approximately 100  $\mu$ s. This is in excellent agreement with the independent Bloch-McConnell RD data of Figure 6.1b. In addition to confirming the time scale of the process, NERRD data provide information about the fluctuation angle of the NH motion of D52 and R54, which is found to be of the order of 7 to 11 degrees, depending on the assumed populations (see Figure 6.S1). Interestingly, this angle is similar to the difference between the orientations that these two bonds have in  $\beta$ I and  $\beta$ II states, which is 12 and 14 $^\circ$  for the two residues (comparing PDB entries 3ONS and 1UBI). Taken together, both Bloch-McConnell RD and NERRD data show that the  $\beta$ -turn dynamics in MPD-ub occurs on a time scale of 90 $\pm$ 25  $\mu$ s; this exchange process generates chemical-shift fluctuations in the  $\beta$ -turn region (Figure 6.1b), and angular fluctuations of residues D52 and R54, as revealed by NERRD (Figure 6.1d). Bloch-McConnell RD fits of cubic-PEG-ub point to an exchange time scale of 125 $\pm$ 45  $\mu$ s. (All exchange parameters are listed in Table 6.S3.)

It is interesting to note that the two crystals, which have very different crystal neighbors around the exchanging region, have similar exchange kinetics, but that these kinetics are very different to the case of solution, where the exchange process has a

time constant of  $\tau_{\text{ex}}=6 \pm 2 \mu\text{s}$ , i.e. about 15 times faster [30]. This significant slowdown of the process in the crystals is not due to the increased viscosity induced by the precipitation agent [47]. As noted above, the "ground state" is different between the two crystals, being  $\beta\text{I}$  in solution and cubic-PEG-ub, and  $\beta\text{II}$  in MPD-ub. This shows that the nature of the ground state does not dictate the kinetics, because in solution and cubic-PEG-ub, which have both  $\beta\text{I}$  as major state, the kinetics are very different.

#### 6.4.2 The relative populations of $\beta\text{I}/\beta\text{II}$ states differ in the different crystals

Inspection of the RD data in Figure 6.1b reveals a noticeable difference between the two crystals: the amplitude of the Bloch-McConnell RD profiles for residues in cubic-PEG-ub is considerably larger than those of their counterparts in MPD-ub; as we will show here, this is due to differences in the relative populations of exchanging states. Bloch-McConnell RD profiles depend, in addition to exchange kinetics, on populations and the chemical-shift difference of given nucleus in the exchanging states. These two quantities cannot be disentangled in the present "fast" exchange regime, and only the parameter  $\phi_{\text{ex}}$  can be obtained, which is the product of the (squared) chemical-shift change between the states,  $\Delta\omega$ , and the populations of the two involved states,  $\phi_{\text{ex}}=p_{\beta\text{I}} \cdot p_{\beta\text{II}} \cdot |\Delta\omega|^2$ . Attempts to lift this degeneracy of solutions by simultaneously fitting NERRD data were unsuccessful (data not shown), and are additionally complicated by the fact that for cubic-PEG-ub the NERRD data report on an additional process (described further below). Therefore, only the residue-wise  $\phi_{\text{ex}}$  values can be obtained [38], and they are represented in Figure 6.1c for five residues the data of which are available in both crystal forms. Residue-wise  $\phi_{\text{ex}}$  values are significantly larger in cubic-PEG-ub than in MPD-ub, on average by a factor  $3.3 \pm 1.0$ . Larger  $\phi_{\text{ex}}$  values are due to a higher population of "minor" state in cubic-PEG-ub, and/or to larger chemical-shift differences  $\Delta\omega$  between the exchanging states. The chemical-shift differences are not a priori known, such that it is not possible to conclude with certainty about the populations in the two crystal forms. Several arguments convince us that these  $\phi_{\text{ex}}$  values indeed show that the minor-state population is larger in cubic-PEG-ub than in MPD-ub: (i) the two states in exchange are in both cases similar ( $\beta\text{I}$ ,  $\beta\text{II}$ ). Therefore, one can safely assume that, at least when considering the average over several residues, the  $\Delta\omega$  values are similar in the two crystal forms; (ii) calculated chemical-shift differences, derived from MD simulations of the two crystal forms, discussed below, are indeed similar. If we thus assume that the larger  $\phi_{\text{ex}}$  in cubic-PEG-ub derives from a larger minor-state population, we can use the ratio of  $\phi_{\text{ex}}$  values in the two crystals to estimate the respective population levels: given that the minor state in cubic-PEG-ub is, by definition, populated to  $<50\%$ , and given the averaged ratio of  $\phi_{\text{ex}}$  in the two crystal forms, the estimated upper limit of the minor-state population of MPD-ub is  $\approx 15\%$ . When assuming smaller minor-state populations for cubic-PEG-ub (20-50%), the resulting range of minor-state populations in MPD-ub is 3-12% (see Table 6.S1).

In order to obtain complementary insights and mechanistic details, we turned to all-atom molecular dynamics simulations. MD simulations using explicit crystal lat-

tices have previously been shown to successfully reproduce experimental ssNMR and XRD data, and they may contribute powerful information about mechanistic details that are difficult to obtain experimentally [6,36,48–50]. We simulated an explicit arrangement of ubiquitin molecules according to the crystal lattices of MPD-ub and cubic-PEG-ub, built of 24 and 48 molecules, respectively, over 2  $\mu$ s. The explicit simulation of a large number of molecules improves the sampling statistics considerably such that we effectively have multi-microsecond-long observations of dynamics in the different crystals. In addition, we performed an 8  $\mu$ s simulation of ubiquitin in solution, which we find to be in excellent agreement with a recent 1 ms trajectory [32]. Along all these trajectories we observed multiple  $\beta$ I  $\leftrightarrow$   $\beta$ II transitions, enabling us to extract the relative populations of the two conformations, which are reported in Table 6.S3. The minor-state populations are qualitatively in good agreement with the data derived from Bloch-McConnell RD data, i.e. MPD-ub has  $\approx$ 4-fold lower minor-state population than cubic-PEG-ub (chain B), mirroring the lower  $\phi_{\text{ex}}$  values. Experimentally the signals from the two chains overlap in most cases, making it impossible to distinguish the two chains. The fact that our experimental data reflect better the behavior found by MD for chain B (in terms of populations) suggests that our BMCRD profiles are dominated by the larger dispersion profiles of residues chain B. Alternatively, it may also be that the experimental averaging of BMCRD profiles over the two chains actually leads to an underestimation, and that the dispersions of the "pure" chain B would be larger than what we observe.

Furthermore, we have also simulated a third crystal form, rod-PEG-ub, employing a crystal lattice composed of 24 molecules, and a trajectory length of 2  $\mu$ s. Remarkably, we did not observe any transition from the  $\beta$ I to the  $\beta$ II conformer along this trajectory. Although we cannot exclude that the process is too infrequent to be detected within our limited time window, this finding suggests that rod-PEG-ub only exists in the  $\beta$ I conformation. This prediction, as well as the predictions made above concerning MPD-ub and cubic-PEG-ub, are confirmed by recent room-temperature X-ray structures (see Discussion section).

It is noteworthy that the time scale of the exchange process in the MD simulations tends to be faster than observed experimentally (see Table 6.S3). This finding suggests that the force field employed in this study does not accurately reflect the transition-state free-energy barrier. Similar findings have been reported before [51]. This observation is perhaps not surprising, because force fields were parameterized primarily against experimental equilibrium data, and not against kinetic data. Accordingly, MD simulations tend to reproduce more correctly the populations of exchanging states than the kinetic rate constants.

Taken together, NMR and MD techniques independently establish that the relative populations of  $\beta$ I and  $\beta$ II conformations differs widely in different crystals and in solution. The  $\beta$ I :  $\beta$ II populations span essentially the full range from 10:90 (MPD-ub) to 57:43/91:9 (cubic-PEG-ub), 90:2 (solution) and  $\approx$ 100:0 (rod-PEG-ub). The extracted population ratios can be converted into free energy differences between the  $\beta$ 1 and  $\beta$ 2 forms that range from  $dG \approx -2.5$  kcal/mol in solution to  $\approx +1.5$  kcal/mol in MPD-ub.

### 6.4.3 Molecules in cubic-PEG-ub undergo tens-of- $\mu$ s overall rocking motion

We also recorded NERRD experiments on cubic-PEG-ub crystals, in order to obtain insight into the amplitude of  $\mu$ s dynamics. It shall be reminded that in MPD-ub only few residues (in particular D52 and R54; see Fig. 6.1d) show increased  $R_{1\rho}$  close to the rotary resonance condition (i.e. non-flat NERRD profiles), and that the time scale obtained from a fit of these data is in excellent agreement with Bloch-McConnell RD data. In contrast to this finding, essentially all residues in cubic-PEG-ub exhibit non-flat NERRD profiles, exemplified in Figure 6.2. Given that this trend is observed for all residues, we ascribe the underlying motion to a global process, involving the entire molecule, such as an overall "rocking" motion of the molecules in the crystal. We have proposed the existence of such motions in cubic-PEG-ub recently [36], based on the previous observation at a single RF-field strength (15 kHz) of elevated  $R_{1\rho}$  rate constants across the entire cubic-PEG-ub protein compared to MPD-ub [36]. These previous single-RF-field ssNMR  $R_{1\rho}$  measurements, however, did not allow determining the amplitude, and the time scale could be determined only roughly (tens of ns to about hundred  $\mu$ s). In order to obtain quantitative insight into the underlying global motional process, we have performed a common fit of a dynamic model with a global time scale to NERRD data from 22 well-resolved residues in cubic-PEG-ub, excluding residues in the  $\beta$ -turn region, as for those residues one expects superposition of local and global motion. The best-fit time scale is of the order of tens of microseconds (Figure 6.2b). Residue-wise fluctuation angles, expressed in a two-site jump model, are in the range 4-7° (Figure 6.S3). Such a relatively small amplitude is expected for an overall-rocking motion within the confines of the lattice; MD simulations revealed an amplitude of overall motion in cubic-PEG-ub of  $\approx 8^\circ$  [36].

It is interesting to note that the tens-of-microseconds overall-motion time scale of overall motion, found experimentally, is similar to the time scale of the  $\beta$ -turn motion in cubic-PEG-ub ( $125 \pm 45 \mu$ s). This similarity may be fortuitous, but it might also point to some correlation between the two processes, local and overall motion. Indeed, microsecond-long MD simulations, discussed in the next section point to such a correlation of internal and overall motion.

The fact that essentially all residues exhibit non-flat NERRD profiles hampers the study of the  $\beta$ -turn exchange process using this technique, because residues in the  $\beta$ -turn region sense both the overall and the local motion, making it difficult to disentangle the two processes from NERRD data.

Finally we note that in cubic-PEG-ub crystals we detect non-flat BMCRD dispersions also for residues outside the "canonical"  $\beta$ -turn region, i.e. we see more residues with evidence for  $\mu$ s motion than in MPD-ub. Different reasons may be identified for such additional chemical-shift fluctuations. (i) It has been recently reported by solution-state NMR and MD studies that the  $\beta$ -turn exchange process leads to rearrangements across the molecule; consequently, small but detectably non-flat solution-NMR RD profile were seen for residues across the structure, namely in  $\beta 1$ ,  $\beta 2$ , the  $\alpha$ -helix,  $\beta 3$ ,  $\beta 5$  (as well as the canonical  $\beta$ -turn region). One of the reasons for the detection of

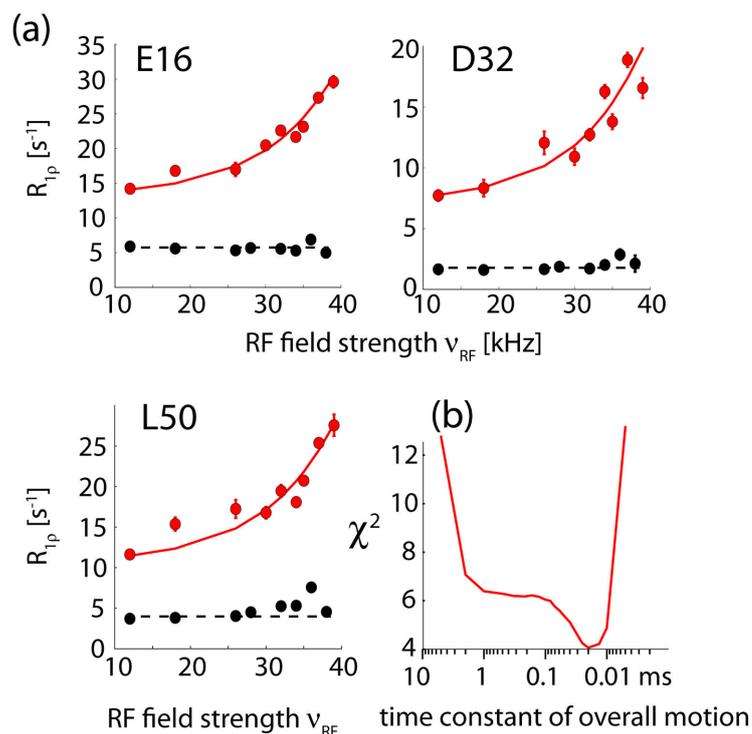


Figure 6.2: (a) NERRD data of representative amide sites outside the canonical  $\beta$ -turn region shown for MPD-ub (black) and cubic-PEG-ub (red). The solid red lines correspond to a two-state fit of data from 22 amide sites with a common exchange time constant and residue-wise fluctuation angles (all shown in Figure S\_PEGNERRD). (b) shows the  $\chi^2$  value of this fit as a function of the jump time constant.

non-flat BMCRD for more residues in PEG-ub than in MPD-ub may therefore be due to the higher population of minor state in cubic-PEG-ub: it might be that from a concerted process, involving large parts of the molecule, only the residues with largest chemical-shift change (i.e. those in the  $\beta$ -turn region) are detectable, while more residues show detectable non-flat BMCRD profiles in cubic-PEG-ub due to the higher population. (ii) Rocking motion, revealed by the NERRD data (Figure 6.2) may also induce chemical-shift fluctuations. (iii) There may also be other motional modes, either intra-molecular or inter-molecular, which are absent in MPD-ub; we identify at least one such instance, H68, discussed further below.

#### 6.4.4 Identifying intermolecular contacts that alter the conformational equilibrium

As noted above,  $\beta$ I conformation in ubiquitin represents a "default" state, which is found in solution and in the majority of crystal structures [27,32,52,53]. At the same time the  $\beta$ II conformation has also been observed in a number of crystals, such as MPD-ub [33]. This leads us to conclude that the crystalline environment can act as a switch converting the protein from  $\beta$ I to  $\beta$ II. Our MD simulations of ubiquitin crystals are well suited to identify intermolecular interactions that are responsible for this transformation.

In our initial analysis of the MPD-ub trajectory we have identified two such putative intermolecular contacts. One is the four-way junction involving a pair of D52 side chains and a pair of K63 side chains, all of which belong to different ubiquitin molecules but come together in the crystal lattice (Figure 6.3a). The other is the intermolecular hydrogen bond between the (neutral) carboxyl groups of E64 and E24, which appears to stabilize the coupling between E24 and G53 (Figure 6.3b). According to the MD simulation data, both of these interactions show clear preference for the  $\beta$ II over the  $\beta$ I form.

However, additional simulations convinced us that this is the case of correlation rather than causation. Specifically, we have recorded two separate 2- $\mu$ s-long trajectories of MPD-ub crystals, one of which carried K63A mutation and the other E64A mutation. As it turned out, these trajectories produced exactly the same proportion of  $\beta$ I to  $\beta$ II species as the wild-type crystal simulation. This led us to conclude that the intermolecular contacts involving K63 and E64 side chains are not the key to the  $\beta$ -turn conformation. At the same time we have found that a E24A mutation produced a complete reversal of  $\beta$ II to  $\beta$ I ratio compared to the wild-type simulation, from 90%:10% to 9%:91%. In other words, E24A mutation has restored the "default" state of the system, which is dominated by  $\beta$ I. Therefore it must be the hydrogen bond between the carboxyl group of E24 and the backbone amide group in G53 which is directly responsible for prevalence of the  $\beta$ II form in the wt MPD-ub crystal. But this hydrogen bond is the intramolecular hydrogen bond – so how is it related to crystal packing?

Inspection of the trajectories of ubiquitin in solution and in the cubic-PEG-ub crystal revealed that in these simulations the side chain of E24 is mostly immersed in solvent. In solution, E24 side chain is solvated 64% of the entire simulation time. In the cubic-PEG-ub, it is solvated 60% of the time for chain A and 51% of the time for chain B. In contrast, it is solvated at the level of only 8% in the MPD-ub simulation (see Table 6.S4). As it turns out, the packing of ubiquitin molecules in the MPD-ub crystal lattice is such that E24 side chain does not have the space to extend outwards. This is illustrated in Figures 6.3c and d, which compare the volume of space available to E24 side chain in MPD-ub and in cubic-PEG-ub crystals. In particular, the intrinsically favorable rotameric state ( $\chi_1=180^\circ$ ,  $\chi_2=180^\circ$ ) which is prominent in other simulations, is completely obstructed in MPD-ub. We conclude that typically the E24 side chain enjoys the entropic benefit of being immersed in solvent, but in the MPD-ub it is denied this opportunity and therefore opts for the hydrogen bond with G53.

The propensity of the carboxylic group in E24 to form a hydrogen bond with the G53 amide group is also influenced by its protonation state. The MPD-ub crystal has been obtained from the mother liquor at pH 4.2. Under these conditions the carboxylic group of E24, as positioned in the crystallographic structure, is expected to be mostly protonated [54]. To elucidate the role of this factor, we have altered E24's status from neutral to cationic and then recorded an additional 1- $\mu$ s-long simulation of the MPD-ub crystal. As it turns out, in this trajectory E24 is partially recruited into salt bridge with K27 (the presence of this salt bridge, 21%, is in line with what is found in other crystals, see Table 6.S4). This happens at the expense

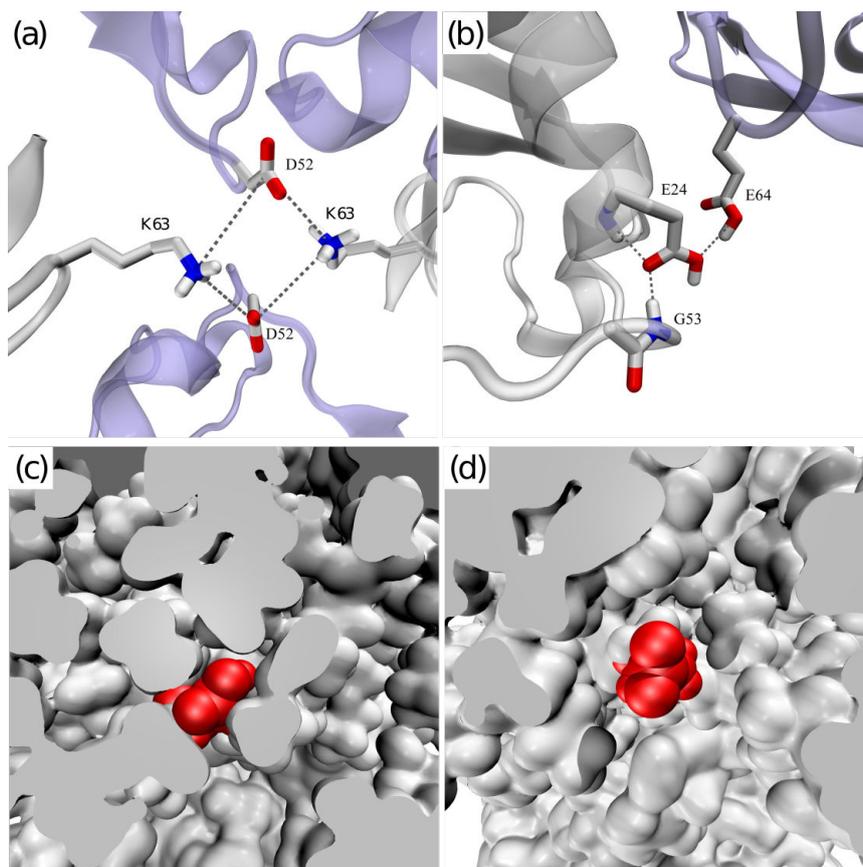


Figure 6.3: The surrounding of E24 side chain in different crystal forms of ubiquitin. (a) A rare frame from the MPD-ub trajectory where E24 side chain (painted red) is projected toward solvent. (b) A typical frame from cubic-PEG-ub trajectory where E24 from ubiquitin molecule chain A is immersed into solvent.

of the E24-G53 hydrogen bond, which drops to the level of 36%. Accordingly, the ratio of  $\beta$ II to  $\beta$ I in this trajectory changes from 90%:10% to 75%:25%. This finding demonstrates that the protonation state of E24 also plays a significant role in the stabilization of the  $\beta$ II species in MPD-ub crystal, albeit not as important as the steric confinement factor.

#### 6.4.5 Coupling between rocking and $\beta$ I $\leftrightarrow$ $\beta$ II exchange

As discussed above, the experimental data suggest that rocking motion in cubic-PEG-ub occurs on a time scale similar to the  $\beta$ I  $\leftrightarrow$   $\beta$ II conformational exchange. (MD simulations also suggest that rocking occurs on the same time scale as  $\beta$ I  $\leftrightarrow$   $\beta$ II exchange, although this time scale is considerably shorter than observed experimentally.) Is this a mere coincidence, or does it point toward a certain common underlying mechanism? Here we would like to argue that the latter is true.

