



**HAL**  
open science

# Etude moléculaire des étapes initiales d'importation de protéines mitochondriales

Iva Sucec

► **To cite this version:**

Iva Sucec. Etude moléculaire des étapes initiales d'importation de protéines mitochondriales. Biologie structurale [q-bio.BM]. Université Grenoble Alpes [2020-..], 2022. Français. NNT : 2022GRALV052 . tel-04175258

**HAL Id: tel-04175258**

**<https://theses.hal.science/tel-04175258>**

Submitted on 2 Aug 2023

**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 L'UNIVERSITÉ GRENOBLE ALPES**

École doctorale : CSV- Chimie et Sciences du Vivant

Spécialité : Biologie Structurale et Nanobiologie

Unité de recherche : Institut de Biologie Structurale

## **Etude moléculaire des étapes initiales d'importation de protéines mitochondriales**

### **Atomic-level studies of the initial mitochondrial-protein import steps**

Présentée par :

**Iva SUCEC**

#### Direction de thèse :

**Paul SCHANDA**

Directeur de thèse

**Beate BERSCH**  
cnrs

Co-directrice de thèse

#### Rapporteurs :

**Nadia IZADI-PRUNEYRE**

DIRECTRICE DE RECHERCHE, CNRS délégation Paris Centre

**Sebastian HILLER**

PROFESSEUR ASSOCIE, Universität Base

#### Thèse soutenue publiquement le **1 août 2022**, devant le jury composé de :

**Paul SCHANDA**

PROFESSEUR, Institute of Science and Technology, Austria

Directeur de thèse

**Beate BERSCH**

CHARGE DE RECHERCHE HDR, CNRS délégation Alpes

Co-directrice de thèse

**Johannes HERRMANN**

PROFESSEUR, Technische Universität Kaiserslautern

Examineur

**Karen FLEMING**

PROFESSEUR, Johns Hopkins University

Examinatrice

**Nadia IZADI-PRUNEYRE**

DIRECTRICE DE RECHERCHE, CNRS délégation Paris Centre

Rapporteuse

**Sebastian HILLER**

PROFESSEUR ASSOCIE, Universität Base

Rapporteur

**Franz BRUCKERT**

PROFESSEUR DES UNIVERSITES, Grenoble INP

Président







UNIVERSITÉ GRENOBLE ALPES



DOCTORAL THESIS

---

# Atomic-level studies of the initial mitochondrial-protein import steps

---

*Author:*  
Iva Sučec

*Supervisor:*  
Prof. Dr. Paul Schanda  
*Co-supervisor:*  
Dr. Beate Bersch

*A thesis submitted in fulfillment of the requirements  
for the Doctor of Philosophy*

*of the*

École doctorale - Chimie et Sciences du vivant  
Biologie Structurale et Nanobiologie

July 29, 2022



## *Acknowledgements*

To start with, I would like to thank the reviewers of this manuscript, Professors Nadia Izadi-Pruneye and Sebastian Hiller, for taking the time to critically read and review the report of the work done during my PhD training. Also, I would like to thank other members of my thesis defense jury, Professors Karen Fleming, Johannes Herrmann and Franz Bruckert for your valuable insights and productive discussion during the defense.

I would like to express my gratitude to the members of my PhD thesis committee, Prof. Eva Pebay-Peyroula, Prof. Doron Rapaport and Prof. Martin Weik, for their helpful and valuable feedback throughout my PhD training.

The IBS NMR Spectroscopy group, you all made my time in the group very pleasant and I learned a lot from all of you. I will mention only few people in particular. Isa, Lionel, Karine and Rida, thank you for being there and helping with every technical question and issue around the lab, for keeping everything (and everyone) in place and most importantly, for making our wet-lab such a fun and pleasant place to work at. Catherine and Bernhard, thank you for your time to answer my questions and sharing your wisdom. Cedric, Alicia and Adrien, thank you for your technical support, all de-bugging and for being so approachable whenever I had a question and needed help. Nina, Ricarda, Tiago and Faustine, it was such a pleasure to share the office/lab, ideas and occasional frustrations with you!

I would like to acknowledge everyone who contributed and helped with the work presented in this thesis. Katharina, even if I never got a chance to meet you, thank you for all the work you did during your PhD that laid the bases of my projects. Undina, thank you for all discussions that helped in wet-lab protocol optimizations, and all the protein samples you prepared at IBS and IST. Jakob and Lukas at ISTA, thank you for continuing this beautiful project and for the experiments performed in the last few months, with the help from Olena Tsvyetkova. Also, I would like to thank Antoine for the experiments performed during his internship in the NMR group.

Most importantly, I would like to thank Paul and Beate for their guidance, supervision and all discussions during my PhD training. Thank you Beate for constantly encouraging critical thinking and for always being available to help and discuss. Thank you Paul for all the knowledge you so happily shared and for your passion towards the science that is so easy to absorb. Additionally, thank you for recognizing when I felt overwhelmed or lost and for directing me back on track. Thank you both for keeping me motivated through the years, which admittedly was not so easy all the time, and for supervisor-complementarity which made my training and projects more fluent and always so interesting.

To the Grenoble Gang, you guys are the best! Undina, even if it is hard sometimes to exclude work from our talks, I am so thankful for sharing the time outside of lab with you and for having you as a friend. Arthur, Aggeliki, Arijit, Ricarda, Aldo, Laura and Nina thank you guys for sharing the time, thoughts, laughter and the most beautiful hikes with me! You made my time in Grenoble unforgettable and even more valuable!

Zrinka, vjerujem da su tvoj optimizam, strpljivost prema mladom Padawanu i ljubav prema znanosti iznimno utjecali na moj put u znanstveno istraživanje. Hvala ti što si bila moj prvi mentor!

Beskrajno bi se htjela zahvaliti svom ocu Igoru. Tata, hvala ti što vjerujes u mene i na tvojoj beskrajnoj podršci u svemu. Hvala cijeloj mojoj obitelji, bez vas, bez svih dobrih i loših trenutaka koji su oblikovali moje odluke, sve ovo ne bi bilo moguće. Hvala vam što ste uvijek uz mene, bez obzira na kojem dijelu svijeta sam.



UNIVERSITÉ GRENOBLE ALPES

*Abstract*

École doctorale - Chimie et Sciences du vivant

Biologie Structurale et Nanobiologie

**Atomic-level studies of the initial mitochondrial-protein import steps**

by Iva Sučec

Mitochondria play an important role in numerous eukaryotic cellular processes. 99% of their ca. 1500 different proteins are encoded by the nuclear genome and synthesized in the cytosol. Consequently, the proper functioning of mitochondria depends on correct import of all these protein as precursor proteins, and their insertion and folding in the right sub-compartment within mitochondria. The translocase of the outer membrane (TOM) complex specifically recognizes the precursor proteins via their targeting signals and enables their translocation through the outer membrane. Among the import machineries downstream from the TOM complex, a sole chaperoning system of the intermembrane space, TIM holdase chaperones, ensures the transport of highly hydrophobic membrane proteins through the intermembrane space.

In the first part I show the characterization of the membrane protein client binding on two TIM chaperones which explains the TIM's chaperones substrate specificity on an atomic level. Previously, the conserved hydrophobic binding site on TIM9.10 chaperone has been revealed integrating solution NMR experiments with different biophysical and biochemical assays and molecular modeling. The "fuzzy" mode of interaction, where the client binds in an unfolded, translocation-competent state, sampling a multitude of conformations, supports apparent contradictory role of TIM chaperones: protecting highly hydrophobic membrane clients from the aggregation by tight binding, and also the transfer of the client to the downstream insertase without any significant energetic barrier. Here we show that a small pool of membrane protein clients with additional soluble domain shows preference in binding to the other, non-essential TIM8.13 chaperone. We showed that for this type of client, the binding is not only mediated by the hydrophobic patches on the chaperone, but that it additionally involves a network of polar interactions at a distinct binding site that only TIM8.13 provides. Integrating NMR, SAXS and molecular dynamic simulations provides us with two models of chaperone-client binding for two structurally very similar chaperones.

In the second part focuses on interaction studies of the three cytosolic receptor domains of the TOM import machinery, Tom20, Tom22 and Tom70. The cytosolic domains of these proteins specifically recognize their clients, the precursor proteins, via the targeting signals of the client, but the mechanisms remain unclear. Because the receptor domains have partially overlapping function, biochemical and cellular experiments conducted in several research groups worldwide have only been partially successful in deciphering the mechanism of client recognition. To provide atomic-level mechanistic insights, I have used solution NMR spectroscopy, biochemical and biophysical techniques to study and characterize: (i) the cytosolic domains of the Tom receptors without client proteins, and (ii) in complexes with client proteins or fragments thereof, (iii) interactions between the individual Tom receptors and (iv) Tom receptor(s) interactions with the cytosolic co-chaperone Xdj1. These are challenging studies due to the dynamic, flexible - and in case of Tom70, relatively big - nature of the Tom receptor domains. Our results reveal an interaction pattern of cytosolic, intrinsically disordered domain of Tom22 with the other two receptors and the cytosolic co-chaperone, suggesting an universal role of Tom22 as a client replacement in the Tom20, Tom70 and Xdj1 client binding site. Such interaction may be required for the release of the client protein and its transfer from the Tom20 receptor (or Tom70 or Xdj1) through the outer membrane Tom40 pore.

Together, the results obtained during my PhD provide functional insight into the sequence of the events during the first steps of the mitochondrial protein import.

UNIVERSITÉ GRENOBLE ALPES

## *Résumé*

École doctorale - Chimie et Sciences du vivant

Biologie Structurale et Nanobiologie

**Étude moléculaire des étapes initiales d'importation de protéines mitochondriales**

Iva Sučec



Chez les eucaryotes, les mitochondries jouent un rôle important dans de nombreux processus cellulaires. 99% des 1500 protéines mitochondriales sont codées par le génome nucléaire et synthétisées dans le cytosol. Par conséquent, le bon fonctionnement des mitochondries dépend de l'importation, de la localisation et du repliement correct de ces protéines dites précurseurs/clientes. Le TOM complexe formant la translocase de la membrane externe (TOM) reconnaît spécifiquement les protéines précurseuses via leurs signaux d'adressage et permet leur translocation à travers la membrane externe. En aval du complexe TOM, un seul système de chaperonnes de l'espace intermembranaire, les chaperonnes TIM, assurent le transport des protéines membranaires hautement hydrophobes à travers l'espace intermembranaire.

Dans la première partie, je montre la caractérisation de la liaison de différentes protéines membranaires clientes à deux chaperonnes TIM, en cherchant à expliquer leur spécificité à un niveau atomique. Auparavant, en intégrant des expériences de RMN en solution avec différentes techniques biophysiques et biochimiques ainsi que la modélisation moléculaire, le site de liaison hydrophobe conservé sur la chaperonne TIM9.10 a été révélé. Le mode d'interaction "flou", où le client se lie dans un état déplié et translocation-compétent, échantillonnant une multitude de conformations, soutient le rôle apparemment contradictoire des chaperonnes TIM: protéger les clients membranaires hautement hydrophobes de l'agrégation par une liaison étroite, en autorisant simultanément leur transfert à l'insertase en aval sans barrière énergétique significative. Nous montrons ici qu'un petit groupe de clients de protéines membranaires possédant un domaine supplémentaire soluble, montre une préférence pour la chaperonne TIM8.13, non essentielle dans la levure. Pour ce type de client, la liaison n'est pas seulement médiée par les patches hydrophobes de la chaperonne, mais elle implique également un réseau d'interactions polaires sur un site de liaison distinct que seul TIM8.13 fournit. L'intégration des données RMN et SAXS avec des simulations de dynamique moléculaire nous fournit deux modèles de liaison chaperonne-client pour deux chaperonnes structurellement très similaires.

Dans la deuxième partie, je présente les études d'interaction des domaines cytosoliques de trois récepteurs de la machinerie d'importation TOM. Les récepteurs Tom20, Tom22 et Tom70 reconnaissent spécifiquement les protéines précurseurs via les signaux d'adressage. En raison du chevauchement partiel des fonctions de ces récepteurs démontrées *in vivo*, le mécanisme de reconnaissance des clients ainsi que l'interaction entre les trois récepteurs reste à caractériser au niveau atomique. Nous avons utilisé la spectroscopie RMN en solution, ainsi que des techniques biochimiques et biophysiques pour étudier et caractériser (i) les domaines cytosoliques des récepteurs en absence de client, (ii) les complexes récepteur-protéine cliente, (iii) les interactions entre les récepteurs et (iv) les interactions entre les récepteurs Tom et une co-chaperonne cytosolique. Ces études constituent un défi en raison de la nature dynamique, flexible et, dans le cas de Tom70, relativement grande des domaines des récepteurs. Nos résultats révèlent un modèle d'interaction du domaine cytosolique intrinsèquement désordonné de Tom22 avec les deux autres récepteurs et avec la co-chaperonne cytosolique Xdj1. Ce modèle suggère un rôle universel de Tom22 qui implique le remplacement du client dans son site de liaison sur Tom20, Tom70 et Xdj1. Une telle interaction peut être nécessaire pour la libération de la protéine cliente et son transfert du récepteur Tom20 (ou Tom70 ou Xdj1) à travers le pore Tom40 de la membrane externe.

Ensemble, les résultats obtenus fournissent des indications fonctionnelles sur les premières étapes de l'importation des protéines mitochondriales.

# Contents

Acknowledgements	iii
Abstract	vi
Résumé	viii
List of Figures	xiii
List of Tables	xv
List of Abbreviations	xvii
<b>1 Introduction 1.: Structural and functional insights into first steps of mitochondrial proteins import</b>	<b>1</b>
1.1 Mitochondrial import machineries	2
1.1.1 Translocase of the outer membrane: overall composition and structure	3
1.1.2 Tom20: interaction with the matrix targeting signal sequence	4
1.1.3 Tom22: the central hub of the TOM complex	6
Interaction of Tom22 and precursor proteins.	6
Interaction of Tom22 and Tom20.	7
1.1.4 Tom70: receptor for precursors with internal-targeting signals	8
Structural studies of the cytosolic domain of Tom70.	8
Client binding sites on Tom70.	9
1.1.5 Chaperoning in the inter-membrane space	10
1.2 How do chaperones bind (partly) unfolded client proteins?	12
<b>2 Introduction 2.: Preparing chaperone-client protein complexes for biophysical and structural studies</b>	<b>13</b>
2.1 Client-chaperone complex formation	14
2.1.1 General considerations	14
2.1.2 Overview of the possibilities to prepare chaperone-client complexes	19
2.2 Different complex-formation approaches in practice	19
2.2.1 Forming complexes in solution by mixing chaperones and their soluble client proteins	19
DnaK complexes	19
Complexes of the chaperone Spy with Im7	20
Hsp90 and Hsp110 complexes	21
2.2.2 Making a client protein bind by adjusting sample conditions	22
Hsp60 (thermosome) complexes formed at high temperature	22
Trigger factor complexes obtained via temperature or disulfide-bond breakage	23

2.2.3	Mutating client proteins to make them chaperone-binding prone (or reduce their aggregation propensity) . . . . .	23
2.2.4	Complex formation upon removal of a denaturant . . . . .	24
	Complex formation by dialysis . . . . .	26
2.2.5	Hampering aggregation by client immobilization: complex formation with a pull-down approach . . . . .	26
2.2.6	Purifying P-C complexes from the cell . . . . .	28
2.2.7	Capturing emerging client proteins in a cell-free system . . . . .	28
2.2.8	Fusing chaperone and client into a single polypeptide chain . . . . .	29
2.2.9	Chaperones bound to protein aggregates . . . . .	29
2.3	Concluding remarks . . . . .	30
<b>3</b>	<b>Open questions addressed in this thesis: molecular mechanisms of chaperones and receptor domains</b>	<b>31</b>
<b>4</b>	<b>Results 1: receptor interactions and precursor protein handling at the mitochondrial import gate</b>	<b>33</b>
4.1	Conformation and dynamics of Tom22 . . . . .	34
4.1.1	The cytosolic Tom22 domain is an intrinsically disordered protein with a transient helix	34
4.2	Interaction of the presequence and the Tom receptors . . . . .	39
4.2.1	Tom22 <sub>cyt</sub> does not bind the precursor presequences . . . . .	39
4.2.2	Presequence binding site on the Tom20 receptor . . . . .	41
4.2.3	Presequence-containing client protein and its interactions with the Tom receptors . .	44
4.3	Interaction of Tom20 and Tom22 cytosolic domains . . . . .	50
4.3.1	Tom22 binds the presequence-binding site of the Tom20 receptor . . . . .	50
4.3.2	Tom22 employs the residues of the cytosolic $\alpha$ -helical region for binding to Tom20 . .	53
4.3.3	Reduced binding of Tom22 with Tom20 variants . . . . .	58
4.4	Testing the possibility of a ternary Tom20-Tom22-presequence complex formation . . . . .	60
4.5	Interaction of Tom22 and Tom70 cytosolic receptor domains . . . . .	64
4.5.1	Tom22 <sub>cyt</sub> employs its small helix for binding to the Tom70 receptor . . . . .	64
4.5.2	Cytosolic helix of Tom22 binds to the putative client binding site of Tom70 receptor .	65
4.6	Determining kinetic parameters of the interactions . . . . .	70
4.7	Conclusions on the molecular mechanism of the Tom receptor domains . . . . .	71
<b>5</b>	<b>Results 2. Molecular basis of the client specificity of the TIM holdases.</b>	<b>77</b>
5.1	Introduction to the Client specificity of mitochondrial TIM chaperones . . . . .	78
5.2	TIM8·13 and TIM9·10 interact differently with membrane precursor proteins . . . . .	80
5.3	The small TIM chaperones use a conserved hydrophobic cleft for membrane precursor protein binding . . . . .	82
5.4	Hydrophilic fragments interact differently with TIM8·13 and TIM9·10 . . . . .	84
5.5	TIM8·13 uses an additional hydrophilic face for protein binding . . . . .	86
5.6	Structural ensembles of chaperone-Tim23 complexes . . . . .	88
5.7	Discussion of the results obtained for the client specificity of mitochondrial TIM chaperones	91
<b>6</b>	<b>Results 3. New client proteins of TIM9·10 dependant carrier import pathway</b>	<b>93</b>
6.1	New protein clients of TIM9·10 dependant carrier import pathway . . . . .	93
<b>7</b>	<b>Material and methods</b>	<b>95</b>
7.1	Production and purification of the TIM chaperones . . . . .	95