The inspection of cubic-PEG-ub trajectory led us to identify a potential coupling mechanism. This mechanism involves an ion pair between K11 (chain A) and D52

Table 6.S4: Important interactions involving E24 side-chain carboxylic group as characterized by their presence in the MD simulations. The contacts involving K27 and G53 are intramolecular. The numbers in each column do not always add to 100% because E24 also forms certain other contacts (both inter- and intramolecular)

2*E24 side chain contacts	2*Solution	2*MPD-ub		2*cubic-PEG-ub	
		E24 neutral	E24 charged	chain A	chain B
Solvent	64%	8%	4%	60%	51%
E24-K27	36%	3%	21%	26%	27%
E24-G53	-	76%	36%	2%	3%

(chain B). In the crystal structure these two side chains are positioned far apart (Figure 6.4a). However, during the MD simulation they frequently form a salt bridge (Figure 6.4b). The formation of this salt bridge is associated with (i) slight reorientation of protein molecules in the crystal lattice and (ii)  $\beta\text{I} \rightarrow \beta\text{II}$  transition mediated by residue D52 in the  $\beta$  turn. Thus the reorientational dynamics of the protein (rocking motion) turns out to be correlated with conformational exchange. This scenario is illustrated in the Supplementary Movie 1. (see movies at: [https://www.dropbox.com/sh/e7asu846fyu9eva/AAC1yo\\_8Mv5LJHfrSY7c2hwUa?dl=0](https://www.dropbox.com/sh/e7asu846fyu9eva/AAC1yo_8Mv5LJHfrSY7c2hwUa?dl=0)) To obtain some insight into this putative mechanism, we turn to the trajectory of cubic-PEG-ub. The intermolecular  $\text{K11}_{\text{chainA}}\text{-D52}_{\text{chainB}}$  salt bridge is found in 26% of all frames in this trajectory; this salt bridge favors  $\beta\text{II}$  conformation over  $\beta\text{I}$  by a factor 4.1. However, the simulation does not allow one to directly observe a correlation between  $\text{K11}_{\text{chainA}}\text{-D52}_{\text{chainB}}$  salt bridge and rocking motion (see Figure 6.S4). Indeed, there is a complex dynamic network of intermolecular interactions in the crystal, involving several dozen hydrogen bonds and salt bridges (see Figure 6.S5). Generally speaking, all of these interactions have a role in rocking dynamics, which inevitably masks the effect of  $\text{K11}_{\text{chainA}}\text{-D52}_{\text{chainB}}$ . Consequently, the correlation between  $\text{K11}_{\text{chainA}}\text{-D52}_{\text{chainB}}$  and rocking motion can only be observed in a statistical sense.

To obtain further insight into the role of the  $\text{K11}_{\text{chainA}}\text{-D52}_{\text{chainB}}$  interaction in rocking, we have additionally recorded long MD trajectories of the cubic-PEG crystals of ubiquitin containing K11A mutation. This mutation abrogates the salt bridge between  $\text{K11}_{\text{chainA}}$  and  $\text{D52}_{\text{chainB}}$ . As a result, the proportion of  $\beta\text{II}$  conformation drops and the amplitude of rocking motion becomes appreciably lower (see Table 6.S5). The latter effect is visualized in Figure 6.5, which shows reorientational correlation functions indicative of rocking dynamics [36] the correlation function  $\text{grock}(\tau)$  of WT cubic-PEG-ub declines significantly, with an order parameter for this overall molecular reorientation of 0.95 (chain B) and 0.98 (chain A) [36]. In the case of K11A, the orientational correlation function does not decline as significantly as the corresponding function from the WT simulation, indicating reduced rocking in this mutant. It is noteworthy that elimination of the intermolecular salt bridge stabilizes the crystal (rather than destabilizing it, which would be a trivial outcome). In addition, we have also simulated cubic-PEG-ub ubiquitin containing G53A mutation. This mutation locks the  $\beta$ -turn in the  $\beta\text{I}$  conformation (inserting a non-glycine residue in this position favors  $\beta\text{I}$  form over  $\beta\text{II}$  [55]). The loss of  $\beta\text{II}$  conformation results in

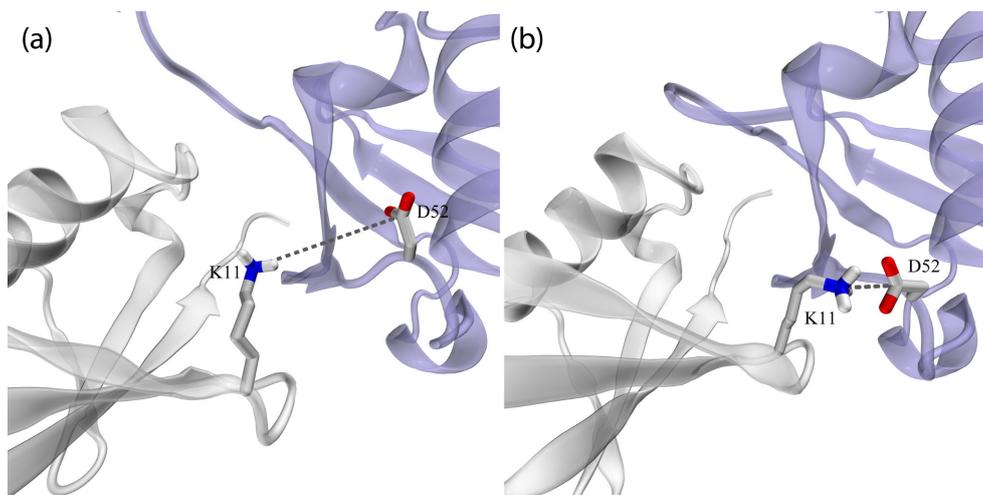


Figure 6.4: Ion pair  $K11_{\text{chainA}} - D52_{\text{chainB}}$  in the crystallographic structure 3N30 [34] (left panel) and in the selected frame from cubic-PEG-ub trajectory (right panel). The dynamic disorder in residue  $K11_{\text{chainA}}$  is reflected in high crystallographic temperature factors; this residue's side chain has not been modeled in our recent structure 4XOL.

the decreased presence of  $K11_{\text{chainA}}-D52_{\text{chainB}}$  salt bridge and small but detectable reduction in rocking (see Table 6.S5 and Figure 6.5).

In summary, the series of mutant simulations suggests that conformational exchange

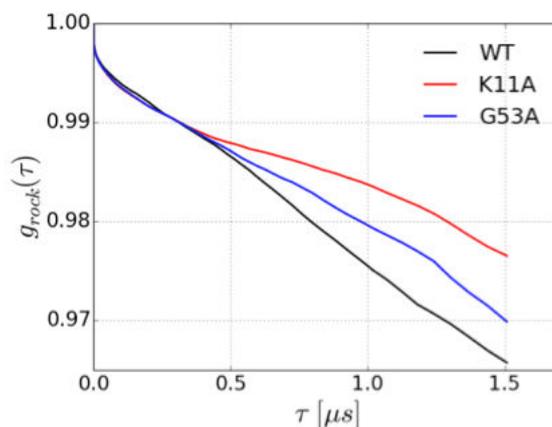


Figure 6.5: Orientational rigid-body correlation functions representative of the rocking motion in the trajectories of cubic-PEG-ub crystals of wild-type ubiquitin, as well as K11A and G53A mutants (see ref. [36] for definition).

in the  $\beta$  turn can indeed be coupled to rocking; the coupling is mediated by the intermolecular salt bridge  $K11_{\text{chainA}}-D52_{\text{chainB}}$ .

One can also envisage a more general scenario whereby conformational exchange is coupled to rocking. Recently, Smith et al. have found that  $\beta\text{I} \leftrightarrow \beta\text{II}$  transitions cause small (within 1 Å) but discernible reshaping of the entire ubiquitin molecule [30]. It is reasonable to suggest that the change in shape leads to "repacking" of ubiquitin molecules in the crystal lattice, entailing slight translation and overall reorientation

Table 6.S5: MD simulations of cubic-PEG-ub crystals containing wild-type ubiquitin and two mutants

	chain	$\beta$ II, %	D52 <sub>chainB</sub> -K11 <sub>chainA</sub> , %	rocking amplitude	$\tau$ ( $\beta$ I $\leftrightarrow$ $\beta$ II), $\mu$ s
WT	B	42	26	5.1°	0.4
K11A	B	18	0	4.1°	0.4
G53A	B	0	14	4.7°	6.9
WT	A	9	0	4.5°	1.1
K11A	A	8	0	4.2°	1
G53A	A	0	0	4.2°	1.8

of the molecules (i.e. causing rocking motion). Conversely, rocking motion leads to subtle changes in the pattern of crystal contacts which in turn triggers  $\beta$ I  $\leftrightarrow$   $\beta$ II transitions (in particular, via allosteric effects [30]). In general, it should not be surprising that a molecule that is conformationally labile produces a poor crystal which suffers, inter alia, from rocking dynamics. The idea that rocking is coupled to conformational dynamics is highly intuitive; when this scenario happens, the two processes are expected to occur on the same time scale.

#### 6.4.6 Probing relaxation dispersion effects

MD simulations on a microsecond time scale can be used to model Bloch-McConnell relaxation dispersion effects. For this purpose MD coordinates are fed, frame by frame, into chemical shift prediction programs such as SHIFTX2 [56]. In a situation when the calculated chemical shift  $\delta$  shows a pattern of modulation on the  $\mu$ s time scale, one may expect to experimentally detect RD effects at the respective site [57]. Ultimately, this approach makes it possible to uncover the dynamic mechanism behind the observable RD profiles. We have used this method to probe the backbone  $^{15}$ N and  $^1$ H $^N$ N sites in MPD-ub and cubic-PEG-ub trajectories. In MPD-ub, relaxation dispersion effects have been predicted for sites in the  $\beta$ -turn 51-54 experiencing  $\beta$ I  $\leftrightarrow$   $\beta$ II exchange, as well as the sites at the N-terminal end of the first  $\alpha$ -helix that are hydrogen-bonded to this  $\beta$  turn. One example of such behavior, residue R54 from MPD-ub, is illustrated in the Supplementary Movie S2 ([https://www.dropbox.com/sh/e7asu846fyu9eva/AAC1yo\\_8Mv5LJHfrSY7c2hwUa?dl=0](https://www.dropbox.com/sh/e7asu846fyu9eva/AAC1yo_8Mv5LJHfrSY7c2hwUa?dl=0)), which illustrates that the  $\beta$ I  $\leftrightarrow$   $\beta$ II transition is accompanied by a  $^{15}$ N chemical-shift change of R54.

Similarly, the analysis of the cubic-PEG-ub trajectory suggests that Bloch-McConnell dispersions can occur in the  $\beta$  turn and its coupled sites (particularly, in the chain B molecules). In addition, we have also found several sites where the origin of predicted RD effects is different. One example of such distinctive behavior is residue H68. Focusing on one specific ubiquitin molecule in cubic-PEG-ub trajectory, we observe H68 side chain engaged in hydrogen bond with Q40 side chain from a neighboring molecule (Fig. 6.6a). Later during the simulation H68 side chain swings sideways and forms hydrogen bond with side chain E64 which belongs to another neighboring molecule (Fig. 6.6b). After ca. 0.8  $\mu$ s it makes a reverse transition, reestablishing hydrogen bond with Q40, and remains in this conformation until the end of the tra-

jectory.

The rotameric transitions of H68 (from  $\chi_1 = -60^\circ$  to  $180^\circ$  and back) cause modulation

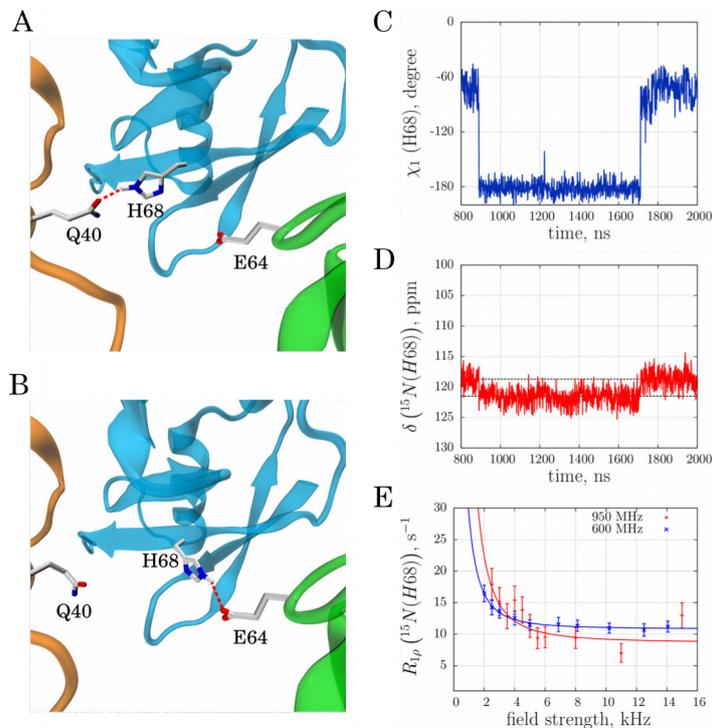


Figure 6.6:  $\mu\text{s}$  time scale dynamics of residue H68 (chain A) in cubic-PEG-ub crystal. (a,b) Intermolecular hydrogen bonds alternatively formed between H68 and Q40 or E64 as seen in the MD simulation of cubic-PEG-ub. (c,d) Time variation of side-chain torsional angle  $\chi_1(t)$  and main-chain  $^{15}\text{N}$  chemical shift  $\delta(t)$  for residue H68 in the selected ubiquitin molecule from the simulated unit crystal cell. (e) Experimental relaxation dispersion data for residue H68 (chain unassigned) in the cubic-PEG-ub crystal; also shown are the results of fitting using fast exchange model (continuous lines). The motional time scale according to the experimental data,  $56 \mu\text{s}$ , is one-to-two orders of magnitude slower than suggested by the MD simulations (the same is true for other manifestations of  $\mu\text{s}$  dynamics in ubiquitin crystals). The animated version of this graph is available as Supplementary Movie S3.

of the backbone chemical shifts in this residue [56]. For the  $^{15}\text{N}$  spin the amplitude of modulation is calculated to be 2.1 ppm (corresponding to the gap between the two dashed horizontal lines in Fig. 6.6d). This is a sizeable effect, which can be reliably reproduced by chemical shift prediction software. Therefore the computational analysis points toward possible RD effect at the amide site in residue H68. This prediction is borne out by the experimental data, as illustrated in Fig. 6.6e.

Of interest, there is a strong possibility that conformational exchange in H68 is influenced by rocking motion. Indeed, slight reorientation of the central protein molecule relative to its neighbors should favor either H68-Q40 or H68-E64 interaction depending on the direction of reorientational fluctuation (see Figures 6.6a,b). Thus, H68 is an example of how RD effects can be linked to the fluctuating pattern of intermolecular contacts and thereby to the rocking motion. In this connection it is noteworthy that RD effects have been experimentally detected at this site in the cubic-PEG-ub

crystal form which is affected by rocking (Figure 6.6e), but not in the MPD-ub crystal which is immune to rocking (not shown). Finally, H68 offers another example of how rocking motion can potentially be coupled to local conformational dynamics (cf. previous section).

## 6.5 Discussion

The advent of novel experimental and analytical tools to decipher the heterogeneity of structures co-existing in the crystal has sparked increased interest in molecular motions in protein crystals. Understanding what governs motions and how the packing may alter motions is of central practical importance. This study provides converging insights from experimental and simulation sources into functionally relevant local dynamics, and overall motions of protein molecules in the crystal lattice. We provide evidence that both local and overall motion are influenced by crystal packing. Concerning the local  $\mu$ s process involving the  $\beta$ I  $\leftrightarrow$   $\beta$ II transition in ubiquitin, we identified the interactions that can shift the equilibrium all the way from the  $\beta$ I state (100:0 in rod-PEG-ub, 98:2 in solution) to a nearly 1:1 coexistence of  $\beta$ I: $\beta$ II states in cubic-PEG-ub, and all the way to a predominantly  $\beta$ II state (10:90 in MPD-ub). As compared to solution, the crystalline environment may thus strongly bias the equilibrium: it can make minor states essentially disappear, such that detection of dynamics with crystal-based methods inevitably fails in these cases, or it can make such states become the major conformer. These findings need to be considered when interpreting dynamic ensembles obtained from crystals in terms of their implications for the behavior in solution.

We have used here a quantitative analysis of NERRD ssNMR data to probe  $\mu$ s motions through the bond vector fluctuation. NERRD analyses have confirmed the time scale of the local  $\beta$ -turn motion in MPD-ub. They have furthermore provided quantitative insight into overall rocking motion in cubic-PEG-ub, which has a significant amplitude ( $\approx$ 4-7 degrees), and occurs on a tens-of-microseconds time scale, similar to the  $\mu$ s local motion. It is remarkable that it is the crystal that has the most abundant local motion, i.e. nearly equal population of the states in the  $\beta$ -turn region and, for example, fluctuating interaction partners of H68, that also has the largest-scale overall motion. Intuitively, it is not surprising that local rearrangements that reshape the molecule will ultimately result in overall motion, and, conversely, that overall motion enables the alternating formation of inter-molecular contacts, which in turn stabilize different local conformational states in an alternating manner. We have highlighted the intermolecular K11-D52 salt bridge and alternating H-bonding of H68, as two prominent routes how such a coupling of local and overall motion can be realized.

Interestingly, the findings and predictions deriving from the NMR and MD results are confirmed by room-temperature crystal structures deposited very recently in the Protein Data Bank. In these structures, alternate conformations were required to improve the fits of the electron density. In particular, the  $\beta$ -turn region has the expected "ground state" ( $\beta$ I in cubic-PEG-ub and rod-PEG-ub,  $\beta$ II in MPD-ub), and alternate conformations are found for residues in the  $\beta$ -turn in MPD-ub (E24, D52,

G53) and cubic-PEG-ub (E51, D52, G53, R54 in chain A and E51, R54 in chain B; a table of all residues modeled with alternate states is provided in Table 6.S2). Furthermore, rod-PEG-ub, which diffracts at highest resolution (1.8 Å) shows no evidence for alternate conformations anywhere in the protein. In the light of the high resolution of this data set this finding strongly suggests that ubiquitin in this crystal packing has the highest structural homogeneity. This finding mirrors the MD simulation results that show that along our multi-microsecond-long trajectory no  $\beta$ -turn conversion occurs (cf Table 6.S3). Interestingly, the different rocking dynamics is also reflected in these room-temperature XRD data: cubic-PEG-ub, which this study revealed to feature the largest structural heterogeneity and largest rocking motion, diffracts at significantly lower resolution (3 Å) than the more homogeneous structures that do not undergo sizeable overall motion. These findings highlight that at least at room temperature the protein motions in the crystal govern the XRD data quality. Our study stresses the importance of integrating the crystalline environment when interpreting XRD-derived dynamics informations, and provides the basis for understanding how different XRD diffraction quality of the same protein can be rooted in local and overall dynamics.

## 6.6 Methods

### 6.6.1 Sample preparation

Recombinant human ubiquitin (without any affinity tag) was overexpressed in *E. coli*, and purified using ion exchange and size exclusion chromatography. Samples used for solid-state NMR were uniformly labeled with  $^2\text{H}$  and  $^{15}\text{N}$  isotopes, resulting in >98% deuteration at non-exchangeable hydrogen sites. In order to have minimal perturbation from remote protons to the relaxation of amide  $^{15}\text{N}$ , the protein was dissolved in  $\text{H}_2\text{O}:\text{D}_2\text{O}$  (20:80) mixture at pH 7, resulting in isotope incorporation of  $^1\text{H}/^2\text{H}$  at exchangeable sites at a 20:80 ratio, and the protein was then lyophilized. Solutions used for crystallization employed the same  $\text{H}_2\text{O}:\text{D}_2\text{O}$  ratio. All three crystal forms used in this study have been obtained using crystallization buffers described previously [36]. Microcrystals were filled into a 1.6 mm Agilent or 1.3 mm Bruker rotor; caps were glued with epoxy glue to avoid dehydration.

### 6.6.2 Solid-state NMR and data analysis

Experiments were performed on (i) a 600 MHz (14.1 T) Agilent VNMRS spectrometer (only BMCRD experiment) equipped with a 1.6 mm triple-resonance MAS probe tuned to  $^1\text{H}, ^{13}\text{C}, ^{15}\text{N}$  or (ii) a 600 MHz or (iii) a 950 MHz (22.3 T) Bruker Avance 2 spectrometer; the latter two were equipped with 1.3 mm MAS probes tuned to  $^1\text{H}, ^{13}\text{C}, ^{15}\text{N}$  frequencies with an auxiliary  $^2\text{H}$  coil. Resonance assignments of MPD-ub and cubic-PEG-ub have been reported elsewhere [36,58]. BMCRD experiments at 600 MHz were performed at 39.5 kHz MAS, while BMCRD data at 950 MHz were collected at 50 kHz MAS. NERRD data on both MPD-ub and cubic-PEG-ub were recorded on the 600 MHz (Bruker) spectrometer at a MAS frequency of 44.053 kHz. In order to enhance resolution, the BMCRD data on cubic-PEG-ub were recorded

as a series of proton-detected 3D hCONH-based experiments. This enabled us to resolve the otherwise overlapped I23 resonance. All other data were collected with 2D proton-detected hNH correlation experiments. In all experiments, cross-polarization steps were used for transfer, with approximately 85 kHz  $^1\text{H}$  RF field, and a  $^{15}\text{N}$  RF field adjusted to the  $n=1$  Hartmann-Hahn condition. Peak volumes in the individual 2D (3D) spectra were measured in NMRView (OneMoon Scientific), and relaxation rate constants were fitted as monoexponential functions, using in-house written python fitting routines. Error bars were estimated from Monte Carlo simulations, based on three times the spectral noise level.

Fits of a two-state exchange model to the BMCRD data were fitted with the program relax, version 4.0 [59]. We have investigated how the choice of residues, and the choice of static magnetic field strengths influences the outcome of these fits; we find that using different sub-clusters among residues in the  $\beta$ -turn region leads to very similar fit results (Figure 6.S2). For the fit shown in Figure 6.1b, c, we chose five residues for which data of both crystals were available: 23 (only 950 MHz, as no 600 MHz data were available for cubic-PEG-ub) and 26, 27, 55, 58 (600 and 950 MHz).