7.2	Production and purification of TIM client proteins . . . . .	96
7.3	Preparation of chaperone-precursor protein complexes . . . . .	97
7.3.1	Competition assays . . . . .	97
7.4	Purification of the Tom receptors . . . . .	97
7.4.1	Production and purification of the Tom70 receptor . . . . .	97
7.4.2	Production and purification of the Tom22 receptor . . . . .	98
7.4.3	Production and purification of the Tom20 receptor . . . . .	98
7.4.4	Purification of presequence-containing client protein . . . . .	99
7.5	Peptide sequences used in the chaperones and receptors interaction studies . . . . .	100
7.6	Cell-free protein production . . . . .	100
7.7	Isothermal Titration Calorimetry experiments . . . . .	101
7.7.1	ITC experiments with the TIM chaperones . . . . .	101
7.7.2	ITC experiments with the Tom receptors . . . . .	101
7.8	SEC-MALS experiments . . . . .	101
7.9	Analytical ultra-centrifugation . . . . .	102
7.10	NMR spectroscopy . . . . .	104
7.10.1	TIM8-13, Tim23 <sub>IMS</sub> , Tom22, Tom20, Tom70 and GB1-pSu9(1-69) resonance assignments	104
7.10.2	NMR titration experiments . . . . .	104
	VDAC titration experiments . . . . .	104
	Tim23 <sub>IMS</sub> titration experiments . . . . .	105
	Titration experiments with Tom20, Tom22 and the presequences . . . . .	105
7.10.3	NMR experiments with the Tim23 <sub>FL</sub> . . . . .	105
7.10.4	Diffusion ordered spectroscopy . . . . .	105
7.11	Small-angle X-ray scattering data collection and analysis . . . . .	105
7.12	Calculations of affinities and populations . . . . .	106
7.12.1	Estimation of the population of Tim23 <sub>IMS</sub> -bound states from ITC-derived $K_d$ . . . . .	106
7.13	NMR relaxation experiments . . . . .	107
7.14	Paramagnetic relaxation enhancement experiments (PRE) . . . . .	107
7.15	Calculating affinities from NMR titration experiments . . . . .	108
<b>A</b>	<b>Appendix A</b> . . . . .	<b>111</b>
A.1	Supporting information for Structural basis of client specificity in mitochondrial membrane-protein chaperones . . . . .	111
<b>B</b>	<b>Appendix B</b> . . . . .	<b>13</b>
B.0.1	Characterization of the yeast Tom70 receptor . . . . .	21
	Assignment of the Ile <sup><math>\delta^1</math></sup> Met <sup><math>\epsilon</math></sup> Thr <sup><math>\gamma</math></sup> methyl groups of the Tom70(39-617). . . . .	23
B.0.2	Kinetic parameters obtained with ITC experiments . . . . .	25
B.0.3	Experimentally determined diffusion coefficients using NMR diffusion-ordered spectroscopy (DOSY). . . . .	26
B.1	Published work . . . . .	26
	<b>References</b> . . . . .	<b>95</b>



# List of Figures

1.1	Mitochondrial protein import pathways. . . . .	3
1.2	Structure of the Translocase of the Outer Membrane (TOM) core complex. . . . .	4
1.3	Cytosolic domain of rat Tom20 bound to the presequence peptide. . . . .	5
1.4	Structural models of the cytosolic domain of Tom70 receptor. . . . .	10
2.1	Schematic overview of the reactions that may occur in a system composed of a chaperone (C) and a client protein (P). . . . .	17
2.2	Schematic overview of the relevant reactions that can lead to the formation of a complex of a chaperone (C) with a soluble client protein (P) that has negligible aggregation propensity. . . . .	20
2.3	Scheme of different approaches for the mitochondrial intermembrane space chaperone complex formation. . . . .	27
2.4	Overview of the main methods used for preparing chaperone-client complexes. . . . .	30
4.1	Intrinsically disordered domain of Tom22. . . . .	35
4.2	Multiple sequence alignment and tertiary structure prediction of Tom22 receptor. . . . .	37
4.3	Structure prediction of Tom22 receptor. . . . .	38
4.4	Tom22 does not interact with the presequences of the mitochondrial precursor proteins . . . . .	41
4.5	Multiple sequence alignment of Tom20 protein from different organisms. . . . .	42
4.6	Yeast Tom20 receptor binds the presequence in the predicted binding site. . . . .	43
4.7	Confirmation of the presequence binding site on the rat Tom20 receptor. . . . .	44
4.8	Output of the automated resonance assignment with FLYA and secondary structure prediction for GB1-pSu9(1-69) construct. . . . .	46
4.9	Heteronuclear NOE relaxation experiments on GB1-pSu9(1-69) construct alone and upon adding ScTom20. . . . .	47
4.10	Interaction of GB1-pSu9(1-69) with Tom20(37-183). . . . .	47
4.11	Interaction of GB1-pSu9(1-69) with Tom20(37-183). . . . .	48
4.12	Yeast Tom20 receptor binds the Tom22 in the presequence binding site. . . . .	51
4.13	Tom20 receptor from rat binds Tom22 in the presequence binding site. . . . .	52
4.14	Cytosolic domain of Tom22 interacts with the cytosolic domain of yeast Tom20. . . . .	54
4.15	Tom22 employs the residues of the $\alpha$ -helical region for binding to Tom20. . . . .	55
4.16	Tom22 secondary chemical shifts in the presence or absence of Tom20. . . . .	57
4.17	Cytosolic domain of Tom22 interacts with the cytosolic domain of yeast Tom20. . . . .	59
4.18	Tom22 observed interaction with Tom20 and pSu9. . . . .	62
4.19	RnTom20 observed interaction with Tom22 and pSu9. . . . .	63
4.20	Binding of the cytosolic domain of Tom22 on the Tom70 receptor. . . . .	65
4.21	Changes in the $^{13}\text{C}^1\text{H}$ HMQC spectrum of Ile $^{\delta 1}$ Met $^{\epsilon}$ Thr $^{\gamma}$ methyl groups of the Tom70(39-617) upon adding cytosolic domain of Tom22. . . . .	66
4.22	Binding site of the cytosolic domain of Tom22 on the Tom70 receptor. . . . .	68

4.23	Scheme of proposed recognition and interaction mechanism of the presequence by Tom20 and Tom22. . . . .	73
4.24	Model of structural feasibility of the proposed mechanism. . . . .	74
5.1	Biochemical characterization of TIM chaperone - membrane protein complexes. . . . .	81
5.2	Solution-NMR and binding of a VDAC fragment to TIM8-13. . . . .	83
5.3	Tim23 has markedly different properties when binding to TIM8-13 and to TIM9-10. . . . .	85
5.4	Tim23 <sub>IMS</sub> and full-length Tim23 differ in their interactions with TIM9-10 and TIM8-13 chaperones. . . . .	87
5.5	Architecture of the TIM8-13 and TIM9-10 holdases in complex with full-length Tim23. . . . .	89
5.6	[Tentative identification of electrostatic interactions from the MD ensemble. . . . .	91
6.1	Interaction of Mpc2 and Mpc3 with the TIM9-10 chaperone <i>in vitro</i> . . . . .	94
7.1	Analysis of <sup>15</sup> N-Tom22(1-74)-Xdj1(CTD1) NMR titration experiments by TITAN software. . . . .	109
S1	Kyte-Doolittle hydrophobicity of different membrane precursor proteins. . . . .	2
S2	The TIM9-10 chaperones binds Ggc1 in a highly dynamic complex. . . . .	3
S3	TIM8-13 and TIM9-10 do not form mixed complexes. . . . .	3
S4	Quantification of the amount of chaperone complexes from mass spectrometry. . . . .	4
S5	Sequence alignment of the small Tims across eukaryotes reveals a conserved hydrophobic motif. . . . .	5
S6	The Tim8 double mutant with increased hydrophobicity does not improve Tim23 binding. . . . .	6
S7	TIM chaperones have a compact rigid core and flexible tentacles. . . . .	7
S8	The linear VDAC <sub>257-279</sub> peptide does not bind to a well-defined binding site of TIM8-13. . . . .	7
S9	Isothermal titration calorimetry (ITC) experiments with TIM8-13 and TIM9-10 chaperones. . . . .	8
S10	Full-length Tim23 interactions with TIM9-10 and TIM8-13 chaperones. . . . .	9
S11	Interaction of trans-membrane Tim23 with the TIM9-10 chaperone. . . . .	10
S12	In the simultaneous presence of Tim23, TIM8-13 and TIM9-10, no ternary complex is observed. . . . .	10
S13	Experimental characterization of the size of TIM chaperones and their precursor protein complexes. . . . .	11
S1	Oligomerization state of the yeast Tom20 constructs. . . . .	13
S2	<sup>15</sup> N <sup>1</sup> H BEST-TROSY spectra of <sup>2</sup> H <sup>15</sup> N-labeled ScTom20(37-183). . . . .	14
S3	<sup>15</sup> N <sup>1</sup> H BEST-TROSY spectra of <sup>2</sup> H <sup>15</sup> N-labeled ScTom20(37-183) alone and upon adding Tom22(1-97). . . . .	15
S4	Assigned <sup>15</sup> N <sup>1</sup> H BEST-TROSY spectrum of <sup>15</sup> N-labeled GB1-pSu9(1-69). . . . .	16
S5	<sup>15</sup> N <sup>1</sup> H BEST-TROSY spectra of <sup>15</sup> N-labeled GB1-pSu9(1-69) in the absence and presence of ScTom20. . . . .	17
S6	Secondary structure propensities of Tom22 in presence and absence of Tom20. . . . .	18
S7	<sup>15</sup> N <sup>1</sup> H BEST-TROSY spectra of <sup>15</sup> N-labeled Tom22(1-74). . . . .	19
S8	Secondary structure propensities of Tom22 in presence and absence of Tom70. . . . .	20
S9	Purification and characterization of Tom70(39-617) protein construct. . . . .	22
S10	NH backbone spectrum and methyl spectrum of Tom70. . . . .	23
S11	Example of the isoleucine assignment process of Tom70. . . . .	24
S12	Chemical shift assignment of the Met, Ile and Thr methyl groups of Tom70(39-617). . . . .	25

# List of Tables

4.1	Dissociation constants for Tom receptors, protein-protein or protein-peptide interactions.	70
7.1	Protein constructs used during the studies for this thesis. . . . .	103
S1	Thermodynamic parameters of TIM8-13 - Tim23 <sub>IMS</sub> interaction from isothermal titration calorimetry measurements. . . . .	1
S1	Thermodynamic parameters of Tom receptor interactions from isothermal titration calorimetry (ITC) measurements. . . . .	25
S2	Experimentally determined diffusion coefficients using NMR diffusion-ordered spectroscopy (DOSY). . . . .	26





# List of Abbreviations

<b>AAC</b>	adenosine diphosphate (ADP)/ adenosine triphosphate (ATP) carrier
<b>ALDH</b>	aldehyde dehydrogenase
<b>AUC</b>	Analytical UltraCentrifugation
<b>BCA assay</b>	bicinchoninic acid assay
<b>BEST-TROSY</b>	<b>B</b> and-selective <b>E</b> xcitation <b>S</b> hort <b>T</b> ransient- <b>T</b> raverse <b>R</b> elaxation <b>O</b> ptimised <b>S</b> pectroscopy: a type of 2D NMR experiment with high sensitivity
<b>BLI</b>	bio-layer interferometry
<b>BPA</b>	p- <b>B</b> enzoyl- <b>L</b> - <b>P</b> henyl <b>A</b> lanine; cross-linking agent
<b>CheSPI</b>	<b>C</b> hemical shift <b>S</b> econdary structure <b>P</b> opulation <b>I</b> nference
<b>CSP</b>	<b>C</b> hemical <b>S</b> hift <b>P</b> erturbation
<b>DHFR</b>	dihydrofolate reductase
<b>EM</b>	<b>E</b> lectron <b>M</b> icroscopy
<b>FLYA</b>	fully automated structure determination of proteins in solution
<b>GB1</b>	the immunoglobulin-binding B1 domain of <i>Streptococcal</i> protein G
<b>GGC</b>	guanosine diphosphate (GDP)/ guanosine triphosphate (GTP) carrier
<b>HADDOCK</b>	<b>H</b> igh <b>A</b> mbiguity <b>D</b> riven protein-protein <b>D</b> OCKing
<b>hetNOE</b>	<b>h</b> eteronuclear <b>N</b> uclear <b>O</b> verhauser <b>E</b> ffect
<b>IDP</b>	<b>I</b> ntrinsically <b>D</b> isordered <b>P</b> rotein
<b>IM</b>	<b>I</b> nnner <b>M</b> embrane
<b>IMS</b>	<b>I</b> nter <b>M</b> embrane <b>S</b> pace
<b>iMTS-L</b>	<b>i</b> nternal- <b>M</b> atrix- <b>T</b> argeting- <b>S</b> ignal like
<b>ITC</b>	<b>I</b> sothermal <b>T</b> itration <b>C</b> alorimetry
<b>MPC</b>	<b>M</b> itochondrial <b>P</b> yruvate <b>C</b> arrier
<b>Ni-NTA</b>	nickel-nitrilotriacetic acid; metal-affinity chromatography matrix
<b>NMR</b>	<b>N</b> uclear <b>M</b> agnetic <b>R</b> esonance
<b>OM</b>	<b>O</b> uter <b>M</b> embrane
<b>OMP</b>	<b>O</b> uter <b>M</b> embrane <b>P</b> rotein
<b>POTENCI</b>	<b>P</b> rediction <b>O</b> f <b>T</b> emperature, <b>N</b> eighbor and <b>pH</b> <b>C</b> orrected shifts for <b>I</b> ntrinsically disordered proteins
<b>PiC</b>	phosphate carrier
<b>PRE</b>	<b>P</b> aramagnetic <b>R</b> elaxation <b>E</b> nhancement
<b>Rn</b>	<b>Rattus norvegicus</b> (in this work sometimes referred to simply as "rat")
<b>SAM</b>	<b>S</b> orting and <b>A</b> ssembly <b>M</b> achinery
<b>SAXS</b>	<b>S</b> mall-angle <b>X</b> -ray <b>S</b> cattering
<b>Sc</b>	<b>Saccharomyces cerevisiae</b> (in this work sometimes referred to simply as "yeast")
<b>SPR</b>	<b>S</b> urface <b>P</b> lasmon <b>R</b> esonance
<b>TALOS</b>	<b>T</b> orsion <b>A</b> ngle <b>L</b> ikelihood <b>O</b> btained from <b>S</b> hifts
<b>TEV</b>	<b>T</b> obacco <b>E</b> tch <b>V</b> irus protease

<b>TIM</b>	Translocase of the <b>I</b> nn <b>e</b> r <b>M</b> embrane
<b>TITAN</b>	software package for the analysis of NMR titrations in two dimensions
<b>TM</b>	Trans <b>M</b> embrane
<b>TOM</b>	Translocase of the <b>O</b> uter <b>M</b> embrane
<b>TPR</b>	<b>T</b> etratrico <b>P</b> eptide <b>R</b> epeats; structural motif
<b>VDAC</b>	<b>V</b> oltage- <b>D</b> eendent <b>A</b> nion <b>C</b> hannel; outer membrane $\beta$ -barrel
<b>WT</b>	wild-type

## Chapter 1

# Introduction 1.: Structural and functional insights into first steps of mitochondrial proteins import

The introduction of this manuscript is divided in two parts:

**1<sup>st</sup> Part. 1** Structural and functional insights into first steps of mitochondrial proteins import.

Literature overview of the protein receptors of the mitochondrial Translocase of the Outer Membrane (TOM) complex and the small TIM chaperones of the mitochondrial intermembrane space.

**2<sup>nd</sup> Part. 2** Preparing chaperone—client protein complexes for biophysical and structural studies.

The second part of the Introduction chapter gives a literature overview of the experimental approaches used for preparation of chaperone–client complexes for structural biology studies. We discuss the challenges and summarize general considerations for successful chaperone–client complex formation. The experimental approaches of preparing chaperone–client complexes reviewed here are about to be published as a chapter in *Biophysics of Molecular Chaperones: Function, Mechanisms and Client Protein Interactions*: I. Sućec, P. Schanda, **Preparing chaperone—client protein complexes for biophysical and structural studies**. (2022).

## 1.1 Mitochondrial import machineries

Mitochondria are double membrane organelles with 4 distinctive sub-compartments – the outer membrane (OM), the folded inner membrane (IM), the intermembrane space (IMS) and the matrix. The central role of mitochondria is the oxidative phosphorylation to produce cellular ATP, but also they are the key sites of lipid metabolism, iron homeostasis and regulation of apoptosis (Lill and Mühlenhoff, 2008; Lemarie and Grimm, 2011; Newmeyer and Ferguson-Miller, 2003). Proteomic studies suggested that there are around 1000 to 1500 proteins in budding yeast *Saccharomyces cerevisiae* and mammalian mitochondria, respectively (Morgenstern et al., 2017). The respiratory chain complex components are located in the IM, and are among a handful of proteins synthesised by the mitochondrial genomic DNA and translated by matrix ribosomes. The majority (99%) of mitochondrial proteins are encoded by the nuclear genome and are post-translationally targeted to the mitochondria, imported and subsequently sorted to their destined sub-compartments. These proteins *en route* to their final destination, are called *precursor proteins*; in this manuscript, I also employ the term *preprotein* for convenience. Some precursor proteins have a cleavable "address tag" (the presequence), but many proteins are recognised by internal sequences only. I use the term precursor protein here irrespectively of whether they do comprise a presequence or not.

Depending on the final destination of the precursor protein, different machineries of the mitochondrial import system are employed (Figure 1.1; reviewed e.g. in Wiedemann and Pfanner, 2017; Dukanovic and Rapaport, 2011; Hansen and Herrmann, 2019; Neupert, 2015; Pfanner, Warscheid, and Wiedemann, 2019; Grevel, Pfanner, and Becker, 2019). Therefore, the mitochondrial import and sorting machineries have an important role in maintaining proper mitochondrial functions, which explains the scientific interest and the numerous research studies performed in the last decades. Additionally, import of mitochondrial membrane precursor proteins requires the assistance of both cytosolic and mitochondrial molecular chaperones, which maintain these highly hydrophobic proteins in the (at least partially) unfolded and ready-to-import state (C.Y. Fan and C. Young, 2011; Bykov et al., 2020; Jores et al., 2018).

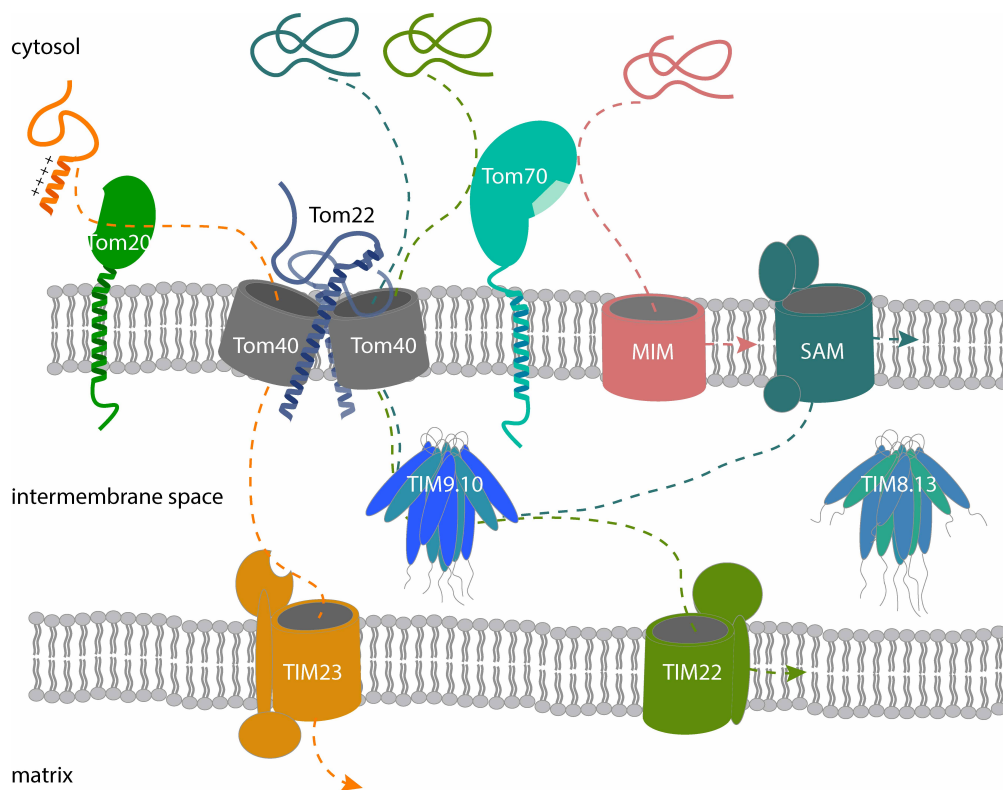


FIGURE 1.1: **Mitochondrial protein import pathways.** Mitochondrial precursors are imported through the outer membrane by the Translocase of the Outer Mitochondrial Membrane (TOM) complex, to which they are targeted by different cytosolic receptors Tom20, Tom22 and Tom70. Presequence-containing precursors, after crossing the outer membrane (OM) through the TOM complex are imported to the matrix by the presequence translocase of the inner membrane (TIM23) (orange pathway). The precursors of the mitochondrial membrane proteins are transferred from the TOM complex, across the inter-membrane space to the insertases located in the inner and outer membranes by the intermembrane space chaperones, TIM9-10 and TIM8-13. The  $\beta$ -barrel proteins are transferred to the Sorting and Assembly Machinery (SAM) for membrane insertion into the outer membrane (OM) (dark green pathway) and  $\alpha$ -helical metabolite carrier proteins are sorted to the carrier translocase (TIM22) of the inner membrane (IM) (green pathway). The precursors of polytopic (multi-spanning) OM  $\alpha$ -helical proteins can be imported via the receptor Tom70 and the Mitochondrial Import Machinery (MIM) (pathway in pink).