NERRD data were fitted using in-house written python programs and numerical spin-evolution simulations using the program GAMMA [60]. Briefly, numerical simulations of the fate of  $^{15}\text{N}$  coherence under a spin-lock was simulated in the presence of a two-site exchange process; hereby, a two-spin ( $^1\text{H}$ - $^{15}\text{N}$ ) system was considered, as described elsewhere [61]. A grid of such simulations with varying populations, jump angles and exchange rate constants was simulated and  $R_{1\rho}$  decay rate constants were extracted by fitting a monoexponential function to the simulated decay. The experimental  $R_{1\rho}$  RD profiles were then compared to each simulated NERRD profile to retrieve the best-fit set of dynamic parameters; hereby, a single additive fit parameter was allowed when comparing each NERRD curve (i.e. a general additive offset,  $R_{1\rho,\text{add}}$ ). This offset takes into account that other motional modes, such as nanosecond motions, contribute to the experimental data, but are not considered in the simulation. A  $\chi^2$  value was obtained by summing deviations  $[(R_{1\rho,\text{exp}} - R_{1\rho,\text{sim}})/\sigma_{\text{exp}}]^2$  over all RD points and over several residues (D52 and R54 in the case of MPD-ub; 22 residues for cubic-PEG-ub), and normalized by the degrees of freedom, to obtain a reduced  $\chi^2$ , which is reported in Figures 6.1e and 6.2b. In the fit of the cubic-PEG-ub data, where we assume that the motion is a diffusive-like motion rather than a jump model to a minor state, we set the population to 50%. We have also investigated the possibility to jointly fit BMCRD and NERRD data points. The fit parameters are similar to the ones of the individual fits of the two regimes; the inclusion of NERRD does not overcome the inherent inability to disentangle populations from  $\Delta\omega$ . This combined approach will be presented elsewhere.

### 6.6.3 MD simulations

The 2- $\mu\text{s}$ -long MD simulations of MPD-ub crystal (24 ubiquitin molecules) and cubic-PEG-ub crystal (48 ubiquitin molecules) recorded under Amber ff99SB\*-ILDN force field [62–64] in the presence of SPC/E water [65] are the same as have been described

previously [36]. Analogous 2- $\mu$ s-long trajectories have been recorded for MPD-ub crystals containing single-residue mutations E24A, G53A, K63A, E64A, or a negatively charged form of E24 as well as for cubic-PEG-ub crystals containing mutations K11A or G53A. All alanine substitutions did not create any conflicts in the initial crystal coordinates. Note that the compound statistics from each crystal trajectory is equivalent to  $2 \times 24 = 48$  or  $2 \times 48 = 96$   $\mu$ s of ubiquitin dynamics. The trajectories exhibit a certain amount of drift, as has been discussed by us earlier [36]. In addition, we have also recorded a 1- $\mu$ s-long trajectory of cubic-PEG-ub crystal, where C $\alpha$  atoms within the secondary structure of ubiquitin have been restrained to their respective positions in the crystallographic structure by means of harmonic restraints with force constant 5 kcal/(mol $\cdot$  $\text{\AA}^2$ ). The rocking motion was largely suppressed in this trajectory, leading to decreased formation of K11<sub>chainA</sub>-D52<sub>chainB</sub> salt bridge and lower content of  $\beta$ II species. However, this trajectory featured a significant proportion of distinctive  $\beta$ II' species, which appears to be an artefact. Therefore the restrained simulation was discontinued. As a control we have also recorded 2- $\mu$ s-long trajectories of MPD-ub crystal and cubic-PEG-ub crystal (wild-type and K11A variant) using newer ff14SB force field [66]. The results of these simulations have confirmed all of our findings (not shown). Finally, the previously reported solution trajectory of ubiquitin has been extended to 8  $\mu$ s.

In processing the MD trajectories we used geometric criteria to identify hydrogen bonds (nitrogen-oxygen distance less than 3.5  $\text{\AA}$ , angle 150 $^\circ$ -180 $^\circ$ ) and salt bridges (at least one nitrogen-oxygen distance less than 4 $\text{\AA}$ , centroid distance 5 $\text{\AA}$ ) [67,68]. To calculate the amplitude of rocking motion, we used the following sequence of steps: the MD frames were resampled with the step of 1 ns and then averaged to obtain reference structure for the simulated crystal cell; protein molecules from each individual MD snapshot were superimposed onto their counterparts in the reference structure (via C $\alpha$  atoms within helices and beta sheets); when superimposing the pair of molecules, rotation matrix has been determined and the amplitude of rotation was extracted; the amplitudes were then averaged over multiple molecules and MD frames. Chemical shifts have been calculated by processing the MD coordinates using SHIFTX+ module from the program SHIFTX2 [56]. Given the limited accuracy of chemical shift prediction software, only the more pronounced modulation effects with  $\Delta\delta(^{15}\text{N})$  on the order of 1 ppm can be reliably identified. Special tests have been performed to evaluate intermolecular contributions into chemical shifts; we have found that these contributions can be safely neglected for the problem at hand.

## Acknowledgements

This work was financially supported by the European Research Council (ERC-Stg-2012- 311318-ProtDyn2Function), the French Research Agency ANR (ANR 10-PDOC-011- 01), as well as Commissariat à l'énergie atomique et aux énergies alternatives (CEA), Centre National de la Recherche Scientifique (CNRS) and Université Grenoble Alpes. This work used the platforms of the Grenoble Instruct Center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for

Structural Biology (PSB). Funding from the Russian Science Foundation to N. R. S. is acknowledged. N.C. is supported by a fellowship from the Fondation France Alzheimer. We thank Dr. Peixiang Ma for exploratory experiments and insightful discussions. We thank Anastasya Shilova and Manfred Burghammer (ESRF Grenoble) for insightful discussions and for sharing results about X-ray structures prior to publication.

## 6.7 Supporting information

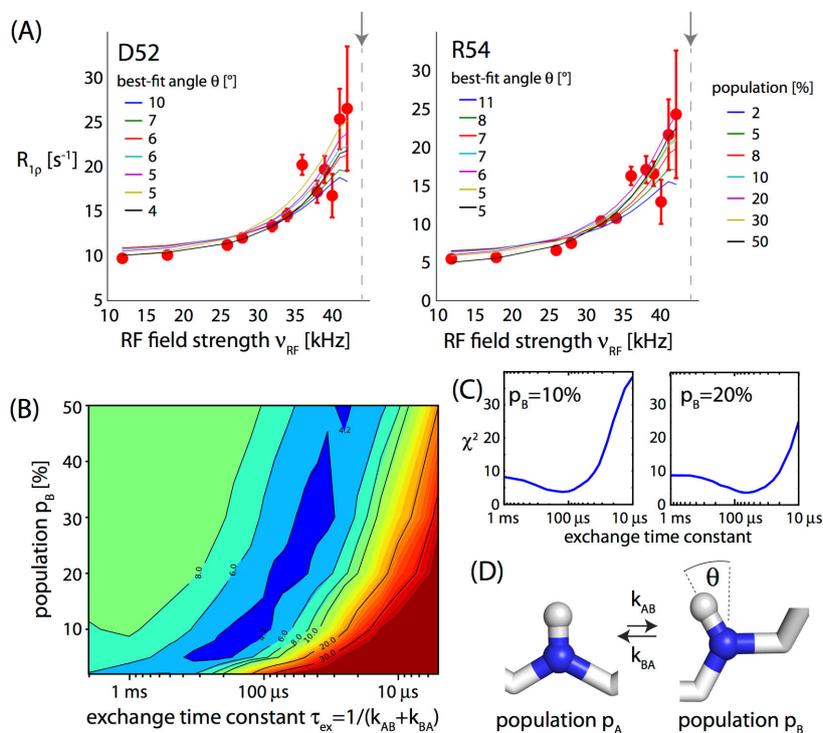


Figure 6.S1: (A) NERRD data of D52 and R54 in MPD-ub crystals, obtained at a MAS frequency of 44.053 kHz. To fit these data, a set of numerical simulations of the  $^{15}\text{N}$  relaxation was performed, as described before<sup>1</sup>, assuming a two-site jump model (shown in panel D). Diffusive-like models cannot be simulated within our simulation framework, but it has been shown that two-site jump models can be related to more general model-free-like models.<sup>2</sup> In this grid search the following parameters were varied: population, time scale of exchange, angle of jump. The data of these two residues were jointly fitted to a single motional process, in which population and exchange time constant are common, while the angle is fitted individually to each residue. To obtain a  $\chi^2$  value, reflecting the deviation from experimental data to simulations, a constant, i.e. RF-field independent offset value was fitted. This offset takes into account fast dynamics, which are not represented in our simulations. Panel (B) shows the  $\chi^2$  for this grid-search procedure as a function of the exchange time constant and the population of the involved states. Panel (C) shows traces at two selected minor-state populations.

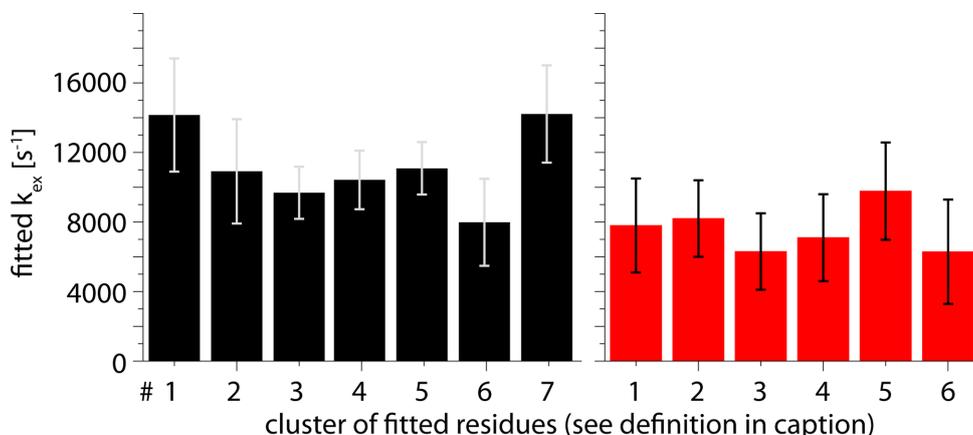


Figure 6.S2: Exchange rate constants from fits of a two-state exchange model to BMCRD data of MPD-ub (black) and cubic-PEG-ub (red), using data from different residues and different  $B_0$  field strengths. These data show that the exact outcome of the fit does not significantly depend on the choice of the subset of data. The data used in these fits are (residue numbers@field strength): MPD-ub: 1: 22, 23, 26, 27, 54, 55, 56, 57, 58@600 MHz; 2: 22, 23, 26, 54, 55, 56, 57, 58@950 MHz; 3: 26,27,55,58@600,950 MHz; 4: 55, 58@600,950 MHz and 23@950 MHz; 5: 27,55,58@600,950 MHz and 23@950 MHz; 6: 26,27,55,58@600 MHz and 23@950 MHz; 7: 23, 26, 27, 55, 58@600 MHz; cubic-PEG-ub: 1: 23, 26, 27, 55, 58@950 MHz; 2: 600MHz: 26, 27,55,58@600 MHz and 23@950 MHz; 3: 55, 58@600,950 MHz and 23@950 MHz; 4: 600+950MHz: 27,55,58 at 600,950 MHz and 23@950 MHz; 5: 26, 27, 50, 55, 58@600 MHz; 6: 20, 55, 58, 59 @ 600,950 MHz.

Table 6.S1: Estimation of the population of minor-state in MPD-ub. These calculations are based on the assumption that the ratio of the  $\phi_{\text{ex}}$  values in cubic-PEG-ub and MPD-ub is, on average, due to different minor-state populations in the two crystals. The experimentally determined ratio is  $\phi_{\text{ex}}(\text{cubic-PEG-ub})/\phi_{\text{ex}}(\text{cubic-PEG-ub}) = 3.3 \pm 1.0$ . Therefore, one can calculate minor-state populations by realizing that  $\phi_{\text{ex}} = p_{\beta\text{I}} \cdot p_{\beta\text{II}} \cdot |\Delta\omega|^2$ , where  $p_{\beta\text{I}} \cdot p_{\beta\text{II}}$  is the product of the populations of the two exchanging states, and with the assumption that  $\Delta\omega$  is identical we have  $p_{\beta\text{I}} \cdot p_{\beta\text{II}}(\text{cubic-PEG-ub})/p_{\beta\text{I}} \cdot p_{\beta\text{II}}(\text{MPD-ub}) = 3.3 \pm 1.0$ . The range of calculated populations in MPD-ub are based on the error bar of  $\phi_{\text{ex}}$ , and represent the range within one standard deviation.

Assumed minor state population in cubic-PEG-ub	Calculated minor state population in MPD-ub
50%	5 to 12%
40%	4.5 to 11.5%
30%	4 to 10%
20%	3 to 7%

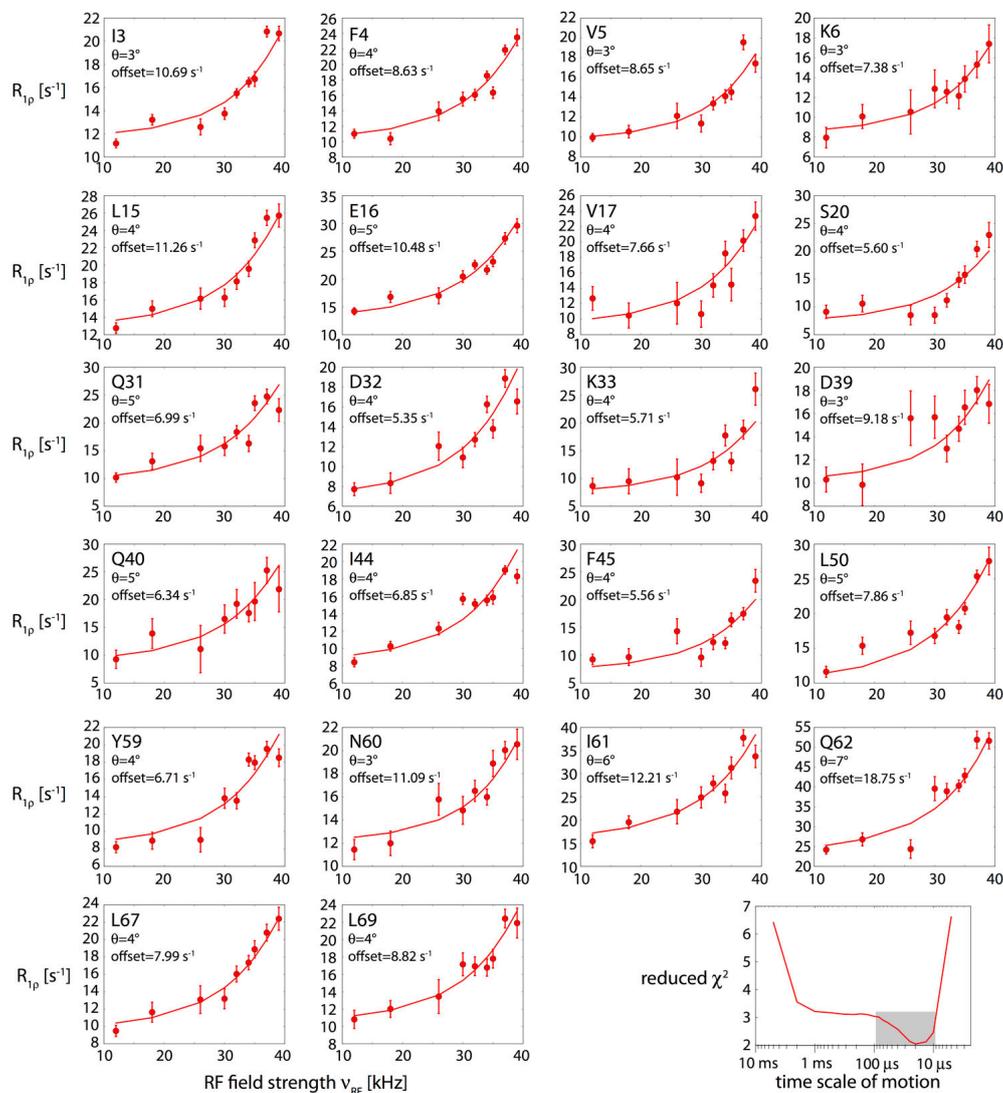


Figure 6.S3: NERRD profiles measured in cubic-PEG-ub, at an MAS frequency of 44.053 kHz. In this selection of 22 residues, we have excluded those residues that are in the  $\beta$ -turn region (22-27 and 51-58), and data with low signal/noise or resonance overlap. The data of all these residues were fitted against numerical spin-dynamics simulations (see Methods section).

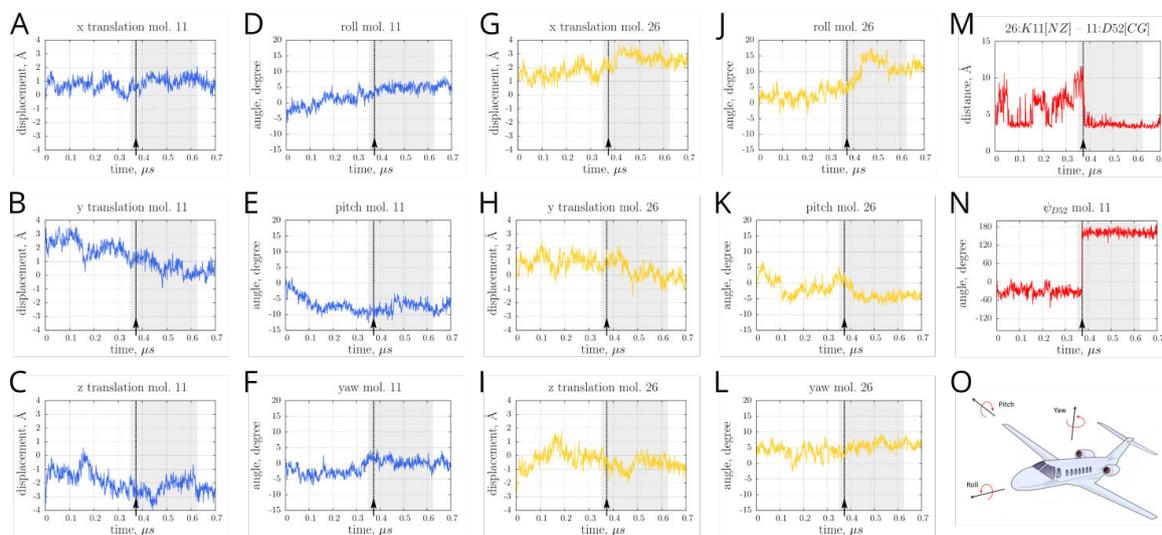


Figure 6.S4: (A-L) Overall motion of the two ubiquitin molecules from the MD trajectory of cubic-PEG-ub crystal. The selected ubiquitin molecules are the same as shown in Film S1, with graphs colored accordingly (gold color for chain A, blue color for chain B). Shown are the displacements of the center of mass of the two ubiquitin molecules relative to their position in the original crystal 3N30, as well as the variables describing the reorientational motion of the molecules (roll, pitch, and yaw). Grey band in the plots indicate the duration of the movie Film S1. The arrow indicates the moment when the salt bridge K11<sub>A</sub>-D52<sub>B</sub> was formed and the  $\beta$ I  $\rightarrow$   $\beta$ II transition occurred. (M) Distance between K11<sub>A</sub> N $\zeta$  and D52<sub>B</sub> C $\gamma$  which is indicative of the salt bridge formation. (N) Torsional angle  $\psi$  in residue D52<sub>B</sub> indicative of the  $\beta$ I/ $\beta$ II conformation. (O) Scheme illustrating the definition of roll, pitch, and yaw.

Table 6.S2: Alternate conformations modeled in room-temperature XRD crystal structures of the three crystal forms discussed in this study. Alternate states are found for the  $\beta$ -turn region in MPD-ub and cubic-PEG-ub, but not in rod-PEG-ub. This mirrors the finding by MD that this latter crystal does not undergo  $\beta$ I-  $\beta$ II transitions.

Protein	MPD 100K (3ONS)	MPD 297K	Cubic 100K (4XOK)	Cubic 297K	Rods 100K (4XOL)	Rods 297K
Residue numbers with alternative conformations	-	6,13,24,29, 34,40,52,53, 63, 69	-	51,52,53,54,72- chain A 42, 51, 54 -chain B	-	72-chain B 21,40- chain C
Total N of alternative conformations	0	10	0	8	0	3

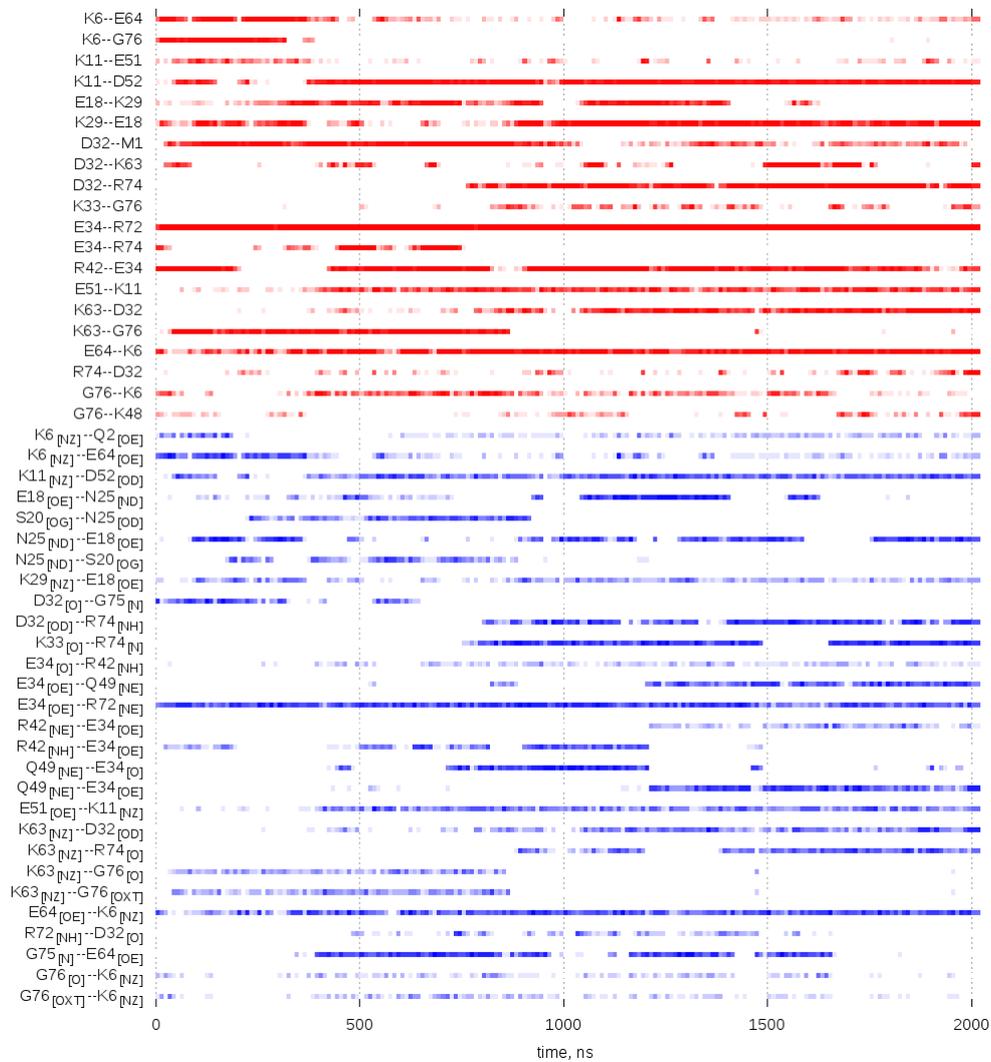


Figure 6.S5: Time evolution of intermolecular salt bridges (blue) and intermolecular hydrogen bonds (red) originating at the molecule #26 in 2- $\mu$ s MD trajectory of cubic-PEG-ub crystal. To generate this plot the trajectory was sampled with 1-ns step and then divided into 20-ns bins. The intensity of color reflects the content of the interaction in a given bin (bright color corresponds to the persistent interaction, pale color corresponds to a sporadic interaction).

## 6.8 References

- (1) Burnley, B. T.; Afonine, P. V.; Adams, P. D.; Gros, P. Modelling Dynamics in Protein Crystal Structures by Ensemble Refinement. *Elife* 2012, 1, e00311.
- (2) de Bakker, P. I. W.; Furnham, N.; Blundell, T. L.; DePristo, M. a. Conformer Generation under Restraints. *Curr. Opin. Struct. Biol.* 2006, 16 (2), 160–165.
- (3) Lang, P. T.; Holton, J. M.; Fraser, J. S.; Alber, T. Protein Structural Ensembles Are Revealed by Redefining X-Ray Electron Density Noise. *Proc. Natl. Acad. Sci. USA* 2014, 111 (1), 237–242.
- (4) Fraser, J. S.; van den Bedem, H.; Samelson, A. J.; Lang, P. T.; Holton, J. M.; Echols, N.; Alber, T. Accessing Protein Conformational Ensembles Using Room-Temperature X-Ray Crystallography. *P Natl Acad Sci U.S.A.* 2011, 108 (39), 16247–16252.
- (5) Van Benschoten, A. H.; Liu, L.; Gonzalez, A.; Brewster, A. S.; Sauter, N. K.; Fraser, J. S.; Wall, M. E. Measuring and Modeling Diffuse Scattering in Protein X-Ray Crystallography. *Proc. Natl. Acad. Sci.* 2016, 113 (15), 201524048.
- (6) Wall, M. E.; Van Benschoten, A. H.; Sauter, N. K.; Adams, P. D.; Fraser, J. S.; Terwilliger, T. C. Conformational Dynamics of a Crystalline Protein from Microsecond-Scale Molecular Dynamics Simulations and Diffuse X-Ray Scattering. *Proc. Natl. Acad. Sci. U. S. A.* 2014, 111 (50), 17887–17892.
- (7) Mittermaier, A.; Kay, L. E. New Tools Provide New Insights in NMR Studies of Protein Dynamics. *Sci. (New York, N.Y.)*, Abstr. = There is Grow. Evid. that Struct. Flex. Play. a Cent. role Funct. protein Mol. Many Exp. data come from Nucl. Magn. Reson. Spectrosc. a Tech. that a 2006, 312 (5771), 224–228.
- (8) Mao, B.; Tejero, R.; Baker, D.; Montelione, G. T. Protein NMR Structures Refined with Rosetta Have Higher Accuracy Relative to Corresponding X-Ray Crystal Structures. *J. Am. Chem. Soc.* 2014, 136 (5), 1893–1906.
- (9) Bax, A. Weak Alignment Offers New NMR Opportunities to Study Protein Structure and Dynamics. *Protein Sci.* 2003, 12 (1), 1–16.
- (10) Ferrari, D.; Merli, A.; Peracchi, A.; Di Valentin, M.; Carbonera, D.; Rossi, G. L. Catalysis and Electron Transfer in Protein Crystals: The Binary and Ternary Complexes of Methylamine Dehydrogenase with Electron Acceptors. *Biochim. Biophys. Acta* 2003, 1647 (1–2), 337–342.
- (11) Pearson, A. R.; Wilmot, C. M. Catching Catalysis in the Act: Using Single Crystal Kinetics to Trap Methylamine Dehydrogenase Reaction Intermediates. *Biochim. Biophys. Acta* 2003, 1647 (1–2), 381–389.
- (12) Ronda, L.; Bruno, S.; Bettati, S.; Mozzarelli, A. Protein Crystal Microspectrophotometry. *Biochim. Biophys. Acta* 2011, 1814 (6), 734–741.
- (13) Merli, A.; Brodersen, D. E.; Morini, B.; Chen, Z.; Durley, R. C.; Mathews, F. S.; Davidson, V. L.; Rossi, G. L. Enzymatic and Electron Transfer Activities in Crystalline Protein Complexes. *J. Biol. Chem.* 1996, 271 (16), 9177–9180.
- (14) Mozzarelli, A.; Rossi, G. L. Protein Function in the Crystal. *Annu. Rev. Biophys. Biomol. Struct.* 1996, 25, 343–365.
- (15) Sahu, S.; Bhuyan, A.; Majumdar, A.; Udgaonkar, J. Backbone Dynamics of Barstar: A N-15 NMR Relaxation Study. *Proteins* 2000, 41 (4), 460–474.
- (16) Kordel, J.; Skelton, N.; Akke, M.; Palmer, A. G.; Chazin, W. Backbone Dynamics of Calcium-Loaded Calbinding-D(9k) Studied by 2-Dimensional Proton-Detected 15N NMR Spectroscopy. *Biochemistry* 1992, 31 (20), 4856–4866.
- (17) Powers, R.; Clore, G.; Garrett, D.; Gronenborn, A. Relationships between the Precision of High-Resolution Protein NMR Structures, Solution Order Parameters and Crystallographic B Factors. *J. Magn. Reson. B* 1993, 101 (3), 325–327.
- (18) Fenwick, R. B.; van den Bedem, H.; Fraser, J. S.; Wright, P. E. Integrated Description of Protein Dynamics from Room-Temperature X-Ray Crystallography and NMR. *Proc. Natl. Acad. Sci. U. S. A.* 2014, 111 (4), E445–54.
- (19) Agarwal, V.; Xue, Y.; Reif, B.; Skrynnikov, N. R. Protein Side-Chain Dynamics as Observed by Solution- and Solid-State NMR Spectroscopy: A Similarity Revealed. *J. Am. Chem. Soc.* 2008, 130 (49), 16611–16621.
- (20) Haller, J. D.; Schanda, P. Amplitudes and Time Scales of Picosecond-to-Microsecond Motion in Proteins Studied by Solid-State NMR: A Critical Evaluation of Experimental Approaches and Application to Crystalline Ubiquitin. *J. Biomol. NMR* 2013, 57 (3), 263–280.
- (21) Yang, J.; Tasayco, M. L.; Polenova, T. Dynamics of Reassembled Thioredoxin Studied by Magic Angle Spinning NMR: Snapshots from Different Time Scales. *J. Am. Chem. Soc.* 2009, 131 (38), 13690–13702.
- (22) Halle, B. Flexibility and Packing in Proteins. *P Natl Acad Sci U.S.A.* 2002, 99 (3), 1274–1279.

- (23) Eastman, P.; Pellegrini, M.; Doniach, S. Protein Flexibility in Solution and in Crystals. *J. Chem. Phys.* 1999, 110 (20), 10141–10152.
- (24) Stocker, U.; Spiegel, K.; van Gunsteren, W. On the Similarity of Properties in Solution or in the Crystalline State: A Molecular Dynamics Study of Hen Lysozyme. *J. Biomol. NMR* 2000, 18 (1), 1–12.
- (25) Rueda, M.; Ferrer-Costa, C.; Meyer, T.; Pérez, A.; Camps, J.; Hospital, A.; Gelpí, J. L.; Orozco, M. A Consensus View of Protein Dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 2007, 104 (3), 796–801.
- (26) Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* 2007, 450 (7172), 964–972.
- (27) Massi, F.; Grey, M. J.; Palmer, A. G. Microsecond Timescale Backbone Conformational Dynamics in Ubiquitin Studied with NMR R1rho Relaxation Experiments. *Protein Sci.* 2005, 14 (3), 735–742.
- (28) Hansen, D. F.; Feng, H.; Zhou, Z.; Bai, Y.; Kay, L. E. Selective Characterization of Microsecond Motions in Proteins by NMR Relaxation. *J. Am. Chem. Soc.* 2009, 131 (16), 16257–16265.
- (29) Salvi, N.; Ulzega, S.; Ferrage, F.; Bodenhausen, G. Time Scales of Slow Motions in Ubiquitin Explored by Heteronuclear Double Resonance. *J. Am. Chem. Soc.* 2012, 134 (5), 2481–2484.
- (30) Smith, C. A.; Ban, D.; Pratihari, S.; Giller, K.; Paulat, M.; Becker, S.; Griesinger, C.; Lee, D.; de Groot, B. L. Allosteric Switch Regulates Protein–protein Binding through Collective Motion. *Proc. Natl. Acad. Sci.* 2016, No. 20, 201519609.
- (31) Mills, J. L.; Szyperski, T. Protein Dynamics in Supercooled Water: The Search for Slow Motional Modes. *J. Biomol. NMR* 2002, 23 (1), 63–67.
- (32) Lindorff-larsen, K.; Maragakis, P.; Piana, S.; Shaw, D. E.; Shaw, D. E. Picosecond to Millisecond Structural Dynamics in Human Ubiquitin. 260 (212), 1–27.
- (33) Huang, K.-Y.; Amodeo, G. A.; Tong, L.; McDermott, A. The Structure of Human Ubiquitin in 2-Methyl-2,4-Pentandiol: A New Conformational Switch. *Protein Sci.* 2011, 20 (3), 630–639.
- (34) Arnesano, F.; Belviso, B. D.; Caliendo, R.; Falini, G.; Fermani, S.; Natile, G.; Siliqi, D. Crystallographic Analysis of Metal-Ion Binding to Human Ubiquitin. *Chem. - A Eur. J.* 2011, 17 (5), 1569–1578.
- (35) Falini, G.; Fermani, S.; Tosi, G.; Arnesano, F.; Natile, G. Structural Probing of Zn(II), Cd(II) and Hg(II) Binding to Human Ubiquitin. *Chem. Commun. (Camb)*. 2008, No. 45, 5960–5962.
- (36) Ma, P.; Xue, Y.; Coquelle, N.; Haller, J. D.; Yuwen, T.; Ayala, I.; Willbold, D.; Colletier, J.; Skrynnikov, N. R. Observing the Overall Rocking Motion of a Protein in a Crystal. *Nat Commun* 2015, 6, 1–24.
- (37) Ma, P.; Haller, J. D.; Zajakala, J.; Macek, P.; Sivertsen, A. C.; Willbold, D.; Boissbouvier, J.; Schanda, P. Probing Transient Conformational States of Proteins by Solid-State R1rho Relaxation-Dispersion NMR Spectroscopy. *Angew. Chemie - Int. Ed.* 2014, 53 (17), 4312–4317.
- (38) Palmer, A. G.; Massi, F. Characterization of the Dynamics of Biomacromolecules Using Rotating-Frame Spin Relaxation NMR Spectroscopy. *Chem. Rev.* 2006, 106 (5), 1700–1719.
- (39) McConnell, H. M. Reaction Rates by Nuclear Magnetic Resonance. *J. Chem. Phys.* 1958, 28 (3), 430.
- (40) Schanda, P.; Ernst, M. Studying Dynamics by Magic-Angle Spinning Solid-State NMR Spectroscopy: Principles and Applications to Biomolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* 2016, 96, 1–46.
- (41) Quinn, C. M.; McDermott, A. E. Quantifying Conformational Dynamics Using Solid-State R1rho Experiments. *J. Magn. Reson.* 2012, 222, 1–7.
- (42) Krushelnitsky, A.; Zinkevich, T.; Reichert, D.; Chevelkov, V.; Reif, B. Microsecond Time Scale Mobility in a Solid Protein as Studied by the 15N R(1rho) Site-Specific NMR Relaxation Rates. *J. Am. Chem. Soc.* 2010, 132 (34), 11850–11853.
- (43) Lewandowski, J. R.; Sass, H. J.; Grzesiek, S.; Blackledge, M.; Emsley, L. Site-Specific Measurement of Slow Motions in Proteins. *J. Am. Chem. Soc.* 2011, 133 (42), 16762–16765.
- (44) Kurbanov, R.; Zinkevich, T.; Krushelnitsky, A. The Nuclear Magnetic Resonance Relaxation Data Analysis in Solids: General R1/R1rho Equations and the Model-Free Approach. *J. Chem. Phys.* 2011, 135 (18), 184104.
- (45) Smith, A. A.; Testori, E.; Cadalbert, R.; Meier, B. H.; Ernst, M. Characterization of Fibril Dynamics on Three Timescales by Solid-State NMR. *J. Biomol. NMR* 2016.
- (46) Maragakis, P.; Lindorff-Larsen, K.; Eastwood, M. P.; Dror, R. O.; Klepeis, J. L.; Arkin, I. T.; Jensen, M.; Xu, H.; Trbovic, N.; Friesner, R. a.; et al. Microsecond Molecular Dynamics Simulation Shows Effect of Slow Loop Dynamics on Backbone Amide Order Parameters of Proteins. *J. Phys. Chem. B* 2008, 112 (19), 6155–6158.
- (47) Tollinger, M.; Sivertsen, A. C.; Meier, B. H.; Ernst, M.; Schanda, P. Site-Resolved Measurement of Microsecond-to-Millisecond Conformational-Exchange Processes in Proteins by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 2012, 134 (36), 14800–14807.
- (48) Kuzmanic, A.; Pannu, N. S.; Zagrovic, B. X-Ray Refinement Significantly Underestimates the Level of Microscopic Heterogeneity in Biomolecular Crystals. *Nat Commun* 2014, 5, 3220.

- (49) Mollica, L.; Baias, M.; Lewandowski, R.; Wylie, B. J.; Sperling, L. J.; Rienstra, C. M.; Emsley, L.; Blackledge, M. Atomic-Resolution Structural Dynamics in Crystalline Proteins from. 2012.
- (50) Xue, Y.; Skrynnikov, N. R. Ensemble MD Simulations Restrained via Crystallographic Data: Accurate Structure Leads to Accurate Dynamics. *Protein Sci.* 2014, 23 (4), 488–507.
- (51) Stafford, K. A.; Robustelli, P.; Palmer, A. G. Thermal Adaptation of Conformational Dynamics in Ribonuclease H. *PLoS Comput. Biol.* 2013, 9 (10).
- (52) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase. *J. Am. Chem. Soc.* 1998, 120 (27), 6836–6837.
- (53) Sidhu, A.; Surolia, A.; Robertson, A. D.; Sundd, M. A Hydrogen Bond Regulates Slow Motions in Ubiquitin by Modulating a  $\beta$ -Turn Flip. *J. Mol. Biol.* 2011, 411 (5), 1037–1048.
- (54) Olsson, M. H. M.; SØndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical P K a Predictions. *J. Chem. Theory Comput.* 2011, 7 (2), 525–537.
- (55) Trevino, S. R.; Schaefer, S.; Scholtz, J. M.; Pace, C. N. Increasing Protein Conformational Stability by Optimizing  $\beta$ -Turn Sequence. *J. Mol. Biol.* 2007, 373 (1), 211–218.
- (56) Han, B.; Liu, Y.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: Significantly Improved Protein Chemical Shift Prediction. *J. Biomol. NMR* 2011, 50 (1), 43–57.
- (57) Xue, Y.; Ward, J. M.; Yuwen, T.; Podkorytov, I. S.; Skrynnikov, N. R. Ms Time-Scale Conformational Exchange in Proteins: Using Long MD Trajectory to Simulate NMR Relaxation Dispersion Data. *J. Am. Chem. Soc.* 2011.
- (58) Schanda, P.; Meier, B. H.; Ernst, M. Quantitative Analysis of Protein Backbone Dynamics in Microcrystalline Ubiquitin by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 2010, 132 (45), 15957–15967.
- (59) Morin, S.; Linnet, T. E.; Lescanne, M.; Schanda, P.; Thompson, G. S.; Tollinger, M.; Teilum, K.; Gagné, S.; Marion, D.; Griesinger, C.; et al. Relax: The Analysis of Biomolecular Kinetics and Thermodynamics Using NMR Relaxation Dispersion Data. *Bioinformatics* 2014, 30 (15), 2219–2220.
- (60) Smith, S.; Levante, T.; Meier, B.; Ernst, R. Computer Simulations in Magnetic Resonance. An Object-Oriented Programming Approach. *J. Magn. Reson. (San Diego, Calif)* 1994, 106, 75–105.
- (61) Kurauskas, V.; Weber, E.; Hessel, A.; Ayala, I.; Marion, D.; Schanda, P. Cross-Correlated Relaxation of Dipolar Coupling and Chemical-Shift Anisotropy in Magic-Angle Spinning R 1 $\rho$  NMR Measurements: Application to Protein Backbone Dynamics Measurements. *J. Phys. Chem. B* 2016, acs.jpcc.6b06129.
- (62) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Structure, Function and Genetics.* 2006, pp 712–725.
- (63) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins Struct. Funct. Bioinforma.* 2010, 78 (8), 1950–1958.
- (64) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* 2009, 113 (26), 9004–9015.
- (65) Cerutti, D. S.; Le Trong, I.; Stenkamp, R. E.; Lybrand, T. P. Simulations of a Protein Crystal: Explicit Treatment of Crystallization Conditions Links Theory and Experiment in the Streptavidin-Biotin Complex. *Biochemistry* 2008, 47 (46), 12065–12077.
- (66) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015, 11 (8), 3696–3713.
- (67) Kumar, S.; Nussinov, R. Relationship between Ion Pair Geometries and Electrostatic Strengths in Proteins. *Biophys. J.* 2002, 83 (3), 1595–1612.
- (68) Xue, Y.; Yuwen, T.; Zhu, F.; Skrynnikov, N. R. Role of Electrostatic Interactions in Binding of Peptides and Intrinsically Disordered Proteins to Their Folded Targets. 1. NMR and MD Characterization of the Complex between the c-Crk N-SH3 Domain and the Peptide Sos. *Biochemistry* 2014, 53 (41), 6473–6495.



## 7 Appendix

### List of publications:

- Kurauskas V, Crublet E, Macek P, Kerfah R, Gauto DF, Boisbouvier J, Schanda P. (2016). Sensitive proton-detected solid-state NMR spectroscopy of large proteins with selective CH<sub>3</sub> labelling: application to the 50S ribosome subunit. *Chemical Communications* 52(61), 9558-61
- Kurauskas V, Weber E, Hessel A, Ayala I, Marion D, Schanda P. (2016). Cross-Correlated Relaxation of Dipolar Coupling and Chemical-Shift Anisotropy in Magic-Angle Spinning R<sub>1ρ</sub> NMR Measurements: Application to Protein Backbone Dynamics Measurements. *The Journal of Physical Chemistry B* 120, 8905-13
- Rodrigues CD, Henry X, Neumann E, Kurauskas V, Bellard L, Fichou Y, Schanda P, Schoehn G, Rudner DZ, Morlot C. (2016). A ring-shaped conduit connects the mother cell and forespore during sporulation in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* 113(41), 11585-11590
- Kurauskas V, Schanda P, Sounier R. (2017). *Methods in Molecular Biology*, In press.