### 1.1.1 Translocase of the outer membrane: overall composition and structure

The Translocase of the Outer Mitochondrial Membrane (TOM) complex is the import channel for most of the nuclear-encoded mitochondrial proteins. The core complex of the import pore, TOMcc, contains the pore forming protein Tom40, the essential Tom22 protein, and three small proteins Tom5, Tom6 and Tom7 (reviewed as early as Lill and Neupert, 1996). The other two proteins of the TOM complex, Tom20 and Tom70, are client specific receptors that are not tightly bound to the core complex. The existence of the multiple translocation pores (Tom40 proteins) in the TOM core complex of *Neurospora crassa* has been shown relatively early by the electron-microscopy (EM) analysis (Ahting et al., 1999) and the presence of two active pores in *Saccharomyces cerevisiae* was functionally confirmed through cation-selective activity measurements (Meisinger et al., 2001). Additionally, Meisinger *et al.* showed that the Tom20 and Tom70 receptors are absent in the detergent extracted TOM complex indicating their loose attachment to the TOMcc (Meisinger et al., 2001). The stoichiometry and the exact composition of the TOMcc was only recently confirmed at the near-atomic level by the cryo-EM structures of the TOMcc from *S. cerevisiae* (Tucker, 2019; Araiso et al., 2019). Araiso *et al.* obtained both the trimer and the dimer form of the TOM complex, however due to higher stability of the dimer form, only the dimer structure was solved showing two molecules each of Tom40, Tom22, Tom5, Tom6 and Tom7, with a two-fold symmetry and dimensions of  $66 \text{ \AA} \times 48 \text{ \AA} \times 65$

Å for each protomer (Figure 1.2; Araiso et al., 2019). The trimer form was only marginally stable, which supports the previous EM-observed TOM complex from a mutant yeast lacking Tom20 which contained only two Tom40 channel structures (Model et al., 2001). The trimer form of the TOMcc was captured by cryo-EM at low resolution (18) (Model, Meisinger, and Kühlbrandt, 2008) and a model of the trimeric TOM complex has been constructed based on observed cross-linking patterns between the Tom subunits (Shiota et al., 2011; Shiota et al., 2015). More recently, high-speed atomic force microscopy studies suggested a dynamic exchange between the trimer and a dimer form of the TOM core complex where the Tom20 receptor could play a role in stabilizing the trimer form (reviewed in Araiso, Imai, and Endo, 2020). The role of transmembrane domain of the Tom22 receptor in stabilizing two Tom40 pores in the active dimeric form of the TOMcc could be deduced from the cryo-EM structures of the TOM core complex (Model, Meisinger, and Kühlbrandt, 2008; Tucker, 2019; Araiso et al., 2019) (Figure 1.2).

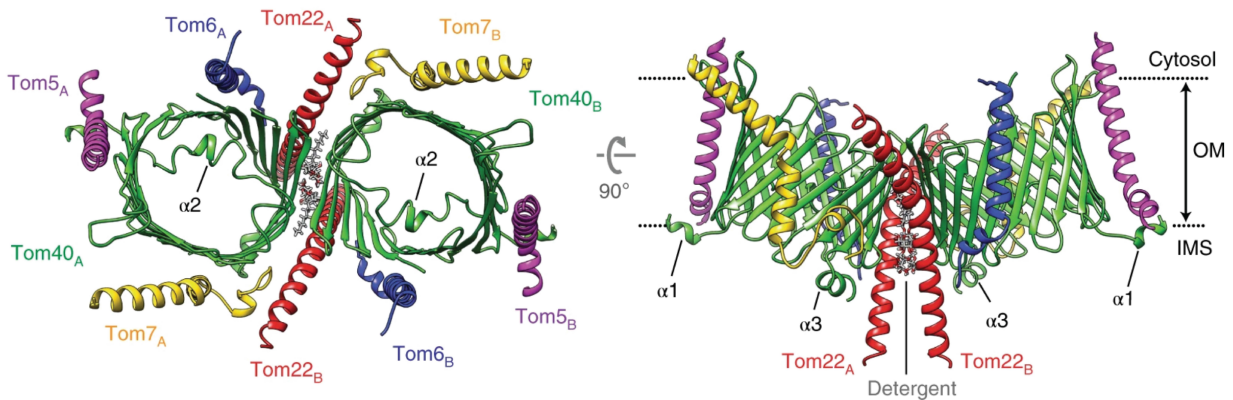


FIGURE 1.2: **Structure of the Translocase of the Outer Membrane (TOM) core complex (Tucker, 2019).** The structure of the dimer form of the TOM core complex from *S. cerevisiae*, isolated by purifying affinity-tagged Tom22 and Tom40 proteins. The absence of the cytosolic domain of Tom22 (residues 1-88) in the electron density map is suggested to be due to its high flexibility. As observed by the authors, Tom22 contains an unusually long (~45 amino acid)  $\alpha$ -helix, of which the middle portion (roughly positions 100–118) spans the membrane (Tucker, 2019). Transmembrane domain of Tom22 is filling a gap towards the intermembrane space between two Tom40 pores which is thought to stabilize the dimer. Each TOM protein subunit is colored differently. The figure was extracted from Tucker, 2019 (panels E and F of Figure 1).

While the interaction of the precursor protein with the Tom40 channel has been shown by cross-linking experiments (Esaki et al., 2003; Esaki et al., 2004) and confirmed by the atomic-resolution structural studies of the TOM complex (Bausewein et al., 2017; Tucker, 2019; Araiso et al., 2019), the interaction patterns and the mechanism of client recognition by the Tom receptors is still unclear. The structural studies of this initial import steps are hampered due to absence of the Tom22 cytosolic domain, presumably due to its high flexibility, and of the Tom20 and Tom70 receptors in the cryo-EM structures of the TOMcc complexes (Tucker, 2019; Araiso et al., 2019). Functional studies, such as *in vivo* or *in organello* studies, are hampered due to apparent overlapping function of the receptors.

### 1.1.2 Tom20: interaction with the matrix targeting signal sequence

The Tom20 protein is anchored in the outer mitochondrial membrane by the N-terminal hydrophobic transmembrane helix while its soluble C-terminal domain is exposed to the cytosol. As shown earlier, the Tom20 and Tom70 subunits are generally lost when isolating the TOM complex (Meisinger et al., 2001); accordingly, the cryo-EM structures of the yeast TOM core complex (Tucker, 2019; Araiso et al., 2019) lack these two proteins. Using single particle tracking microscopy, the mobility of fluorescently-tagged human Tom20 in the mitochondrial membrane has been studied (Bhagawati et al., 2021). Interestingly, it was observed that

Tom20 has decreased mobility in the presence of a substrate protein; this finding was ascribed to association with the TOM core complex (Bhagawati et al., 2021). The structural basis for the association of Tom20 to the TOM complex is not known.

Tom20 is considered a primary receptor for mitochondrial precursor proteins with an N-terminal matrix targeting signal, also called presequence-containing precursor proteins (Pfanner, 2000; Rapaport, 2003; Endo, Yamamoto, and Esaki, 2003). The mitochondrial presequences are unstructured N-terminal extensions of the precursor proteins, ~20-37 amino acids long, that in the presence of the negatively charged environment have the potential to form amphiphilic  $\alpha$ -helices, helices with one hydrophobic and one positively charged side (Schneider et al., 1998; Maduke and Roise, 1996; Pfanner, 2000). Biogenesis of the presequence-containing precursor proteins is extensively studied in yeast (reviewed in Pfanner, 2000; Mossmann, Meisinger, and Vögtle, 2012). It is assumed that after the initial recognition of the presequence by the Tom20 receptor, the presequence proteins are handed to the central TOM receptor Tom22 (Kiebler et al., 1993; Brix, Dietmeier, and Pfanner, 1997) before being translocated across the outer membrane by the  $\beta$ -barrel outer membrane Tom40 pore.

Previous work on the cytosolic domain of rat Tom20 showed that the presequence of mitochondrial aldehyde dehydrogenase (pALDH) binds Tom20 at a groove formed by the hydrophobic patch containing residues Phe70, Leu71, Ile74, Leu106, Val109 and Thr113; in the NMR-derived pALDH-Tom20 complex structure the presequence adopts an amphiphilic  $\alpha$ -helical conformation (Figure 1.3; Abe et al., 2000). From the size of the hydrophobic patch it has been proposed that the rat Tom20 can accommodate an amphiphilic helix that is at most two turns long, and that the sequences recognized by Tom20 are as short as 8 amino acid residues with a weak consensus motif  $\varphi \varphi X X \varphi \varphi$  (where  $\varphi$  is a hydrophobic, and X is any amino acid) (Abe et al., 2000; Saitoh et al., 2007; Muto et al., 2001).