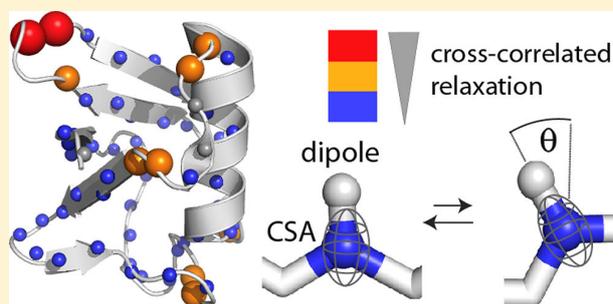
# Cross-Correlated Relaxation of Dipolar Coupling and Chemical-Shift Anisotropy in Magic-Angle Spinning $R_{1\rho}$ NMR Measurements: Application to Protein Backbone Dynamics Measurements

Vilius Kurauskas, Emmanuelle Weber, Audrey Hessel, Isabel Ayala, Dominique Marion,\* and Paul Schanda\*

CEA, CNRS, Université Grenoble Alpes, Institut de Biologie Structurale (IBS), 38027 Grenoble, France

## Supporting Information

**ABSTRACT:** Transverse relaxation rate measurements in magic-angle spinning solid-state nuclear magnetic resonance provide information about molecular motions occurring on nanosecond-to-millisecond (ns–ms) time scales. The measurement of heteronuclear ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ) relaxation rate constants in the presence of a spin-lock radiofrequency field ( $R_{1\rho}$  relaxation) provides access to such motions, and an increasing number of studies involving  $R_{1\rho}$  relaxation in proteins have been reported. However, two factors that influence the observed relaxation rate constants have so far been neglected, namely, (1) the role of CSA/dipolar cross-correlated relaxation (CCR) and (2) the impact of fast proton spin flips (i.e., proton spin diffusion and relaxation). We show that CSA/D CCR in  $R_{1\rho}$  experiments is measurable and that the CCR rate constant depends on ns–ms motions; it can thus provide insight into dynamics. We find that proton spin diffusion attenuates this CCR due to its decoupling effect on the doublet components. For measurements of dynamics, the use of  $R_{1\rho}$  rate constants has practical advantages over the use of CCR rates, and this article reveals factors that have so far been disregarded and which are important for accurate measurements and interpretation.



## INTRODUCTION

The structure of a protein is determined by a large number of individually weak interactions, which constantly reshuffle at ambient temperatures. This dynamic exchange between different conformers is often important for functions such as binding or enzymatic turnover. NMR spectroscopy is very well suited to provide direct atomic-level access to the amplitudes and time scales of such motions. Whereas solution-state NMR is a well-established tool for such studies, magic-angle spinning (MAS) solid-state NMR (ssNMR) is rapidly gaining popularity in studies of biomolecular dynamics, complementing its solution-state counterpart in cases in which limited solubility or a large size hampers the application of solution-state methods. In addition to opening new fields of applications, ssNMR also has theoretical advantages for determining dynamics, compared to solution-state NMR. In particular, the absence of overall tumbling motion enables the study of motions without “blind windows” of time scales, that is, over all time scales from picoseconds to seconds, in contrast to its solution-state counterpart, which is severely challenged on the nanosecond to several-microseconds time scale.<sup>1</sup> In MAS ssNMR, motions can be studied through (1) their averaging effect on anisotropic interactions (in particular, the averaging of dipolar couplings and chemical-shift anisotropies), which report on the amplitude of motion seen by these interactions, and (2)

through the investigation of nuclear spin relaxation rate constants, which depend on amplitudes and time scales of motions. Many biological processes, such as binding, gating, or enzymatic turnover, occur on time scales of hundreds of nanoseconds to milliseconds (ns–ms). Studying these time scales is, therefore, of particular relevance when attempting to decipher the molecular basis of these processes. In MAS ssNMR, motion on these time scales can be accessed by measurements of transverse relaxation rate constants, such as the decay of spin coherence,  $R_2$ . However, quantitative measurements of transverse relaxation rates in ssNMR are generally hampered by the fact that spin coherences in solid samples decay through two mechanisms: (1) due to random field fluctuations (i.e., dynamics), which is what we are interested in herein, and (2) due to evolution under the incompletely averaged anisotropic interactions, in particular due to  $^1\text{H}$ – $^1\text{H}$  dipolar couplings. The latter, termed dipolar dephasing, strongly depends on the presence of (multiple) dipolar couplings, and on the employed MAS frequency. Accessing the dynamics-induced component requires that coherent contributions to signal decay (in particular dipolar

**Received:** June 17, 2016

**Revised:** August 8, 2016

**Published:** August 8, 2016

dephasing) are suppressed to negligible levels. Several reported studies have shown that heteronuclear ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )  $R_{1\rho}$  rate constants, that is, the decay of  $C_x$  ( $N_x$ ) in the presence of a spin-lock radiofrequency (RF) field, can be measured accurately without significant interfering effects from dipolar dephasing if (1) high MAS frequencies and/or (2) extensive deuteration are applied.<sup>2–5</sup>  $R_{1\rho}$  experiments under such conditions have been reported so far for  $^{15}\text{N}$  and  $^{13}\text{C}'$  sites in small model proteins (SH3, GB1, ubiquitin).<sup>2–7</sup>  $R_{1\rho}$  experiments in solids are also particularly informative when the RF field strength approaches the rotary-resonance conditions, as has been shown for small molecules<sup>8,9</sup> and proteins.<sup>4,6</sup>

As an alternative method to probe ns–ms motions, it has been proposed to measure cross-correlated relaxation (CCR) between  $^1\text{H}$ – $^{15}\text{N}$  dipolar/ $^{15}\text{N}$  chemical-shift anisotropy (CSA) interactions in  $R_2'$ -type experiments (i.e., in the absence of a spin-lock RF field). These latter experiments measure the difference of the apparent transverse decay rate constants of the two  $^{15}\text{N}$  scalar-coupling doublet components, that is, the decay of  $N_x\text{--}[H^\alpha]$  and  $N_x\text{--}[H^\beta]$  components in a spin-echo experiment.<sup>10</sup> Even though each individual rate constant itself contains contributions from dipolar dephasing (and, therefore, cannot be interpreted quantitatively), the respective difference can be quantitatively interpreted in terms of dynamics on ns–ms time scales.<sup>11</sup> This experiment suffers, however, from the fact that the  $R_2'$  decay rates in MAS ssNMR are large, because of the dipolar-dephasing contribution. As a consequence, detection of a small differential relaxation rate from two individually large relaxation rate constants corresponding to the two doublet components ( $R_2' + \Gamma^{\text{CSA/D}}$  and  $R_2' - \Gamma^{\text{CSA/D}}$ ) is challenging. Furthermore, the rapid signal loss throughout the experiment, based on scalar-coupling transfers,<sup>10</sup> results in relatively low detection sensitivity.

Here, we investigate the effects of dipolar/CSA CCR in heteronuclear ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )  $R_{1\rho}$  experiments in MAS ssNMR. In solution-state NMR, such CCR effects have been reported<sup>12,13</sup> but have so far been neglected in MAS ssNMR. We use explicit numerical spin dynamics simulations, as well as experimental measurements, and show that CCR rate constants under spin-lock RF fields are measurable and that they report on molecular motion. The CCR rate constant has a similar dependence on dynamics as the in-phase  $R_{1\rho}$  rate constant, as we show with numerical simulations. For several reasons, we assume here that the CCR rate constant in  $R_{1\rho}$  measurements is essentially free from coherent contributions. First, it has been convincingly shown that CCR rate constants, derived from the above-mentioned experiments with free evolution ( $R_2'$ ), primarily reflect dynamics, even though each of the two doublet components itself is not free from coherent contributions.<sup>14</sup> In  $R_{1\rho}$ -based experiments, the situation can be assumed to be even more favorable, because the decay of each of the two components is much less prone to coherent decay.<sup>3</sup> Furthermore, we show that experimentally determined  $R_{1\rho}$ -derived CCR rate constants (referred to here as  $\Gamma_{1\rho}^{\text{CSA/D}}$ ) reveal the expected patterns of motions in a well-studied protein, and correlate with  $R_{1\rho}$  rate constants, which have been shown to contain only minimal coherent contributions.

However, both CCR and  $R_{1\rho}$  contain contributions from stochastic proton flips: we provide evidence that proton spin diffusion reduces the CCR rate constants due to averaging of the doublet components. Furthermore, proton spin diffusion also leads to an increase in the autorelaxation ( $R_{1\rho}$ ) rate constants, in a manner that depends on the RF field strength.

Our experimental data and simulations thus suggest that knowledge of the proton spin-diffusion rate is required for accurately determining dynamics, through either in-phase  $R_{1\rho}$  measurements or CCR measurements under spin-lock. Finally, we also discuss pulse schemes to suppress the effects of CCR and thus eliminate the biexponentiality of decay in in-phase  $R_{1\rho}$  measurements.

## MATERIALS AND METHODS

**Samples and NMR Spectroscopy.** In this study, we investigate the CCR between a heteronuclear CSA tensor ( $^{15}\text{N}$ ) and the  $^1\text{H}$ – $^{15}\text{N}$  dipolar coupling in amide moieties; our conclusion also applies to other  $^1\text{H}$ – $X$  spin pairs, such as  $^1\text{H}$ – $^{13}\text{C}$  in backbone or side-chain sites. We use  $u\text{--}[^2\text{H}, ^{15}\text{H}]$ -labeled ubiquitin, crystallized with 2-methyl-2,4-pentane-diol (MPD),<sup>15</sup> obtained by crystallization from a solution containing an effective  $\text{H}_2\text{O}/\text{D}_2\text{O}$  ratio of 60:40 (taking into account protons originating from MPD). Exchangeable sites, such as amides, are protonated at a level of 60%, in an otherwise deuterated background. Protein samples were filled into 1.6 mm Agilent MAS rotors by ultracentrifugation using a dedicated filling device (in-house built). NMR experiments were performed on a 600 MHz Agilent VNMR5 spectrometer equipped with a 1.6 mm MAS probe tuned to  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  frequencies, spinning at 38 kHz MAS frequency. The effective sample temperature, calibrated with internal standard using the bulk water frequency, was set to 300 K. Pulse sequences used in this study are shown in Figure S1 in the Supporting Information and use  $^1\text{H}$ – $^{15}\text{N}$  out-and-back cross-polarization and  $^1\text{H}$  detection. In particular, the sequence in Figure S1B was used for the CCR measurement, and (in-phase)  $R_{1\rho}$  rate constants with and without  $^1\text{H}$   $\pi$  pulse decoupling are shown in Figure S1A,C, respectively. Finally, the proton spin-diffusion experiment was performed with the sequence shown in Figure S1D.  $^{15}\text{N}$  spin-lock RF field frequencies were calibrated with a nutation experiment: after initial  $^1\text{H}$ – $^{15}\text{N}$  cross-polarization, a 5 ms spin-lock period was applied (to take into account possible detuning during a spin-lock, as this is relevant for a  $R_{1\rho}$  experiment), followed by a  $90^\circ$  phase-shifted nutation pulse. If this pulse is a  $\pi/2$  pulse, then the signal is zero. We determined the amplitude of this pulse by parametric incrementation. We find that two alternative calibration methods result in very similar calibrations: (1) a simple nutation pulse, applied at the beginning of the solvent-saturation period (see Figure S1), that is, disregarding the spin-lock period, and (2) the observation of the rotary resonance condition. For the latter, the signal intensity is monitored as a function of the  $^{15}\text{N}$  spin-lock RF field strength in an experiment that applies a 5 ms spin-lock just after the initial  $^1\text{H}$ – $^{15}\text{N}$  CP transfer (Figure S1A); a minimum is observed when the RF field matches the  $n = 1$  rotary-resonance condition  $\nu_{\text{RF}} = \nu_{\text{MAS}}$ . We find that the three methods provide essentially identical calibration results. Processing and analysis were done with nmrPipe<sup>16</sup> and NMRView (OneMoon Scientific) as well as in-house written scripts in python language. Error estimates of rate constants were based on spectral noise estimates (3 times standard deviation of the signal in empty regions of the spectrum).

**Numerical Simulations.** Simulations of the evolution of spin coherences were performed using in-house written programs using the GAMMA simulation library.<sup>17</sup> The simulation approach includes two spins ( $^1\text{H}$ ,  $^{15}\text{N}$ ), existing in two exchanging subsystems, spanning a composite Liouville

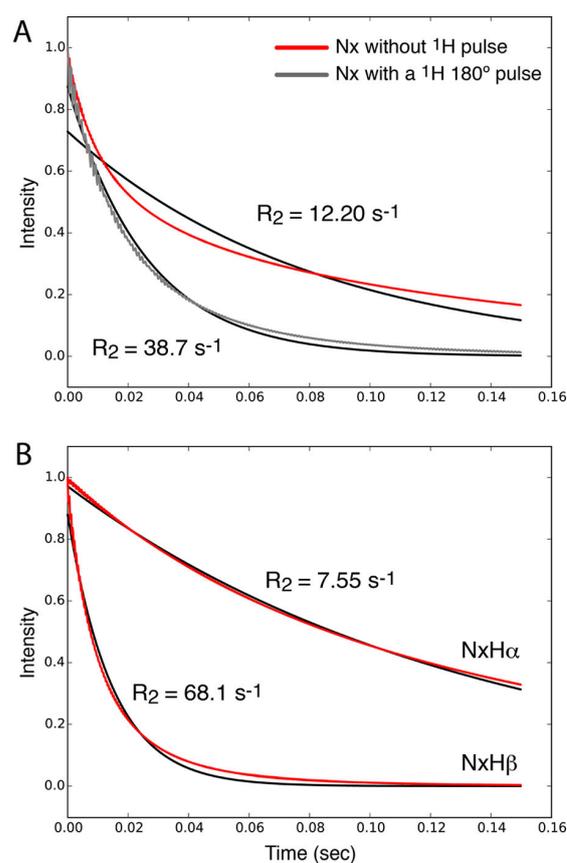
space.<sup>1,11,18,19</sup> The simulation approach and numerous results of similar kinds of simulations have been reported recently.<sup>1</sup> Briefly, the evolution of the appropriate coherences ( $N_x$ ,  $2N_xH^\alpha$ ,  $2N_xH^\beta$ ,  $H_z$ ) is followed over the course of a  $^{15}\text{N}$  spin-lock. The two exchanging substates differ in the orientations of the  $^1\text{H}$ - $^{15}\text{N}$  dipolar coupling tensor and  $^{15}\text{N}$  CSA tensor. The interaction tensor parameters were assumed as follows:  $^1\text{H}$ - $^{15}\text{N}$  dipolar coupling 22 954 Hz (corresponding to 1.02 Å bond length), axially symmetric  $^{15}\text{N}$  tensor with  $\Delta\sigma = -170$  ppm, that is,  $\sigma_{zz} = -113$  ppm (6780 Hz at 14.1 T  $B_0$  field strength). The  $^1\text{H}$  CSA parameters have negligible effect in this context. The time-domain data were fitted with python scripts to obtain relaxation rate constants, using mono- or bi-exponential functions as reported in the text.

## RESULTS AND DISCUSSION

We first aim at providing a quantitative understanding of the effects of CSA/D CCR in  $R_{1\rho}$  experiments by numerical simulations. Although a very well-established theoretical framework for nuclear spin relaxation exists (Redfield theory<sup>20</sup>), this theory is not strictly valid in cases of slow motion, which are the focus of this article. Recent systematic analyses have established that Redfield results may disagree with numerical simulations,<sup>1,7</sup> even in cases where the actual relaxation is more than 1 order of magnitude slower than the dynamic event. To circumvent these problems, we resort to numerical simulations for investigating the properties of  $R_{1\rho}$  relaxation in rotating solids. The simulation approach is based on an explicit jump model of a  $^1\text{H}$ - $^{15}\text{N}$  moiety exchanging between two conformations. The employed composite Liouville-space approach has been used similarly before.<sup>11,18,19</sup>

Figure 1 shows an example of a simulated decay of  $N_x$  coherence for a two-site exchange, using powder averaging and MAS. The observed decay (red curve in Figure 1A) is clearly nonexponential, and the exponential fit (black) unsatisfactorily describes the decay. The nonexponentiality may be ascribed to two effects. The first is that the NMR interactions that lead to relaxation (dipolar coupling, CSA) are orientation dependent, and therefore the relaxation rate constants are also orientation dependent. When summing all contributions from different crystallite orientations across the sample, as required when considering a sample with randomly oriented crystallites, the resulting relaxation is always multiexponential, as has been described and reviewed earlier.<sup>1,21,22</sup> If this is the only reason for nonmonoexponential behavior, then individual crystallites (i.e., with a single orientation within the rotor-fixed frame) should show monoexponential decays. However, we also find that when using single crystallites rather than a powder distribution, the decay is bi-exponential (data not shown), pointing to an additional mechanism.

This second possible origin of nonexponential decay is the presence of CCR due to  $^{15}\text{N}$ -CSA/ $^1\text{H}$ - $^{15}\text{N}$  dipolar-coupling interference. When observing the two components of the initial  $N_x$  coherence, namely, the coherences  $N_x\text{-}[H^\alpha]$  and  $N_x\text{-}[H^\beta]$ , the decays are nearly exponential when integrated over a powder (Figure 1B). If one considers individual crystallites, then the decay of the two coherences is indeed perfectly exponential (Figure S2C,D). One can readily show that the bi-exponential decay of  $N_x$  is due to interference of CSA and dipolar mechanisms: removing either of the two mechanisms (i.e., setting the jump angle for one of the two interactions to zero, or, equivalently, setting the interaction strength to zero) leads to identical relaxation of the two components and

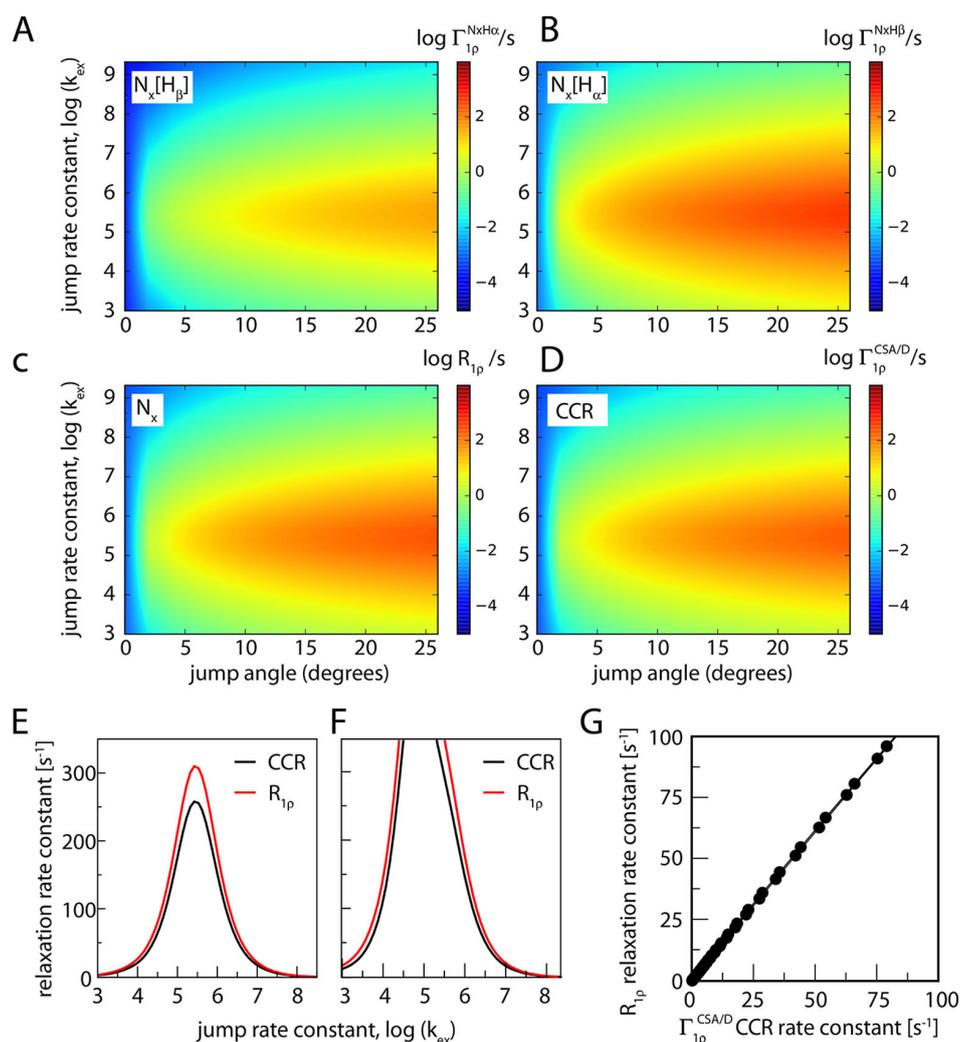


**Figure 1.** Simulated evolution of  $^{15}\text{N}$  spin coherences under a spin-lock RF field. (A) Time evolution of in-phase  $N_x$  coherence (red) and a single-exponential fit, which clearly results in poor agreement with the decay. For the simulation shown in gray, a  $^1\text{H}$   $\pi$  pulse was inserted in the center of the spin-lock period. An exponential fit is shown in black. (B) Simulation of the doublet components,  $N_x\text{-}[H^\alpha]$ ,  $N_x\text{-}[H^\beta]$  (red), and exponential fits (black). The jump angle was  $20^\circ$  and the jump rate constant was  $30\,000\text{ s}^{-1}$ .

monoexponential decay of  $N_x$  (data not shown). Furthermore, if a  $^1\text{H}$   $\pi$  pulse is inserted, thus inverting the sign of the dipolar interaction, the near-monoexponential decay is restored, as shown in Figure 1A (gray line). This is expected, as the effect of the CCR is eliminated.

Numerical simulations also allow determining the difference in the rate constants of the two components, that is, the CSA/D CCR rate constant in the  $R_{1\rho}$  experiment, which we term here  $\Gamma_{1\rho}^{\text{CSA/D}}$ . Figure 2 establishes that the differential relaxation rate constant depends on the amplitude and time scale of the motion, as expected for a CCR, and on the difference in the RF field strength and MAS frequency, again as expected.<sup>1,23</sup> It exhibits a maximum when the motion is on the time scale of microseconds to tens of microseconds, in a manner that depends on the spin-lock RF field strength (Figure 2E,F).

For completeness and illustration purposes, we also provide the analytical expression for the CCR rate constant under spin-lock,  $\Gamma_{1\rho}^{\text{CSA/D}}$ , in the framework of Redfield theory (which is only approximately valid<sup>1</sup>)



**Figure 2.** Simulated relaxation rate constants of a  $^{15}\text{N}$  spin due to its CSA tensor and  $^1\text{H}$ – $^{15}\text{N}$  dipolar coupling, within the two-site jump model. Rate constants for the individual doublet components (A, B), the average rate constant  $R_{1p}$  (C), and the CSA/D CCR rate constant (D), as a function of the jump angle and jump rate constant. (E) shows  $R_{1p}$  and CCR rate constants for a constant jump angle ( $20^\circ$ ) as a function of the jump rate constant, assuming a MAS frequency of 40 kHz and a RF field strength of 12 kHz. (F) uses parameters as in (E) but with an RF field strength of 32 kHz. (G) Correlation of  $R_{1p}$  and CCR rate constants, assuming a jump angle of  $5^\circ$  and a MAS frequency of 40 kHz and a RF field strength of 12 kHz. The jump angle can be related to the often-used order parameter  $S^2$  as  $S^2 = 1/4(3 \cos^2(\theta) + 1)$ .<sup>34</sup> According to this equation, jump angles of 5, 10, 15, 20,  $25^\circ$ , for example, correspond to  $S^2 = 0.994, 0.977, 0.95, 0.91, 0.866$ , respectively.

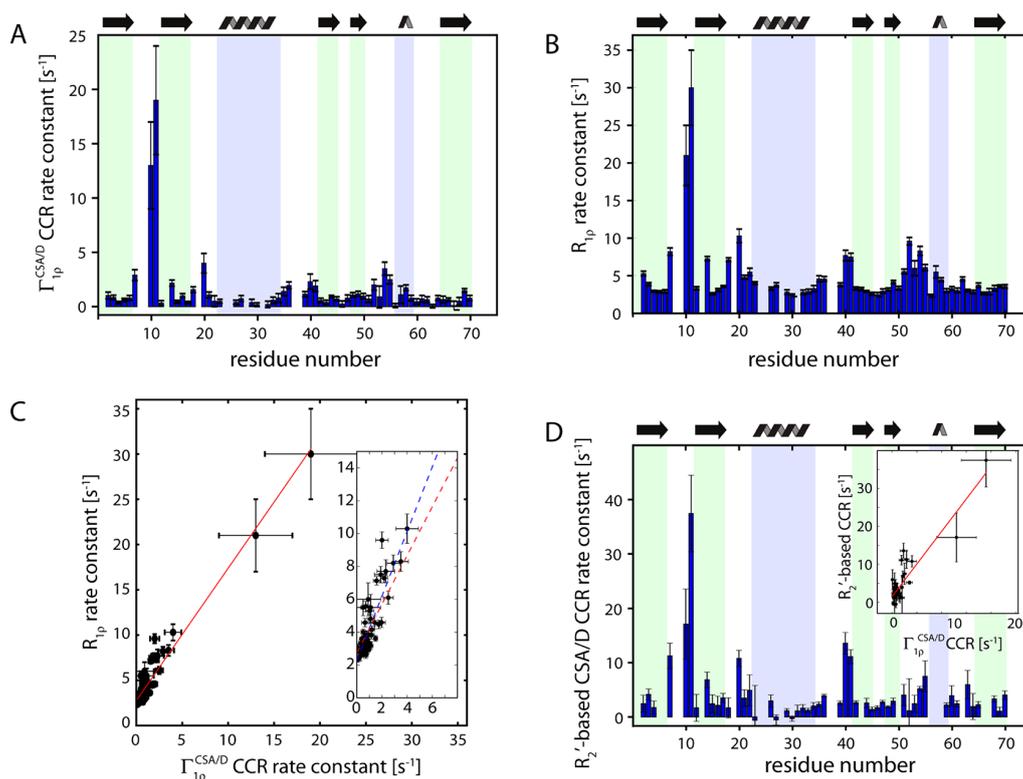
$$\Gamma_{1p}^{\text{CSA/D}} = \delta_{\text{NH}} \cdot \delta_{\text{CSA}} \cdot P_2(\cos(\theta)) \left[ \frac{3}{8} J(\omega_{\text{N}}) + \frac{1}{12} \cdot J(\omega_{\text{RF}} - 2\omega_{\text{MAS}}) + \frac{2}{12} \cdot J(\omega_{\text{RF}} - \omega_{\text{MAS}}) + \frac{1}{12} \cdot J(\omega_{\text{RF}} + 2\omega_{\text{MAS}}) + \frac{2}{12} \cdot J(\omega_{\text{RF}} + \omega_{\text{MAS}}) \right] \quad (1)$$

Here,  $\delta_{\text{NH}}$  is the anisotropy of the dipolar coupling tensor,  $\delta_{\text{CSA}}$  is the CSA, and  $P_2(\cos(\theta))$  is the second Legendre polynomial involving the angle between the principal axes of the dipolar and CSA interactions,  $\theta$ . It is interesting to compare this equation to the expression of the  $R_{1p}$  relaxation rate constant due to the dipolar interaction and CSA, which is given as:<sup>1,23</sup>

$$R_{1p} = R_1 + \sin^2 \theta_{\text{eff}} \left[ R_{1\Delta} - \frac{1}{2} R_1 \right] \quad \text{where} \quad (2)$$

$$R_{1\Delta} = \left\{ \frac{\delta_{\text{NH}}}{4} \right\}^2 \left[ 3 \cdot J(\omega_{\text{N}}) + \frac{1}{3} \cdot J(\omega_{\text{RF}} - 2\omega_{\text{MAS}}) + \frac{2}{3} \cdot J(\omega_{\text{RF}} - \omega_{\text{MAS}}) + \frac{1}{3} \cdot J(\omega_{\text{RF}} + 2\omega_{\text{MAS}}) + \frac{2}{3} \cdot J(\omega_{\text{RF}} + \omega_{\text{MAS}}) \right] + \delta_{\text{CSA}}^2 \left[ \frac{1}{3} \cdot J(\omega_{\text{RF}} - 2\omega_{\text{MAS}}) + \frac{2}{3} \cdot J(\omega_{\text{RF}} - \omega_{\text{MAS}}) + \frac{1}{3} \cdot J(\omega_{\text{RF}} + 2\omega_{\text{MAS}}) + \frac{2}{3} \cdot J(\omega_{\text{RF}} + \omega_{\text{MAS}}) \right]$$

A comparison of (1) and (2) reveals that the two relaxation rate constants provide similar information. In this work, we are interested primarily in motions slower than hundreds of

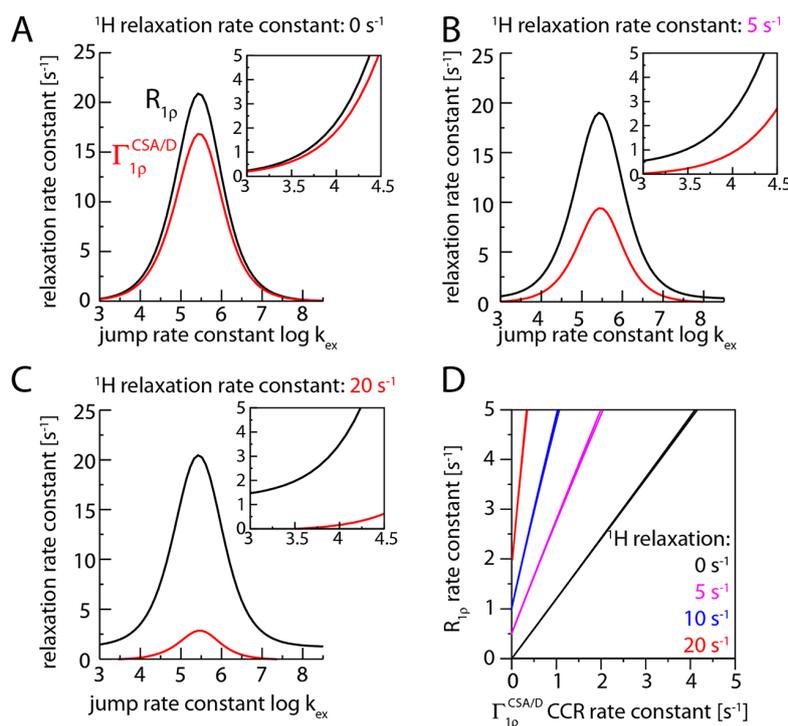


**Figure 3.** Experimental observation of CSA/D CCR in  $^{15}\text{N}$  spin-lock relaxation experiments in deuterated, 60% back-exchanged ubiquitin, obtained at a MAS frequency of 38 kHz. (A) Differential relaxation rate constants as a function of the residue number, determined from the relaxation rate constant of the two doublet components. (B)  $R_{1\rho}$  relaxation rate, that is, obtained from the average of the two rate constants. In (C), a correlation of the two rate constants is shown. The red line corresponds to a correlation obtained over all available residues, with a slope of 1.48 and an intercept of  $2.74\text{ s}^{-1}$ . The inset focuses on the data points at low rate constants. The blue line has a slope of  $1.96\text{ s}^{-1}$ . (D) shows previously published<sup>27</sup> CSA/D CCR relaxation rate constants, obtained with a spin-echo (delay –  $\pi$  pulse – delay) sequence instead of a spin-lock, at the same  $B_0$  field strength. The inset shows the correlation of these data with the  $R_{1\rho}$ -based CCR rate constants of (A).

nanoseconds (because in this range  $\Gamma_{1\rho}^{\text{CSA/D}}$  and  $R_{1\rho}$  are largest, as shown below); in this regime, one can neglect the  $J(\omega_N)$  term; furthermore,  $R_1$  is very small compared to  $R_{1\rho}$ . Neglecting thus the  $J(\omega_N)$  terms, one can see that the two rate constants depend on the same spectral density functions (see eqs 1 and 2). We thus expect that the two rate constants are correlated, an observation that is indeed confirmed both experimentally and through simulations (discussed further below, Figures 2G and 3C). In the rest of this article, we do not use analytical calculations but use exclusively numerical simulations, because this allows taking into account additional factors, such as relaxation of the proton spin due to external mechanisms, as described below. By using numerical simulations, we also circumvent the fact that the Redfield theory is outside its range of validity for slow motions (microseconds).<sup>1</sup>

**Experimental Detection of CSA/D CCR.** To experimentally verify the effects of CCR in  $R_{1\rho}$  experiments, we prepared a sample of  $^2\text{H}, ^{15}\text{N}$ -labeled microcrystalline ubiquitin. The protein was crystallized in a buffer that results in a reprotonation of exchangeable sites, such as the amide sites that we are interested in here, at a level of 60%. In such highly deuterated samples, narrow  $^1\text{H}$  line widths can be achieved, and the simplicity of the spin system, which can be approximated as a two-spin system as far as  $^{15}\text{N}$  relaxation is concerned, also allows for straightforward analysis of relaxation data.<sup>24–26</sup> Of particular relevance here, narrow  $^{15}\text{N}$  lines are obtained even without  $^1\text{H}$  decoupling and therefore one can separately detect

the decay of the doublet components. Figure 3 shows such  $^{15}\text{N}$   $R_{1\rho}$  data for the  $^{15}\text{N}$  doublet components of amide sites in ubiquitin, obtained at a MAS frequency of 38 kHz, and a  $^{15}\text{N}$  RF field strength of 12 kHz. For the majority of residues, the decay rate constants of the two components are detectably different. The differential relaxation rate constant of the two components,  $\Gamma_{1\rho}^{\text{CSA/D}} = \{R_{1\rho}([N_x\text{H}^\alpha]) - R_{1\rho}([N_x\text{H}^\beta])\}/2$ , varies across the protein sequence, ranging from close to zero to about  $20\text{ s}^{-1}$  for different residues (Figure 3A). Residues located in secondary structure elements have small values, below  $1\text{ s}^{-1}$  in most cases. Largest values of  $\Gamma_{1\rho}^{\text{CSA/D}}$ , exceeding  $10\text{ s}^{-1}$ , are observed for residues G10 and K11 in the  $\beta 1$ – $\beta 2$  loop, which has been demonstrated before to undergo nanosecond motion in the microcrystals we are studying.<sup>27,28</sup> CCR rates are furthermore enhanced for residues located in the loop connecting  $\beta 2$  and ubiquitin's  $\alpha$ -helix, the  $\alpha$ - $\beta 3$  loop, and the loop connecting  $\beta 4$  to the short  $3_{10}$  helix. These loop regions have also elevated  $R_{1\rho}$  rate constants (Figure 3B). Figure 3C shows the correlation of residue-wise  $\Gamma_{1\rho}^{\text{CSA/D}}$  and  $R_{1\rho}$  rate constants. Because the majority of residues are located in secondary structure elements, residue-wise variations are generally small. Basically, only residues located in loop regions have significantly enhanced  $\Gamma_{1\rho}^{\text{CSA/D}}$  and  $R_{1\rho}$  rate constants. This limits the precision with which the correlation can be determined. The slope of the correlation is 1.48 when taking into account all residues (red line) and 1.96 when neglecting the two data points with largest rate constants. These



**Figure 4.** Effect of proton relaxation on  $^{15}\text{N}$  transverse relaxation parameters, from numerical simulations with an ad hoc added random field fluctuation leading to proton relaxation. Shown are the simulated relaxation rate constants of  $N_x$  (i.e., the relaxation with a  $^1\text{H}$   $\pi$  pulse in the center of the spin-lock relaxation period),  $R_{1\rho}$ , and the CCR rate constant,  $\Gamma_{1\rho}^{\text{CSA/D}}$ , obtained from the differences of relaxation of the two doublet components. Different rate constants of the  $^1\text{H}$  relaxation are assumed:  $0\text{ s}^{-1}$  (A),  $5\text{ s}^{-1}$  (B), and  $20\text{ s}^{-1}$  (C);  $R_{1\rho}$  and  $\Gamma_{1\rho}^{\text{CSA/D}}$  rate constants are shown in black and red, respectively. (D) Correlation of the  $R_{1\rho}$  and  $\Gamma_{1\rho}^{\text{CSA/D}}$  relaxation rate constants for different  $^1\text{H}$  relaxation rate constants. The jump angle was set to  $5^\circ$ , corresponding to an order parameter of  $S^2 = 0.994$ . More simulations with a large motional amplitude are shown in Figure S3.

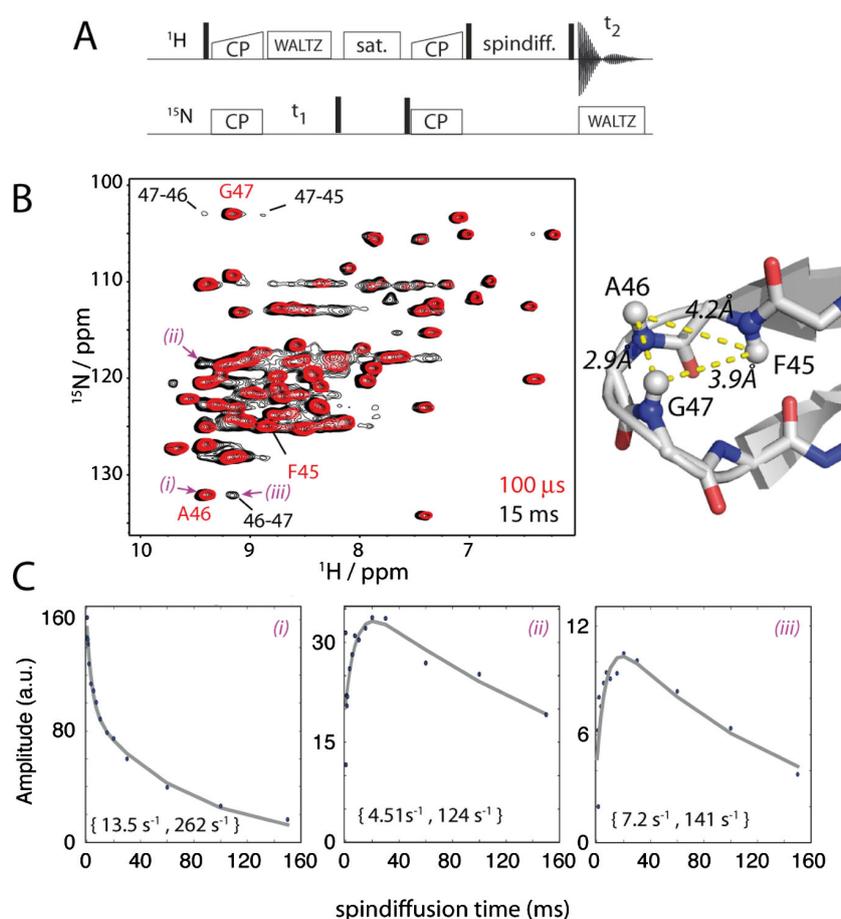
experimental data unequivocally establish that the correlation has a nonzero offset: whereas  $R_{1\rho}$  rate constants exceed  $2\text{ s}^{-1}$  for all residues, many residues have a close-to-zero CSA/D CCR rate constant. This finding cannot be understood on the basis of the simulations in Figure 2. To understand this offset, one needs to take into account partial proton self-decoupling, as shown in the following section.

For completeness, we have also shown previously determined CSA/D CCR rate constants, determined with a spin-echo element instead of a spin-lock period. The drawback of the former experiment is that the decay with free evolution ( $R_2'$ ) is significantly faster than the decay under spin-lock (in ubiquitin by up to 1 order of magnitude). Consequently, it is more difficult to determine small CCR rates from the difference of the two comparably large decay rate constants corresponding to the two doublet components. Qualitatively, the three observables highlight the same residues undergoing slow motions.

**Proton Spin Diffusion and Its Effect on Heteronuclear Transverse Relaxation.** Longitudinal relaxation of the proton spin, or spin diffusion (i.e., magnetization exchange between protons), leads to interconversion of the spin states,  $\alpha$  and  $\beta$ . Accordingly, the states  $N_x[\text{H}^\alpha]$  and  $N_x[\text{H}^\beta]$  interconvert, and this would result in (partial) self-decoupling of these two  $^{15}\text{N}$  doublet components. In the context of CCR, such self-decoupling is an important parameter, because it is expected to reduce the CCR rate constant.<sup>29</sup> In this section, we study the consequences of proton longitudinal relaxation or spin diffusion on heteronuclear relaxation and CCR, using numerical simulations. Different from the simulations shown in Figure

2, we added an ad hoc random field contribution acting on the proton spin. The effect of this random field is to relax the proton spin and therefore interchange spin states  $\text{H}^\alpha$  and  $\text{H}^\beta$ . In principle, one might also attempt to explicitly simulate the proton spin diffusion or relaxation by using a large spin system with multiple protons. In practice, however, such simulations are prohibitively long if the spin system is larger than about 4–5 spins, and the accuracy of the proton evolution is not likely better than that in our ad hoc external relaxation approach.

Figure 4 shows the results of numerical simulations of the heteronuclear  $R_{1\rho}$  experiment in the presence of this additional proton relaxation, simulated here for the case of small-amplitude motion of the HN bond. Simulations with larger amplitude motions are shown in Figure S3. These simulations establish that proton longitudinal relaxation induced by external relaxation mechanisms indeed reduces the CCR rate constant. The reduction of the  $\Gamma_{1\rho}^{\text{CSA/D}}$  rate constant depends on the rate constant of the proton relaxation, as can be seen by comparing the red curves shown in panels (A)–(C). This reduction of  $\Gamma_{1\rho}^{\text{CSA/D}}$  can be understood by the averaging of the two components. Figure 4 also shows that the  $^1\text{H}$  spin flips have an effect on the  $R_{1\rho}$  rate constants. For the range of  $R_{1\rho}$  rate constants that are observed experimentally, that is, in the range of  $R_{1\rho}$  of a few  $\text{s}^{-1}$  to tens of  $\text{s}^{-1}$  at most, the proton relaxation leads to an increase in  $R_{1\rho}$ . The increase in  $R_{1\rho}$  depends on the RF field strength and the MAS frequency (Figure S4). The origin of this increased  $R_{1\rho}$  is the fact that stochastic proton spin flips counteract the MAS-induced recoupling of the dipolar interaction, as discussed in more detail in the Supporting Information. Taken together, the simulations of Figure 4



**Figure 5.** Experimental investigation of proton relaxation and spin diffusion in the deuterated, 60% back-exchanged ubiquitin sample at 38 kHz MAS (14.1 T). (A) Pulse sequence used to detect  $^1\text{H}$  relaxation/spin diffusion. (B) Overlay of two spectra obtained with different mixing times as indicated. Selected amide peaks and the corresponding spin-diffusion cross-peaks are indicated, and the distances between these sites are indicated on the right. (C) Representative time traces observed at peak positions corresponding to amide sites (diagonal peaks), and two-spin-diffusion cross-peaks, obtained at spectral positions at which the intensity is observed only with a mixing time. The gray lines correspond to bi-exponential fits, and the fitted rate constants are indicated in brackets. The identity of the peaks for which the graphs are shown in (C) is indicated in the spectrum with (i)–(iii).

establish that the effect of the additional  $^1\text{H}$  spin relaxation is (1) a reduction of the CCR rate constant and (2) an increase in the  $R_{1\rho}$  rate constant. The decreased CCR and increased  $R_{1\rho}$  rate constants change the slope of the correlation of the two rate constants and result in an offset of the two rate constants (Figure 4D). It is remarkable that this situation is qualitatively found in the experimental data, shown in Figure 3C, which reproduce an offset of  $R_{1\rho}$  rate constants. It shall be noted that Figure 4 assumes a single amplitude and varies the time scale of the motion, a situation that is not likely to apply to the range of amide sites across a protein. Nonetheless, qualitatively the behavior observed in Figure 4 applies irrespectively of the precise motional model.