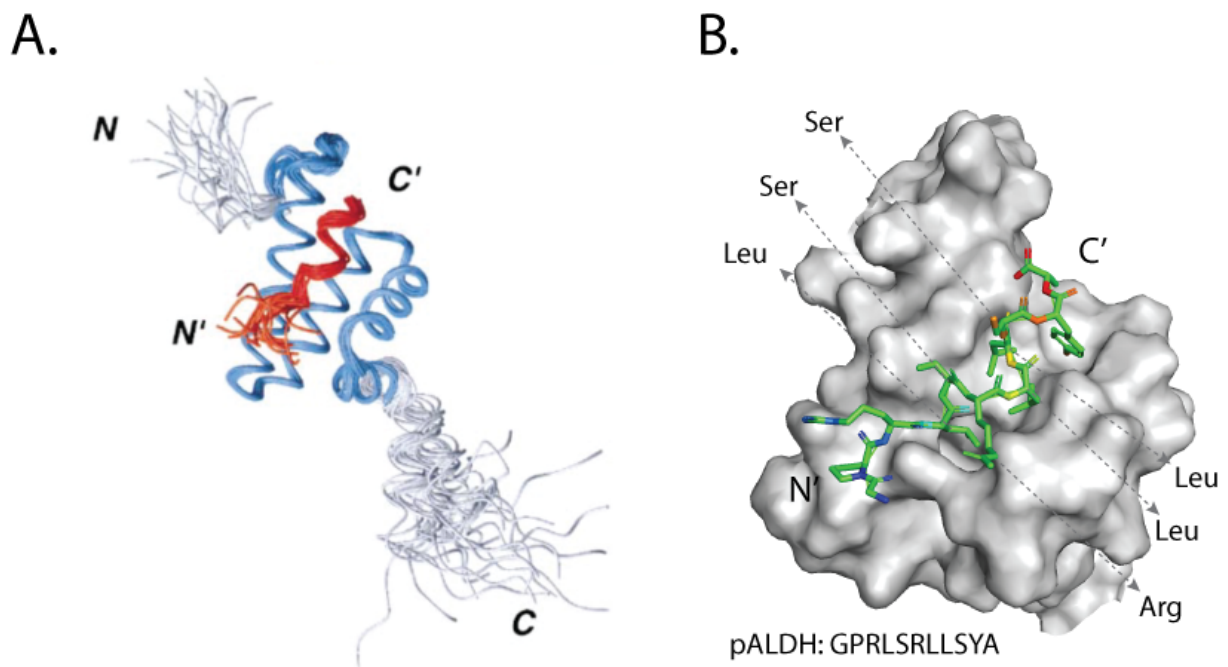


FIGURE 1.3: Cytosolic domain of rat Tom20 bound to the presequence peptide. **A.** Structures of the cytosolic domain of rat Tom20(51-145) (blue/gray) bound to the pALDH presequence peptide (red). Structures obtained by NMR studies (Abe et al., 2000). This panel is extracted from Figure 3 (panel B) of (Abe et al., 2000).

**B.** Molecular surface of the presequence binding site of rat Tom20 (PDB ID: 1om2; (Abe et al., 2000)), with indicated residues of the pALDH presequence forming a weak motif  $\varphi \varphi X X \varphi \varphi$ .



NMR relaxation studies revealed motion on the sub-millisecond timescale in the hydrophobic binding site of Tom20 in the presence of the presequence peptide and the peptide mobility inside the binding groove was proposed (Saitoh *et al.*, 2007). To obtain a complex stable enough for crystallography and NMR studies Saitoh *et al.* used a disulphide-bonded fusion-construct of the two.

Other precursor proteins that are recognized by the Tom20 receptor are the  $\beta$ -barrel OM precursor proteins, which are not synthesized with the cleavable presequence, but instead have targeting information within the mature protein sequence. It was shown that for the efficient import of the mitochondrial  $\beta$ -barrel OM precursor proteins, some kind of secondary structure, more specifically  $\beta$ -hairpin, serves as a targeting signal (Doron Rapaport and Walter Neupert, 1999; Jores *et al.*, 2016). In the absence of Tom20,  $\beta$ -barrel OM precursor proteins can be recognized by the Tom70 receptor, although not with the same efficiency (Yamano *et al.*, 2008b).

It was shown that Tom20 proteins from human and rat mitochondria exhibit overall sequence identities of  $\sim 30\%$  with yeast Tom20, and can complement the functional defects of Tom20-deficient yeast cells (Swie Goping, Millar, and Shore, 1995; Iwahashi *et al.*, 1997).

### 1.1.3 Tom22: the central hub of the TOM complex

The Tom22 protein is the only receptor of the TOM complex whose deletion has a strong impact on precursor-protein import and yeast cell viability (Van Wilpe *et al.*, 1999; Lithgow *et al.*, 1994). Tom22 has an opposite orientation from the other two receptors with an N-terminal soluble cytosolic domain (not visible in the cryo-EM structures presumably due to large-amplitude flexibility (Tucker, 2019; Araiso *et al.*, 2019)), a transmembrane helix anchoring the receptor to the OM (and maintaining the TOM complex integrity, Kiebler *et al.*, 1993; Van Wilpe *et al.*, 1999; Shiota *et al.*, 2011), and a C-terminal domain protruding to the intermembrane space. The latter is involved in the interaction with the Tim50 protein of the TIM23 inner-membrane translocase (Shiota *et al.*, 2011). Cytosolic domain of Tom22 receptor contains several conserved negatively charged residues (protein sequence alignment in Figure 4.2) and is proposed to have a dual role.

#### Interaction of Tom22 and precursor proteins.

Tom22 is generally referred to as a receptor protein of the TOM complex and its interaction with presequences was addressed by numerous independent studies. Kiebler *et al.* investigated the import of radiolabeled precursor proteins that were known to require either Tom20 or Tom70 (precursors with and without cleavable presequence of proteins of all mitochondrial compartments). They demonstrated that import was blocked when antibodies against the cytosolic domain of Tom22 were added. Interestingly, the interaction of precursors with the receptors Tom20 and Tom70 was still detectable (Kiebler *et al.*, 1993). Kiebler *et al.* noticed that the inhibitory effect of the anti-Tom22 antibodies was almost identical to that of anti-Tom20 antibodies. Therefore, it was suggested that Tom22 acts in the transfer of precursors between receptors and the import pore, perhaps by transiently binding to the positively charged side of the amphipathic  $\alpha$ -helix (Kiebler *et al.*, 1993).

*In vitro* pull-down experiments performed by Brix *et al.*, showed binding of presequence-containing precursors from a mixture of  $^{35}\text{S}$ -labeled preproteins that were loaded onto a Ni-NTA resin column with immobilized cytosolic domain of Tom22 (Brix, Dietmeier, and Pfanner, 1997). The experiment indicates stronger interaction of Tom22 with cytochrome c1 than with the presequence (pSu9-DHFR). However, in presence of a presequence peptide, reduced binding of the above mentioned clients to Tom22 cytosolic domain was noticed. This suggests competition in binding of the presequence peptide and the client proteins to Tom22 cytosolic domain. Additionally, repeating the pull-down experiment at low salt concentration increased

the binding of the presequence, a pSu9-DHFR construct, to Tom22. This result suggests an ionic interaction of Tom22 and the precursor protein (Brix, Dietmeier, and Pfanner, 1997). However, in a systematic study on the fungus *N. crassa* 15 out of the 19 negative charges of Tom22's cytosolic domain were removed, yet the import capacity was not dramatically altered (Nargang et al., 1998). Subsequent studies therefore addressed the identification of the Tom22 recognition sequences on precursor proteins.

Decomposing a protein's sequence into short, overlapping peptides followed by binding assays (peptide scan) can allow identification of those regions that are involved in protein-protein interaction. Using such an approach it was suggested that the amino-terminal region of CoxIV (first 10 peptides) preferentially bound to the cytosolic domain of Tom22 (Brix et al., 1999). The binding region was shifted by approximately 9 amino acids away from the very N-terminus compared to the recognition site of Tom20. However in the same study, no specific interaction between peptides covering the phosphate carrier (PiC) sequence and the cytosolic domain of Tom22 could be detected: instead, scattered spots over various regions of the preprotein were observed (Brix et al., 1999). All peptides that bound to Tom22 had a positive charge, opening the question whether the observed interactions were specific.

More recently, Shiota *et al.* performed *in vivo* cross-linking study using the BPA cross-linking agent on the cytosolic domain of Tom22, located either at residue 48 (acidic region) or at residue 75. They showed that in the presence of over-expressed presequence the cross-linking of Tom22 and Tom20 is reduced, but only if the BPA is at position 48, not if it is at position 75 (Shiota et al., 2011). They also observed cross-linking of the presequence pSu9-DHFR to Tom22 when the BPA-tag was in the acidic (negatively charged) region of Tom22, at residue 48. This data points out the interaction of Tom22 and the presequence. However, considering the distance span of BPA cross-linking agent of 9.6 Å, it cannot be ruled out that the pSu9-DHFR-Tom22 cross-linking is due to proximity of the Tom20-bound presequence to Tom22, rather than a direct binding of the presequence to Tom22.

Considering the experimental observations made, two different mechanisms for the function of Tom22 within the TOM complex have been proposed:

- (i) the preproteins with a presequence are first recognized by Tom20 and subsequently by Tom22 (Pfanner and Geissler, 2001) or
- (ii) Tom20 and Tom22 simultaneously bind to the presequence, where the hydrophobic side would be bound by Tom20 and the positively charged side would be bound by Tom22 (Shiota et al., 2011; Araiso, Imai, and Endo, 2022). The amphipathic nature of the  $\alpha$ -helix formed by the presequence could enable different types of interactions to occur during the precursor import.