The amount by which the CCR rate is reduced depends on the rate at which the  $^1\text{H}$  spin flips (Figure 4). We therefore need to estimate which spin diffusion rate constants are realistic in our experimental setting, that is, at a MAS frequency of 38 kHz and with the deuteration level we employed (60% back-exchange). Figure 5 shows experimental measurements of spin diffusion between amide protons. By comparison of the spectra obtained with very short mixing time (100  $\mu\text{s}$ , red) and 15 ms mixing (black), one can observe that there are additional peaks

pointing to  $^1\text{H}$ – $^1\text{H}$  spin diffusion. Figure 5B highlights a few cases of well-resolved cross-peaks, corresponding to spin diffusion in a  $\beta$ -turn region (residues 45–47); for the cases of these cross-peaks, the corresponding amide sites are within  $\sim 4.2$  Å at most. Figure 5C shows the time evolution of representative diagonal- and cross-peaks. The diagonal peaks show nonmonoexponential decay, which we ascribe to contributions of spin diffusion/cross-relaxation (fast component) and relaxation of the magnetization, presumably at the most dynamic sites, or through exchange with solvent. When fitting these curves with bi-exponential functions, respective rate constants are several tens per second for the fast rate (about 30–40% amplitude) and about 5–10 s $^{-1}$  (60–70% amplitude), respectively. On attempting to describe the decay with a single rate constant, approximated by the time point where the signal decays to  $1/e \sim 37\%$ , apparent rate constants of 20 s $^{-1}$  are obtained. The observable cross-peaks show a buildup within tens of milliseconds, similar to the fast rate observed in the diagonal-peak decay curves. Taken together, these data allow one to estimate that proton relaxation occurs at a rate of approximately 10–20 s $^{-1}$ . (We note that the spin-diffusion rate is expected to decrease at higher MAS frequencies

and higher deuteration levels. A very recent report, which became available during the revision of this article,<sup>30</sup> used higher MAS frequencies (60 kHz), but lower deuteration (100% back-protonation of exchangeable sites), and values of about 10–20 s<sup>-1</sup> were reported for the majority of the residues. Whereas this is in rather good agreement with the present estimates, no apparent multiexponential decay has been reported there, nor did the authors report on spin-diffusion cross-peaks.)

Using this estimate of <sup>1</sup>H relaxation, one can simulate its effect on measured  $\Gamma_{1\rho}^{\text{CSA/D}}$  and  $R_{1\rho}$  rate constants, and possibly correct for them. It has been proposed in solution-state NMR that the <sup>1</sup>H–<sup>1</sup>H nuclear Overhauser effect (i.e., relaxation-induced spin diffusion) can be taken into account when analyzing CCR,<sup>31</sup> using an analytical approach. In principle, such an approach could also be applied in solids. However, the effect of the proton spin flips on the <sup>15</sup>N relaxation depends on the RF field and the MAS frequency, and an appropriate formalism is currently not developed. However, one may resort to numerical simulations, such as the ones we use here, to evaluate the effects of proton spin flips on heteronuclear  $R_{1\rho}$  or CCR rate constants. Experimentally determined proton decay rate constants, or at least global estimates, may be used to this end.

We illustrate this with the data shown in Figure 4. For example, assume that the experimentally determined  $R_{1\rho}$  rate constant is 3 s<sup>-1</sup>, that is, including the effects of proton relaxation. The actual  $R_{1\rho}$  rate constant, corrected for a <sup>1</sup>H relaxation rate constant of 20 s<sup>-1</sup> (10 s<sup>-1</sup>), is 1.8 s<sup>-1</sup> (2.6 s<sup>-1</sup>), that is, the excess decay induced by the <sup>1</sup>H relaxation is of the order of 0.4–1.2 s<sup>-1</sup>. In principle, one can take into account these effects, and use the corrected values when fitting motional models to transverse relaxation data.

A similar estimation for  $\Gamma_{1\rho}^{\text{CSA/D}}$  suggests, for example, that in the presence of a <sup>1</sup>H relaxation with a rate constant of 10 s<sup>-1</sup>, a measured  $\Gamma_{1\rho}^{\text{CSA/D}}$  of 1 s<sup>-1</sup> points to an actual  $\Gamma_{1\rho}^{\text{CSA/D}}$  rate constant of 3.6 s<sup>-1</sup>. Arguably, it is more difficult to correct  $\Gamma_{1\rho}^{\text{CSA/D}}$  for <sup>1</sup>H relaxation effects, because the values of this rate constant are smaller than those of  $R_{1\rho}$ , and <sup>1</sup>H relaxation has a stronger effect on  $\Gamma_{1\rho}^{\text{CSA/D}}$  than on  $R_{1\rho}$ .

It is noteworthy that such a correction appears more difficult for the CCR rate constant, which may become very small and thus error prone, than for  $R_{1\rho}$ . Of course, one may attempt to suppress the effects of proton spin flips even more, by using higher MAS frequencies and/or higher deuteration levels. Increasing the MAS frequency will, however, not suppress proton magnetization decay, which is actually due to dynamics or chemical exchange, such that it appears unlikely that one may suppress its effects completely.

**Suppression of the Effect of CSA/D CCR in  $R_{1\rho}$  Experiments.** Finally, we note that the presence of CSA/D CCR might hamper accurate extraction of  $R_{1\rho}$  rate constants. In solution-state NMR, schemes for suppression of the effect of CSA/D CCR have been reported.<sup>13,32</sup> They generally require irradiation of <sup>1</sup>H spins. Suppression of CCR in  $R_{1\rho}$  experiments, even in solution-state NMR, is complicated by the fact that <sup>1</sup>H RF fields applied during heteronuclear spin-lock may partially reintroduce the scalar-coupling Hamiltonian in solution and solids,<sup>12,13</sup> and it may also reintroduce the dipolar coupling in the solid state. In MAS ssNMR, particular care has to be taken to avoid rotary resonance conditions, that is, <sup>1</sup>H multipulse or CW schemes, which counteract the averaging brought about by MAS. To investigate these effects, we have tested two different

suppression schemes, namely, (1) a single <sup>1</sup>H  $\pi$  pulse, applied in the center of the relaxation period and (2) a train of <sup>1</sup>H  $\pi$  pulses, spaced by 5 ms delays. Both schemes, or variants thereof, are used in solution-state NMR.<sup>12,13,33</sup> Interestingly, we find that in the multipulse scheme, even if the <sup>1</sup>H pulse spacing (5 ms) is very long compared to the rotor period ( $\sim 20\text{--}30\ \mu\text{s}$ ), the heteronuclear coherences decay significantly faster than those in the absence of <sup>1</sup>H pulses or with a single <sup>1</sup>H  $\pi$  pulse (Figure S5). The use of a single <sup>1</sup>H  $\pi$  pulse appears as a straightforward method to suppress the effects of CCR. We note, however, that even with a suppression of CCR, one does not expect monoexponential decay, due to the inherent multiexponential nature of decays in solids, outlined above.

## CONCLUSIONS

We have shown here that CSA/D CCR can be measured in MAS ssNMR of deuterated proteins. This relaxation rate constant reports on dynamics on nanosecond-to-microsecond time scales, as we illustrate with an application to microcrystalline ubiquitin. In many practical applications, however, the more easily accessible  $R_{1\rho}$  rate constant of the decoupled line has advantages over the measurement of CSA/D CCR under spin-lock, in particular due to higher sensitivity and resolution in a decoupled versus a nondecoupled spectrum. We also provided evidence, by numerical simulations, that rapid <sup>1</sup>H spin flips, as mediated by <sup>1</sup>H–<sup>1</sup>H spin diffusion, reduce the CCR rate constant and also contribute to the observed  $R_{1\rho}$  rate constant. This finding also shows that for accurate quantitative analyses of  $R_{1\rho}$  rate constants, one should correct the measured rate constants for the contribution arising from proton relaxation. The present report of CCR adds to the increasing arsenal of techniques to study protein dynamics on long time scales by MAS ssNMR and stresses the importance of considering hitherto ignored effects when interpreting relaxation rate constants in ssNMR.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.6b06129.

Figures showing the pulse sequences used in this study; figures with additional simulations, and experimental data; extended discussion about the increase in rate constants due to proton spin relaxation (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: dominique.marion@ibs.fr (D.M.).

\*E-mail: paul.schanda@ibs.fr (P.S.).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was financially supported by the European Research Council (ERC-StG-311318 ProtDyn2Function) and the Commissariat à l'énergie atomique et aux énergies alternatives (CEA), the Centre National de la Recherche Scientifique (CNRS), and the Université Grenoble Alpes. This work used the platforms of the Grenoble Instruct Center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology

(PSB). We thank Matthias Ernst for stimulating discussions and for assistance with implementing the proton random field fluctuations in the  $\gamma$  simulations.

## REFERENCES

- Schanda, P.; Ernst, M. Studying Dynamics by Magic-Angle Spinning Solid-State NMR Spectroscopy: Principles and Applications to Biomolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *96*, 1–46.
- Krushelnitsky, A.; Zinkevich, T.; Reichert, D.; Chevelkov, V.; Reif, B. Microsecond Time Scale Mobility in a Solid Protein as Studied by the  $^{15}\text{N}$   $R_{1\rho}$  Site-Specific NMR Relaxation Rates. *J. Am. Chem. Soc.* **2010**, *132*, 11850–11853.
- Lewandowski, J. R.; Sass, H. J.; Grzesiek, S.; Blackledge, M.; Emsley, L. Site-Specific Measurement of Slow Motions in Proteins. *J. Am. Chem. Soc.* **2011**, *133*, 16762–16765.
- Ma, P.; Haller, J. D.; Zajackala, J.; Macek, P.; Sivertsen, A. C.; Willbold, D.; Boisbouvier, J.; Schanda, P. Probing Transient Conformational States of Proteins by Solid-State R1rho Relaxation-Dispersion NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2014**, *53*, 4312–4317.
- Lamley, J. M.; Lougher, M. J.; Sass, H.-J.; Rogowski, M.; Grzesiek, S.; Lewandowski, J. R. Unraveling the Complexity of Protein Backbone Dynamics with Combined  $^{13}\text{C}$  and  $^{15}\text{N}$  Solid-state NMR Relaxation Measurements. *Phys. Chem. Chem. Phys.* **2015**, *17*, 21997–22008.
- Lamley, J. M.; Öster, C.; Stevens, R. A.; Lewandowski, J. R. Intermolecular Interactions and Protein Dynamics by Solid-State NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2015**, *54*, 15374–15378.
- Ma, P.; Xue, Y.; Coquelle, N.; Haller, J. D.; Yuwen, T.; Ayala, I.; Willbold, D.; Colletier, J.; Skrynnikov, N. R. Observing the Overall Rocking Motion of a Protein in a Crystal. *Nat. Commun.* **2015**, *6*, No. 8361.
- Quinn, C. M.; McDermott, A. E. Quantifying Conformational Dynamics Using Solid-State  $R_{1\rho}$  Experiments. *J. Magn. Reson.* **2012**, *222*, 1–7.
- Quinn, C. M.; McDermott, A. E. Monitoring Conformational Dynamics with Solid-State  $R_{1\rho}$  Experiments. *J. Biomol. NMR* **2009**, *45*, 5–8.
- Chevelkov, V.; Diehl, A.; Reif, B. Quantitative Measurement of Differential  $^{15}\text{N}$ -H A/ $\beta$  T2 Relaxation Rates in a Perdeuterated Protein by MAS Solid-State NMR Spectroscopy. *Magn. Reson. Chem.* **2007**, *45*, S156–S160.
- Skrynnikov, N. Asymmetric Doublets in MAS NMR: Coherent and Incoherent Mechanisms. *Magn. Reson. Chem.* **2007**, *45*, S161–S173.
- Massi, F.; Johnson, E.; Wang, C.; Rance, M.; Palmer, A. G. NMR  $R_{1\rho}$  Rotating-Frame Relaxation with Weak Radio Frequency Fields. *J. Am. Chem. Soc.* **2004**, *126*, 2247–2256.
- Korzhev, D. M.; Skrynnikov, N. R.; Millet, O.; Torchia, D. A.; Kay, L. E. An NMR Experiment for the Accurate Measurement of Heteronuclear Spin-Lock Relaxation Rates. *J. Am. Chem. Soc.* **2002**, *124*, 10743–10753.
- Chevelkov, V.; Reif, B. TROSY Effects in MAS Solid-State NMR. *Concepts Magn. Reson., Part A* **2008**, *32A*, 143–156.
- Igumenova, T. I.; McDermott, A. E.; Zilm, K. W.; Martin, R. W.; Paulson, E. K.; Wand, A. J. Assignments of Carbon NMR Resonances for Microcrystalline Ubiquitin. *J. Am. Chem. Soc.* **2004**, *126*, 6720–6727.
- Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPIPE - a Multidimensional Spectral Processing System Based on Unix Pipes. *J. Biomol. NMR* **1995**, *6*, 277–293.
- Smith, S.; Levante, T.; Meier, B.; Ernst, R. Computer Simulations in Magnetic Resonance. An Object-Oriented Programming Approach. *J. Magn. Reson., Ser. A* **1994**, *106*, 75–105.
- Abergel, D.; Palmer, A. G. On the Use of the Stochastic Liouville Equation in Nuclear Magnetic Resonance: Application to  $R_{1\rho}$  Relaxation in the Presence of Exchange. *Concepts Magn. Reson., Part A* **2003**, *19A*, 134–148.
- Kubo, R. Stochastic Liouville Equations. *J. Math. Phys.* **1963**, *4*, 174–183.
- Redfield, A. G. On the Theory of Relaxation Processes. *IBM J. Res. Dev.* **1957**, *1*, 19–31.
- Torchia, D. A.; Szabo, A. Spin-Lattice Relaxation in Solids. *J. Magn. Reson. (1969)* **1982**, *49*, 107–121.
- Giraud, N.; Blackledge, M.; Goldman, M.; Böckmann, A.; Lesage, A.; Penin, F.; Emsley, L. Quantitative Analysis of Backbone Dynamics in a Crystalline Protein from Nitrogen-15 Spin-Lattice Relaxation. *J. Am. Chem. Soc.* **2005**, *127*, 18190–18201.
- Kurbanov, R.; Zinkevich, T.; Krushelnitsky, A. The Nuclear Magnetic Resonance Relaxation Data Analysis in Solids: General R1/R1rho Equations and the Model-Free Approach. *J. Chem. Phys.* **2011**, *135*, No. 184104.
- Reif, B. Ultra-High Resolution in MAS Solid-State NMR of Perdeuterated Proteins: Implications for Structure and Dynamics. *J. Magn. Reson.* **2012**, *216*, 1–12.
- Schanda, P.; Huber, M.; Boisbouvier, J.; Meier, B. H.; Ernst, M. Solid-State NMR Measurements of Asymmetric Dipolar Couplings Provide Insight into Protein Side-Chain Motion. *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 11005–11009.
- Knight, M. J.; Webber, A. L.; Pell, A. J.; Guerry, P.; Barbet-Massin, E.; Bertini, L.; Felli, I. C.; Gonnelli, L.; Pierattelli, R.; Emsley, L.; et al. Fast Resonance Assignment and Fold Determination of Human Superoxide Dismutase by High-Resolution Proton-Detected Solid-State MAS NMR Spectroscopy. *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 11697–11701.
- Schanda, P.; Meier, B. H.; Ernst, M. Quantitative Analysis of Protein Backbone Dynamics in Microcrystalline Ubiquitin by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **2010**, *132*, 15957–15967.
- Haller, J. D.; Schanda, P. Amplitudes and Time Scales of Picosecond-to-Microsecond Motion in Proteins Studied by Solid-State NMR: A Critical Evaluation of Experimental Approaches and Application to Crystalline Ubiquitin. *J. Biomol. NMR* **2013**, *57*, 263–280.
- Hu, K.; Vögeli, B.; Clore, G. M. Interference between Transverse Cross-Correlated Relaxation and Longitudinal Relaxation Affects Apparent J-Coupling and Transverse Cross-Correlated Relaxation. *Chem. Phys. Lett.* **2006**, *423*, 123–125.
- Smith, A. A.; Testori, E.; Cadalbert, R.; Meier, B. H.; Ernst, M. Characterization of Fibril Dynamics on Three Timescales by Solid-State NMR. *J. Biomol. NMR* **2016**.
- Reif, B.; Diener, A.; Hennig, M.; Maurer, M.; Griesinger, C. Cross-Correlated Relaxation for the Measurement of Angles between Tensorial Interactions. *J. Magn. Reson.* **2000**, *143*, 45–68.
- Massi, F.; Johnson, E.; Wang, C.; Rance, M.; Palmer, A. G. NMR R1rho Rotating-Frame Relaxation with Weak Radio Frequency Fields. *J. Am. Chem. Soc.* **2004**, *126*, 2247–2256.
- Palmer, A. G.; Massi, F. Characterization of the Dynamics of Biomacromolecules Using Rotating-Frame Spin Relaxation NMR Spectroscopy. *Chem. Rev.* **2006**, *106*, 1700–1719.
- Clore, G. M.; Driscoll, P. C.; Wingfield, P. T.; Gronenborn, A. M. Analysis of the Backbone Dynamics of Interleukin-1 Beta Using Two-Dimensional Inverse Detected Heteronuclear  $^{15}\text{N}$ - $^1\text{H}$  NMR Spectroscopy. *Biochemistry* **1990**, *29*, 7387–7401.



Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

## Sensitive proton-detected solid-state NMR spectroscopy of large proteins with selective CH<sub>3</sub> labelling: application to the 50S ribosome subunit

Vilius Kurauskas,<sup>a,b,c</sup> Elodie Crublet<sup>\*,a,b,c,e</sup>, Pavel Macek,<sup>a,b,c</sup> Rime Kerfah,<sup>d</sup> Diego F. Gauto,<sup>a,b,c</sup> Jérôme Boisbouvier<sup>a,b,c</sup> and Paul Schanda<sup>\*,a,b,c</sup>

**Solid-state NMR spectroscopy allows the characterization of structure, interactions and dynamics of insoluble and/or very large proteins. Sensitivity and resolution are often major challenges for obtaining atomic-resolution information, in particular for very large protein complexes. Here we show that the use of deuterated, specifically CH<sub>3</sub>-labelled proteins result in significant sensitivity gains compared to previously employed CHD<sub>2</sub> labelling, while line widths only marginally increase. We apply this labelling strategy to a 468 kDa-large dodecameric aminopeptidase, TET2, and the 1.6 MDa-large 50S ribosome subunit of *Thermus thermophilus*.**

Biomolecular magic-angle spinning (MAS) solid-state NMR (ssNMR) has made substantial progress over the last decade in terms of sensitivity, resolution and information content, and established as an important technique in structural biology, addressing structure and dynamics in systems with increasing complexity.<sup>1-3</sup> The progress was enabled by improved NMR instrumentation, in particular probes allowing fast MAS (up to 100 kHz), improved pulse sequences, as well as better sample preparation and isotope labelling techniques. Experiments based on proton-detection are particularly powerful for the study of protein structure, interactions and dynamics. The large gyromagnetic ratio of protons translates to high intrinsic detection sensitivity, and ability to measure long-range distances. However, the large <sup>1</sup>H gyromagnetic ratio and the corresponding strong <sup>1</sup>H-<sup>1</sup>H dipolar couplings also result in large <sup>1</sup>H line widths. Proton-detected experiments are, therefore, particularly powerful when protons are introduced

only at selected sites of interest in an otherwise deuterated environment, and when high MAS frequencies are employed, leading to efficient averaging of <sup>1</sup>H-<sup>1</sup>H interactions. Methyl groups are particularly interesting reporters of protein structure and dynamics in proteins, due to their widespread abundance in proteins, often located in important hydrophobic core regions<sup>4</sup>. In solution-state NMR, methyl-directed studies have provided a wealth of information about dynamics, interactions and structure in particular for large complexes (up to 1 MDa), because in this case methyls are the only observable NMR reporters.<sup>5,6</sup> Protocols for introducing methyls of all iso-topomers (CH<sub>3</sub>, CH<sub>2</sub>D, CHD<sub>2</sub>) into otherwise deuterated proteins<sup>4</sup> have been established. The three isotopomers differ in the number and nature of interactions, and thus in the spin relaxation properties, sensitivity and resolution. In solution-state NMR, the most commonly employed isotopomer is CH<sub>3</sub>, which generally provides the highest sensitivity due to the three-fold spin multiplicity. As long as optimized pulse schemes are used<sup>7</sup>, the spectra of all isotopomers exhibit similar line widths.<sup>8</sup> In ssNMR, only few studies have reported the use of specifically methyl-labelled samples for studying dynamics and structure.<sup>3,6-8</sup> All these studies have employed CHD<sub>2</sub> labelling, with the exception of a brief note of CH<sub>3</sub> labelling.<sup>9</sup> The primary argument for using the CHD<sub>2</sub> isotopomer is related to the fact that there are no intra-methyl <sup>1</sup>H-<sup>1</sup>H dipolar couplings in CHD<sub>2</sub>, and that these interactions are expected to lead to strong line broadening; this assumption was supported by pioneering experiments on stochastically labelled proteins at low (20 kHz) MAS frequency,<sup>10</sup> but not systematically investigated. Here we show that, in experiments performed at high MAS frequency (50-60 kHz) the line widths in CHD<sub>2</sub> and CH<sub>3</sub> methyl are comparable, and that CH<sub>3</sub> labelling results in improved sensitivity compared to CHD<sub>2</sub>. The enhanced sensitivity opens possibilities to study even very large objects, as demonstrated with the 1.6 MDa-large 50S ribosome, where sensitivity is a great challenge.

- a. Université Grenoble Alpes, Institut de Biologie Structurale, Grenoble.  
b. CEA, Institut de Biologie Structurale, F-38044 Grenoble, France.  
c. CNRS, Institut de Biologie Structurale, F-38044 Grenoble France.  
d. NMR-Bio. 5 place Robert Schumann, F-38025 Grenoble, France.  
e. Present address: NMR-Bio. 5 place Robert Schumann, F-38044 Grenoble, France.

e-mail: [crublet@nmr-bio.com](mailto:crublet@nmr-bio.com) and [schanda@ibs.fr](mailto:schanda@ibs.fr)

Electronic Supplementary Information (ESI) available: comprehensive experimental details. See DOI: 10.1039/x0xx00000x

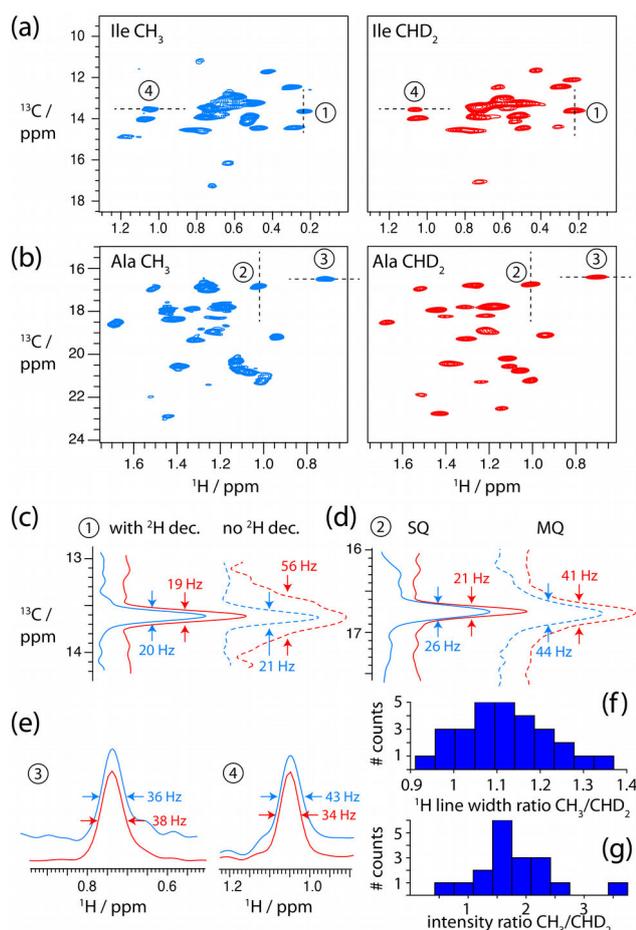


Fig. 1. Proton-detected 2D methyl  $^1\text{H}$ - $^{13}\text{C}$  spectra of  $\text{CH}_3$  (blue) and  $\text{CHD}_2$  (red) methyls of (a) isoleucines and (b) alanines in TET2, obtained at 14.1 T (600 MHz  $^1\text{H}$  Larmor frequency). These spectra were obtained with the CP-based pulse sequence shown in ESI Figure S1a. Traces along the positions indicated with dashed lines and numbers are shown in (c)-(e). Dashed 1D traces were obtained at the indicated positions in the spectra, but different pulse sequences, either omitting  $^2\text{H}$  decoupling, or using MQ evolution (see Figure S1). Line-widths (full-width at half maximum) are indicated for these selected peaks. (f) shows the ratio of line widths observed for  $\text{CH}_3$  and  $\text{CHD}_2$  correlations, merged for both Ala and Ile samples. (g) Intensity ratios for cross-peaks observed in the Ile spectrum, taking into account the protein amounts in  $\text{CH}_3/\text{CHD}_2$  samples. For all these spectra, resolution enhancement was applied using a 70-degree shifted sine-bell square function (90 degrees would correspond to a cosine, i.e. no resolution enhancement), see ESI.

In order to investigate the properties of the different methyl isotopomers in ssNMR, we compared four samples of the enzyme TET2 from *Pyrococcus horikoshii*, a dodecameric (12x39 kDa) aminopeptidase involved in the cellular protein quality control, with different methyl labelling schemes. (Details about sample preparation can be found in the ESI.) With 353 residues per monomer and a total oligomer mass of 468 kDa, TET2 is among the largest proteins studied so far by ssNMR, making it thus a considerable challenge in terms of resolution and sensitivity. In the samples we incorporated either  $\text{CH}_3$  or  $\text{CHD}_2$  methyl groups on isoleucine residues or on alanines, respectively, in an otherwise deuterated background. The different lengths and flexibility of Ala and Ile side chains

generally lead to different dynamics, reflecting the range of methyl cases encountered in proteins. Figure 1a,b shows comparisons of  $^1\text{H}$ - $^{13}\text{C}$  correlation spectra of these four samples, obtained with cross-polarization transfer steps at a MAS frequency of 57 kHz. High-resolution spectra are obtained for this very large protein assembly for both isotopomers and both residue types. The spectra are similar to solution-state spectra of TET, showing that structure and dynamics in our solid-state preparations are conserved in the solid state, and allowing us to transfer previously known assignments from solution-state NMR<sup>11</sup> (see Supporting Information Figure S2). In  $\text{CHD}_2$  groups, the observed line widths are of the order of 30 – 45 Hz ( $^1\text{H}$ ) and 20-30 Hz ( $^{13}\text{C}$ ), respectively (Figure 1c-e). These values are in good agreement with previously reported line widths obtained in fast-MAS solid-state NMR experiments on smaller model proteins with  $\text{CHD}_2$  labelling<sup>9,12,13</sup>. For the  $\text{CH}_3$  isotopomer, the line widths along the  $^{13}\text{C}$  dimension are similar to the ones observed with  $\text{CHD}_2$  groups (20-30 Hz), suggesting that the intra-methyl heteronuclear  $^1\text{H}$ - $^{13}\text{C}$  dipolar coupling is not the dominant factor for line width under the chosen conditions (heteronuclear decoupling, fast MAS). This is expected, given that fast-MAS, the high deuteration level of the surroundings, and the 3-fold reduced dipolar  $^1\text{H}$ - $^{13}\text{C}$  coupling that results from the rapid methyl rotation, make decoupling very efficient. Remarkably, the  $^1\text{H}$  line widths in the  $\text{CH}_3$  sample are only slightly larger than the ones in  $\text{CHD}_2$  samples (Figure 1e,f). The fact that reducing the number of methyl proton spins from three ( $\text{CH}_3$ ) to one ( $\text{CHD}_2$ ) has only a modest effect on line widths suggests that the intra-methyl dipolar couplings are not determining the  $^1\text{H}$  line widths. Sample inhomogeneity, dipolar interactions to remote protons as well as dynamics may be important contributions to line widths, such that intra-methyl  $^1\text{H}$ - $^1\text{H}$  dipolar interactions do not dominate. Of note, given that the line widths observed here are similar to microcrystalline samples,<sup>9,12</sup> our conclusions hold also for those systems. In more heterogeneous systems (fibrils, membrane proteins) the differences in line widths between the two isotopomers are expected to be similar or smaller than here, because line widths due to sample heterogeneity is independent of the labelling scheme. Importantly, in terms of sensitivity the  $\text{CH}_3$  labelling scheme outcompetes  $\text{CHD}_2$  labelling, on average by a factor of about 1.9 (Figure 1g). This is due to the proton spin multiplicity (expected gain of 3), reduced by the somewhat faster relaxation of  $\text{CH}_3$  methyls (see below). The experiments shown in Figure 1a,b were recorded with  $^2\text{H}$  decoupling during chemical-shift evolution periods, which requires a probehead able to tune to  $^2\text{H}$  frequency. Such equipment is not standard in contemporary MAS probes. As expected, while omission of  $^2\text{H}$  decoupling has virtually no effect on line widths and sensitivity in  $\text{CH}_3$ -labelled methyls (Figure 1c), due to the absence of  $^1\text{J}_{\text{CD}}$  couplings, in  $\text{CHD}_2$  groups the undecoupled  $^{13}\text{C}$  line widths are dramatically increased due to the scalar coupling ( $\sim 20$

Hz) multiplet structure.<sup>12</sup> Therefore, if  $^2\text{H}$  decoupling is impossible with the available hardware, the  $\text{CH}_3$  isotopomer is clearly favoured in terms of resolution and sensitivity. The effects of  $^2\text{H}$  decoupling on  $^1\text{H}$  line widths are negligible, as the  $^2\text{J}_{\text{HD}}$  coupling is small ( $\sim 1\text{--}3\text{ Hz}$ ).<sup>12</sup> We have also investigated two different types of  $^{13}\text{C}$  evolution, single-quantum (SQ), and multiple-quantum (MQ) evolution. In solution-state NMR, in the case of slow overall molecular tumbling, MQ coherence life times are significantly longer than SQ life times<sup>14</sup>. In contrast, Figure 1d establishes that MQ lines in our ssNMR experiments are almost two-fold larger than their SQ counterparts. This can be understood by considering that due to the absence of overall tumbling, cross-correlated relaxation is small (reducing the potential benefit of MQ evolution), and dephasing due to remote protons outweighs the potential gains from cross-correlated relaxation. We also investigated the relative sensitivity obtained with INEPT-transfer and CP-transfer. INEPT-based transfer is favoured by large-scale dynamics and extensive deuteration, because both factors increase coherence life times and thus INEPT transfer efficiency. We find indeed that for the more dynamic Ile residues INEPT outcompetes CP-transfer, while for Ala we find similar sensitivity (Figure S4).

Line widths are related to the life times of spin coherences, which are determined by spin relaxation (i.e., dynamics) and dephasing that is due to incompletely averaged dipolar interactions; the latter contribution depends on the MAS frequency. To further gain insight into the properties of the different labellings, we performed systematic analyses of apparent relaxation rate constants. These measurements, shown in Figure 2, corroborate the trends shown in Figure 1: (i)  $^1\text{H } R_2'$  relaxation rates are similar in  $\text{CH}_3$  and  $\text{CHD}_2$  samples, with values of  $\sim 100\text{ s}^{-1}$  and  $\sim 80\text{ s}^{-1}$ , respectively, at high MAS frequency (55–57 kHz). The expected line widths from these numbers ( $R_2'/\pi$ ,  $\sim 25\text{--}30\text{ Hz}$ ) are smaller than the experimentally observed values ( $\sim 35\text{--}45\text{ Hz}$ ), showing that part of the line width originates from sample and/or magnetic-field heterogeneity. (ii)  $^{13}\text{C } R_2'$  of  $\sim 20\text{--}40\text{ s}^{-1}$  are found at high MAS frequencies ( $\sim 55\text{ kHz}$ ). The line widths predicted from these numbers (6–12 Hz) are significantly smaller than the experimentally observed line widths in this study and crystalline ubiquitin before<sup>12</sup> ( $\sim 18\text{--}25\text{ Hz}$ ), showing that heterogeneity limits the resolution of  $^{13}\text{C}$  lines. (iii) Omission of  $^2\text{H}$  decoupling leads to significantly shorter  $^{13}\text{C } T_2'$ , in  $\text{CHD}_2$  (but not  $\text{CH}_3$ ), in agreement with the larger line widths (Figure 1c) (iv) MQ coherences decay faster than their SQ counterpart, again in line with data shown in Figure 1d. (v) Furthermore, proton  $R_1$  rate constants are  $\sim 15\%$  higher in  $\text{CH}_3$  compared to  $\text{CHD}_2$ . This allows higher repetition rate of experiments, and thus slightly increased sensitivity for  $\text{CH}_3$  samples (see Figure S4). (vi) We also measured  $^{13}\text{C } R_1$  and  $^{13}\text{C } R_{1p}$  rate constants, and find that they are not MAS-frequency dependent. This finding establishes that coherent dephasing (i.e. contribution to the decay that is not due to dynamics, and

which would be MAS dependent) is negligible, allowing quantitative interpretation of methyl dynamics. While  $^{13}\text{C } R_1$  has been used for this purpose before,<sup>13</sup> the possibility to quantitatively interpret  $R_{1p}$  enables the study of slow (ns) motion.<sup>15</sup>

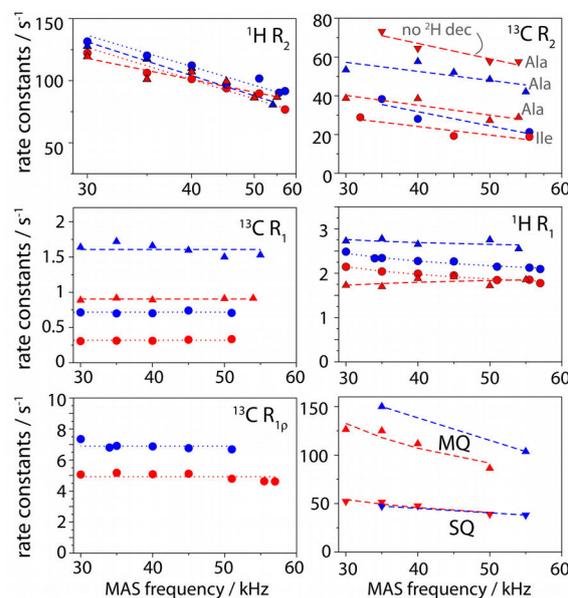


Fig. 2. MAS-frequency dependence of various spin relaxation parameters in Ile-labelled (circles) and Ala-labelled (triangles) TET2, with  $\text{CH}_3$  (blue) and  $\text{CHD}_2$  (red) isotopomers. Bulk methyl relaxation rate constants were obtained from a series of 1D spectra, using pulse sequences shown in ESI Figure S1.

Having thus established  $\text{CH}_3$  labelling provides enhanced sensitivity compared to  $\text{CHD}_2$ , with similar resolution, we turned to a very challenging application, the 50S ribosome particle from *Thermus thermophilus*. Ribosomes are among the most complex biological machineries, composed of RNAs and proteins, with a total molecular mass of 1.6 MDa (for the 50S subunit). Unlike TET2, where the 12-fold symmetry improves sensitivity and reduces spectral complexity, ribosomes are not symmetric. Sensitivity and resolution are thus a great challenge. Earlier studies by solution-state NMR have observed only highly mobile isotope-labelled proteins attached to ribosomes<sup>16</sup>, or flexible nascent polypeptides exiting the ribosome,<sup>17,18</sup> while less flexible proteins are invisible due to unfavourable relaxation properties. Recently, ssNMR spectra have been reported for tightly ribosome-bound proteins, which were isotope labelled and bound to unlabelled 50S and 30S subunits, respectively.<sup>19,20</sup> This selective labelling of a single protein overcomes the severe resonance overlap problem but requires reconstitution of ribosomes from labelled/unlabelled subunits, which is not generally possible. We investigate here the possibility to study uniformly labelled ribosome (50S subunit, 1.6 MDa), obtained from bacterial cells grown in uniformly deuter-

ated, Ala-CH<sub>3</sub>-labelled media, and sedimentation into a 1.3 mm ssNMR rotor<sup>21</sup>. The HSQC-type <sup>1</sup>H-<sup>13</sup>C correlation spectrum obtained in a total experimental time of 14 hours is shown in Figure 3 (black). As expected, the large number of non-equivalent Ala methyl sites – 337 in total – results in significant resonance overlap. Nonetheless, for the resolved cross-peaks the line widths are similar to TET2 (~25–30 Hz for <sup>13</sup>C and 40–50 Hz for <sup>1</sup>H). Signal-to-noise ratios of these cross-peaks are of the order of 3 to 8. Given the small amount of sample (1.2 nmoles) these results are encouraging, and demonstrate that methyl labelling is a route to gain residue-specific information of very large complexes. We also performed an additional experiment, using a shorter <sup>13</sup>C evolution time and larger number of scans to favour sensitivity and confirm that the detected cross peaks are indeed reproducible (red spectrum in Figure 3). Resonance assignment and the spectral complexity are of course challenging. Solutions may involve unlabelling of protein components that can be reconstituted, or chemically methyl-labelled amino acid with low abundance, such as methylated cysteines<sup>22</sup> or lysines. The data shown here – to our knowledge the first reported ssNMR spectra of uniformly methyl-labelled ribosomes – establish that in terms of sensitivity and line widths, studies of ribosomes are feasible.

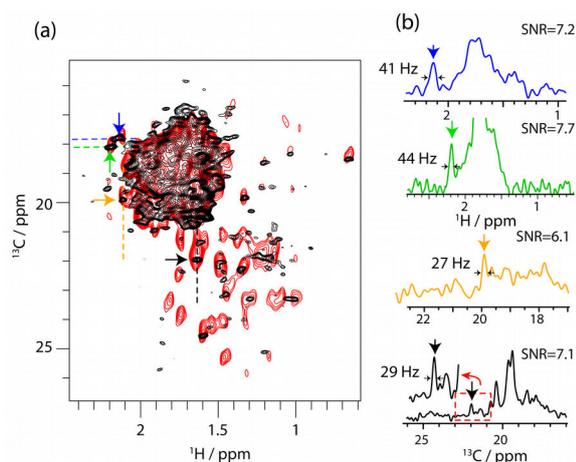


Fig. 3. Alanine methyl ssNMR spectra of *Thermus thermophilus* ribosome, obtained with 1.5 mg of deuterated, Ala-<sup>13</sup>CH<sub>3</sub> labelled material. (a) 2D spectra obtained with INEPT-transfer (black, pulse sequence Figure S1E), obtained in a total time of 14 hours. The spectrum in red was obtained in 9 hours, using CP-transfers (sequence Figure S1A) with increased number of scans at the expense of resolution (see ESI for details). (b) 1D traces of the black spectrum along the <sup>1</sup>H and <sup>13</sup>C dimensions were extracted at the positions indicated with dashed lines, and the colours of the traces correspond to the colours of the arrows.

In conclusion, we have shown here that CH<sub>3</sub>-methyl-labelling allows for high-resolution proton-detected ssNMR at high MAS frequencies. In terms of sensitivity, CH<sub>3</sub> labelling significantly outcompetes CHD<sub>2</sub> labelling. We believe that CH<sub>3</sub> methyl labelling will be very useful for further ssNMR studies of structure, dynamics and

interactions, in particular for large systems, as exemplified here.

This work was supported by the European Research Council (ERC-StG-2012-311318 and ERC-StG-2010-260887) and used the platforms of the Grenoble Instruct Center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

- 1 A. Loquet, N. G. Sgourakis, R. Gupta, K. Giller, D. Riedel, C. Goosmann, C. Griesinger, M. Kolbe, D. Baker, S. Becker and A. Lange, *Nature*, 2012, **486**, 276–9.
- 2 C. Wasmer, A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek and B. H. Meier, *Science*, 2008, **319**, 1523–6.
- 3 L. B. Andreas, M. Reese, M. T. Eddy, V. Gelev, Q. Z. Ni, E. A. Miller, L. Emsley, G. Pintacuda, J. J. Chou and R. G. Griffin, *J. Am. Chem. Soc.*, 2015, **137**, 14877–14886.
- 4 R. Sprangers, A. Velyvis and L. E. Kay, *Nat. Methods*, 2007, **4**, 697–703.
- 5 A. M. Ruschak and L. E. Kay, *J. Biomol. NMR*, 2010, **46**, 75–87.
- 6 D. Sheppard, R. Sprangers and V. Tugarinov, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **56**, 1–45.
- 7 V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw and L. E. Kay, *J. Am. Chem. Soc.*, 2003, **125**, 10420–10428.
- 8 J. E. Ollerenshaw, V. Tugarinov, N. R. Skrynnikov and L. E. Kay, *J. Biomol. NMR*, 2005, **33**, 25–41.
- 9 L. B. Andreas, T. Le Marchand, K. Jaudzems and G. Pintacuda, *J. Magn. Reson.*, 2015, **253**, 36–49.
- 10 V. Agarwal, A. Diehl, N. Skrynnikov and B. Reif, *J. Am. Chem. Soc.*, 2006, **128**, 12620–12621.
- 11 C. Amero, M. Asunción Durá, M. Noirclerc-Savoye, A. Perollier, B. Gallet, M. J. Plevin, T. Vernet, B. Franzetti and J. Boisbouvier, *J. Biomol. NMR*, 2011, **50**, 229–236.
- 12 M. Huber, O. With, P. Schanda, R. Verel, M. Ernst and B. H. Meier, *J. Magn. Reson.*, 2012, **214**, 76–80.
- 13 V. Agarwal, Y. Xue, B. Reif and N. R. Skrynnikov, *J. Am. Chem. Soc.*, 2008, **130**, 16611–16621.
- 14 J. E. Ollerenshaw, V. Tugarinov and L. E. Kay, *Magn. Reson. Chem.*, 2003, **41**, 843–852.
- 15 P. Schanda and M. Ernst, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2016, **96**, 1–46.
- 16 F. A. A. Mulder, L. Bouakaz, A. Lundell, M. Venkataramana, A. Liljas, M. Akke and S. Sanyal, *Biochemistry*, 2004, **43**, 5930–5936.
- 17 L. D. Cabrita, et al, *Nat. Struct. Mol. Biol.*, 2013, **23**, 278–285.
- 18 S.-T. D. Hsu, P. Fucini, L. D. Cabrita, H. Launay, C. M. Dobson and J. Christodoulou, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16516–16521.
- 19 E. Barbet-Massin, C. T. Huang, V. Daebel, S. T. D. Hsu and B. Reif, *Angew. Chemie - Int. Ed.*, 2015, **54**, 4367–4369.
- 20 I. Gelis, V. Vitzthum, N. Dhimole, M. A. Caporini, A. Schedlbauer, D. Carnevale, S. R. Connell, P. Fucini and G. Bodenhausen, *J. Biomol. NMR*, 2013, **56**, 85–93.
- 21 I. Bertini, C. Luchinat, G. Parigi, E. Ravera, B. Reif and P. Turano, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 10396–10399.
- 22 T. L. Religa, A. M. Ruschak, R. Rosenzweig and L. E. Kay, *J. Am. Chem. Soc.*, 2011, **133**, 9063–9068.