To distinguish between these two models we will try to investigate the molecular aspects of the interaction between Tom22 and mitochondrial precursors in more detail.

### **Interaction of Tom22 and Tom20.**

Tom22 was initially proposed to interact with the Tom20 by hydrophobic surface patches, similarly to the interaction of Tom20 with the hydrophobic side of the amphipathic helix formed by the presequence (Brix, Dietmeier, and Pfanner, 1997). More recently, the interaction of yeast Tom22 and Tom20 has been mapped from cross-linking experiments (Shiota et al., 2011). Using the BPA cross-linking agent which can reach up to 9.6 Å, on different positions of the cytosolic domain of Tom22, the majority of cross-linking products involved a region of Tom22 with conserved acidic residues (Shiota et al., 2011) (see alignment in Fig.4.5; residues 42-52). From the Tom20 side, the cross-linking of Tom22 to Tom20 was successful when the BPA-tag was near the presequence binding site of Tom20 (Shiota et al., 2011). These experiments suggest the

proximity of the presequence and Tom22 binding sites on the Tom20 receptor and that probably negatively charged residues of Tom22 might be involved in binding to Tom20.

It is thought that Tom22 plays a role of the central TOM receptor and in recruiting Tom20 and Tom70. Tom22 was found to be essential for the *in vitro* assembly of the Tom20 receptor into the TOM core complex. More precisely, in the absence of the Tom22 cytosolic domain, the assembly efficiency of Tom20 is decreased (Yamano et al., 2008b).

Interestingly, deletion of either Tom22 or Tom20 cytosolic domains leads to similar effect of presequence import impairment (Yamano et al., 2008b). Keeping in mind the requirement of Tom22 for the recruitment of Tom20 to the TOM complex (Yamano et al., 2008b), it is unclear whether Tom22 is directly involved in the presequence recognition or if it modulates the binding of Tom20 and the presequence.

#### 1.1.4 Tom70: receptor for precursors with internal-targeting signals

The third receptor, Tom70, is N-terminally anchored in the OM by a transmembrane helix, and it plays a major role in the import of the large family of hydrophobic mitochondrial metabolite carriers (Wiedemann, Pfanner, and Ryan, 2001). Carrier proteins of the IM are delivered to Tom70 by cytoplasmic chaperones and it is assumed that they are handed to Tom22 before getting imported through the Tom40 pore. Recognition and interaction of Tom70 with the cytosolic chaperone-precursor complex is crucial for the import through the outer membrane (Young, Hoogenraad, and Hartl, 2003). The Tom70 receptor interacts with the molecular chaperones Hsp70/Hsc70 (both yeast and mammalian receptors) and Hsp90 (mammalian receptors) via a cleft formed by one of Tom70's N-terminal tetratricopeptide repeats (TPR), which binds the highly conserved EEVD motif at the C-terminus of the Hsp70/Hsc70/Hsp90 chaperone. The binding mode is known in detail from the crystal structure of *S. cerevisiae* Tom71 (a Tom70 paralog in *S. cerevisiae* that arose from a gene duplication event) (Li et al., 2009). It is proposed that the hydrophobic residues upstream from the chaperone binding site direct the specificity of the chaperone binding (Scheufler et al., 2000).

Interestingly, Tom20 also interacts with the TPR clamp domain of Tom70 via its conserved C-terminal DDVE motif (note the similar physico-chemical properties to Hsp70's EEVD motif). While this interaction is not necessary for the import of the receptor-protein precursors to the OM, it could suggest a competition mechanism for the Hsp70 binding site on Tom70 (Fan et al., 2011).

##### Structural studies of the cytosolic domain of Tom70.

In the crystal structure of the cytosolic domain of the yeast protein, Tom70 forms a homodimer with the interface formed mostly by the N-terminal domain (PDB: 2gw1) (Wu and Sha, 2006). A carrier precursor binding pocket in the C-terminal domain was proposed from this structure and from the conservation of the residues forming a C-terminal cleft (Wu and Sha, 2006). Additional structural information is provided from crystal structures of yeast Tom71 which shares 53% sequence identity with Tom70 and an overlapping function (Koh, Hájek, and Bedwell, 2001; Kondo-Okamoto, Shaw, and Okamoto, 2008). Tom71 has been crystallized as a monomer in the apo form (PDB: 3fp3; Li et al., 2009) and in presence of a short C-terminal Ssa1 peptide (yeast Hsp70 (Mashaghi et al., 2016)) (PDB: 3fp2 and 3lca). Comparing three Tom71 crystal structures reveals that the C-terminal domain of the Tom70/Tom71 receptor is adopting different conformations. An apparent *closed* state is observed in the apo Tom70 protein (PDB: 2gw1) while Tom71 in complex with the Ssa1 peptide is observed either adopting an *open* (PDB: 3fp2 (Li et al., 2009)) or an *intermediate* (PDB: 3lca; Li, Cui, and Sha, 2010) conformation.

Comparison of the *closed* (apo-) Tom70 and the *open* (Hsp70 peptide-bound) Tom71, suggests that the interaction with the cytosolic chaperone would induce significant conformational changes. This conformational

change was proposed to bring the precursor binding pocket closer to the chaperone-bound precursor (Li et al., 2009). Somehow controversially, SAXS studies (Mills et al., 2009) point out that the apparent Tom70 *closed* dimer conformation (PDB: 2qw1) could be interpreted differently; the structure may correspond to two Tom70 proteins in the *open* conformation, crossing and obscuring their chaperone and precursor binding sites.

The two conformations of Tom71 receptor observed with the bound Ssa1 peptide (*intermediate* and *open*) suggest flexibility of the receptor that was proposed to be needed for the interaction with different precursor proteins (Li, Cui, and Sha, 2010). The difference between the two conformations is in the 12° rotation of the C-terminal domain around the helix A7, a helix present in the interface of the N- and C-terminal domains of Tom70 (see Figure 1.4).

### Client binding sites on Tom70.

The potential clientome of the Tom70 receptor seems to be quite broad since its role in import of many different precursor proteins has been shown. First discovered and most abundant type of Tom70 clients are the precursors of the carriers of the IM. Most of the precursor proteins targeted to the TIM22 insertase, among which are the carriers of the IM, are recognized by the Tom70 receptor through the targeting elements scattered over the primary structure of a precursor (Hines and Schatz, 1993; Brix et al., 1999). However, it is still unclear what are the exact characteristics of these internal targeting signals and how the specificity in recognition is achieved. Affinity measurements, by determining fluorescence anisotropy with labeled precursor-peptides, showed binding of Tom70 to phosphate-carrier precursor peptide (PiC) with a  $K_D \sim 70 \mu\text{M}$  which is not affected by the presence of the Hsp70 EEVD containing-peptide (Mills et al., 2009).

It is assumed that the internal targeting elements cooperate in binding to the receptor Tom70 (Wiedemann, Pfanner, and Ryan, 2001; Backes et al., 2018; Kreimendahl et al., 2020). A reason for the presence of these scattered recognition/interaction sites along the precursor could be that several Tom70 receptors keep the long precursors unfolded and import-competent during their translocation into mitochondria (Backes et al., 2018). From this, a role of Tom70 as a membrane bound co-chaperone was proposed.

Tom70-dependent import of presequence-containing precursor proteins (Yamamoto et al., 2009; Melin et al., 2015) and of precursors targeted to mitochondrial matrix, carrying internal-matrix-targeting-signal-like sequences (iMTS-L) (Backes et al., 2018) was shown. In case of  $\beta$ -barrel protein precursors, the targeting signal has been shown to be a  $\beta$ -hairpin motif, which is primarily recognized by the Tom20 receptor (Jores et al., 2016). Interestingly, formation of cross-linking adducts was also shown with other components of the TOM machinery, such as Tom40, Tom22 but also Tom70 (Jores et al., 2016), indicating possible recognition of the  $\beta$ -barrel protein precursors by the Tom70 receptor.

The Tom70 binding site for a carrier-precursor was first suggested from the structure of the apo- protein and from the conservation of the hydrophobic residues forming the C-terminal cleft of the cytosolic domain of Tom70 (see Figure 1.4 cyan surface) (Wu and Sha, 2006). This carrier-precursor binding site comprises the residues of so-called Tom70 core domain, the smallest cytosolic domain of Tom70 that was sufficient for client binding (Brix et al., 2000). This core domain is stably folded and contains a single tandem TPR motif. *In vitro* cross-linking studies with purified Tom70 were established in order to identify residues involved in presequence recognition. A pALDH presequence was modified with a BPA on the hydrophobic (pL<sub>19</sub>B) side of the amphipathic helix as a photoreactive peptide. Tom70 fragments crosslinked to pALDH were identified by mass spectrometry after trypsin digestion of the photoadducts (Melin et al., 2015). Tom70 residues Met216, Met551, Ile604 and Met617 were found to bind the presequence. The importance of these

residues was confirmed by observed impairment of the Mdl1 (a presequence-containing IM protein) import in mitochondria with Tom70(M551R) variant (Melin et al., 2015). These presequence binding residues are located in a lower part of the cytosolic C-terminal cleft, compared to the proposed carrier-precursor binding site (see Figure 1.4 residues in magenta) (Melin et al., 2015).

Interestingly, recent studies demonstrated interaction of human Tom70 and the Orf9b protein from SARS-CoV-2. Cryo-EM and crystal structures were obtained that show binding of this viral peptide in both the upper, carrier-precursor binding site and the lower, presequence binding site, occupying the full C-terminal cleft of the Tom70 cytosolic domain. (see Figure 1.4; (Gao et al., 2021; Gordon et al., 2020a). This interaction appears to be hydrophobic, and Orf9b adopts an  $\alpha$ -helical conformation. Additionally, Gao and colleagues determined the dissociation constant of the Orf9b peptide and Tom70 in the nanomolar range and showed a decreased binding affinity between the Hsp90 EEVD motif and Tom70 when Tom70 was associated with Orf9b(Gao et al., 2021).

Thus, Tom70 is involved in binding of many different precursor proteins however, detailed atomic-resolution characterization of the binding site, or potentially multiple binding sites, is limited.

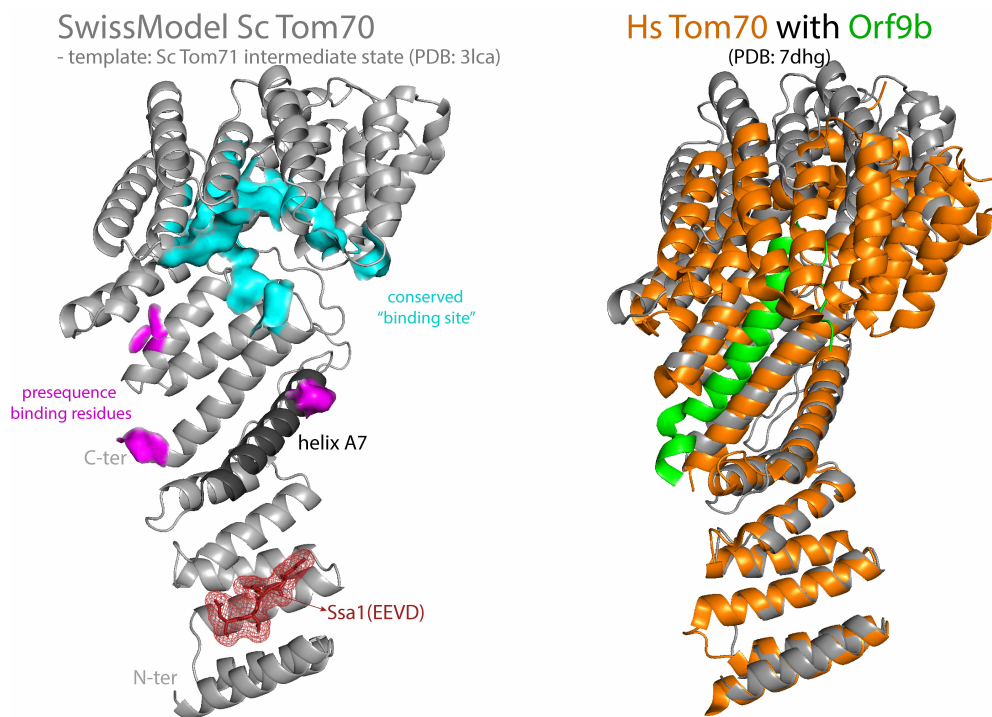


FIGURE 1.4: **Structural models of the cytosolic domain of Tom70 receptor.** Model of the cytosolic domain of the yeast Tom70 protein obtained with SwissModel server (<https://swissmodel.expasy.org/>) using the crystal structure of yeast Tom71 protein in presence of a short C-terminal Ssa1 peptide (PDB:3lca, an intermediate state) (gray structures on both sides). The Ssa1 peptide is shown in red to indicate the chaperone binding site. The surface area in cyan indicates the proposed carrier-precursor binding site (Wu and Sha, 2006) and the presequence-binding residues are shown in magenta. Helix A7 (shown in black) is in the interface of the N- and C-terminal domains of the cytosolic Tom70 domain. The orange colored structure is the crystal structure of the human Tom70 protein bound to Orf9b SARS-CoV-2 protein (PDB:7dhg). The C-terminal domain of the Orf9b-bound Tom70 seem to be tilted around the helix A7, which could correspond to the more closed state, compared to the Ssa1-bound intermediate state.

### 1.1.5 Chaperoning in the inter-membrane space

After passage of the precursor proteins through the TOM complex, several routes are possible, depending on the nature and final destination of the precursor protein (see Fig. 1.1). For membrane proteins *en route* to the TIM22 insertase ( $\alpha$ -helical, such as mitochondrial carriers) or the SAM complex ( $\beta$ -barrel, such as VDAC)

the small TIM chaperones are in charge of guiding the presequences to the respective insertases. In their apo state, these chaperones form hetero-hexameric complexes of ~65 kDa, composed of alternating subunits of Tim8 and Tim13 [TIM8·13] (Beverly et al., 2008) or Tim9 and Tim10 [TIM9·10] (Webb et al., 2006), or Tim9, Tim10 and Tim12 [TIM9·10·12] (Gebert et al., 2008). It is interesting to note that the TIM chaperones are in continuous exchange, whereby subunits co-exist between the hexameric state (predominant at ambient temperature) and monomeric forms. At room temperature, approximately 10% of the Tim9 and Tim10 subunits coexist as monomers, together with the hexameric TIM9·10 complex. NMR experiments have established that the integration of subunits into the hexamer and release of subunits from the hexamer is slow, on time scales of many tens of minutes (Weinhäupl et al., 2021). As the only known chaperone system of the mitochondrial IMS, TIM chaperones are crucial for the recognition and transfer of most of the mitochondrial membrane precursor proteins (Morgenstern et al., 2017). They recognize a broad range of membrane proteins and transfer them in an unfolded state from the mitochondrial outer membrane pore (TOM, with its central pore formed by Tom40), through the aqueous mitochondrial intermembrane space, towards the insertases of the inner membrane (TIM22) or outer membrane (SAM) (Koehler et al., 1999; Bauer et al., 1996; Bauer et al., 2000; Rehling et al., 2001; Paschen et al., 2003; Kozjak et al., 2003; Gentle et al., 2004; Lithgow and Schneider, 2010).

The structural basis of the chaperone action of the small TIMs has been resolved first by crystal structures of the apo TIM9·10 chaperone (Webb et al., 2006) and the TIM8·13 chaperone (Beverly et al., 2008), and then later by structural and dynamical investigation of the holo complexes, performed in the host group (Weinhäupl et al., 2018). In particular, Weinhäupl *et al.* obtained the first structure of a complex of a small TIM chaperone, TIM9·10, with the mitochondrial guanosine diphosphate (GDP)/guanosine triphosphate (GTP) carrier (Ggc1). The structure, composed of two chaperone complexes holding one precursor protein, revealed a highly dynamic ensemble of Ggc1 conformers that form multiple short-lived and rapidly interconverting (<1 ms) interactions with a hydrophobic binding cleft of the chaperone. The TIM9·10-Ggc1 complex can be described as a “fuzzy complex,” in which the high overall affinity is driven by a multitude of individually weak interactions with the hydrophobic TM parts of its clients.

TIM8·13 and TIM9·10 are structurally highly similar but have different substrate binding preferences. Although TIM8·13 is not essential in yeast (Koehler et al., 1999), yeast cells depleted of Tim8 and Tim13 show conditional lethality (Paschen, 2000). In addition, mutations in the human Tim8a protein have been identified as the cause of a neurodegenerative disorder known as Mohr-Tranebjærg syndrome or deafness-dystonia-optic neuropathy syndrome (Roesch et al., 2002; Kang et al., 2019).

*In vivo* experiments, predominantly in yeast, have identified mitochondrial membrane proteins whose biogenesis depends on small TIM chaperones. TIM9·10 is believed to interact with all members of the mitochondrial carrier (SLC25) family, which comprises more than 50 members in humans, such as the ADP/ATP carrier (Aac in yeast) and GTP/GDP carrier (GGC1); TIM9·10 furthermore transports the central components of the TIM22 and TIM23 insertion machineries (Tim23, Tim17, Tim22) as well as outer-membrane  $\beta$ -barrel proteins (Morgenstern et al., 2017). TIM8·13 has been shown not to bind the inner-membrane proteins ADP/ATP carrier (Aac) nor Tim17 (Paschen et al., 2000), instead it was shown to bind the precursors of the inner-membrane proteins Tim23 (Paschen et al., 2000; Leuenberger et al., 1999; Davis et al., 2000) and Ca<sup>2+</sup>-binding aspartate-glutamate carriers (Roesch et al., 2004), as well as the outer-membrane  $\beta$ -barrel proteins VDAC/Porin, Tom40 (Hoppins and Nargang, 2004) and Tob55/Sam50 (Habib et al., 2005). The inner-membrane proteins that have been reported to interact with TIM8·13 have a hydrophilic domain in addition to trans-membrane domains but this does not hold true for the outer-membrane  $\beta$ -barrels.

The structural basis for the different specificity of TIM9·10 compared to TIM8·13 has not been elucidated yet.

## 1.2 How do chaperones bind (partly) unfolded client proteins?

During my PhD thesis I contributed in writing a review article on the molecular mechanisms that chaperones employ for stabilizing their clients. The motivation for writing this literature overview came from our studies on the chaperones of the mitochondrial intermembrane space, the TIM chaperones. The review article discusses the interactions that contribute to the complex formation of several different chaperones and their clients, along with the known structures and physico-chemical properties of these complexes. The main focus is on interactions of the dynamic client–chaperone complexes where a client is at least partially unfolded and the chaperone is ATP-independent.

The way how proteins interact with other proteins spans a wide range of behaviors. The interaction of folded and rather rigid globular proteins is certainly the most-studied type of interaction. Seen in thousands of crystal structures, it is often the enthalpy of favorable interactions (e.g. hydrophobic or charge-charge interactions) that drives the binding. On the extreme other end of the spectrum, proteins may also interact - and tightly interact - even though they lack any kind of stable structure. Cases of such tight complexes of intrinsically disordered proteins have been reported (Borgia et al., 2018). In such cases, it may be complementary electrostatic charges of the interacting protein which drives the interaction. The complexes of chaperones or receptor domains along import pathways (such as in mitochondrial import) and their client proteins is a field that only starts being studied at the atomic level. Therefore, we only begin understanding these interactions.

The full article can be found in the Appendix B.1 of this manuscript: I. Sućec, B. Bersch and P. Schanda. **How do chaperones bind (partly) unfolded client proteins?** *Front. Mol. Biosci. Rev.*

The section "Chaperoning in the mitochondrial intermembrane space" of this review article is giving an overview of the studies of TIM chaperones and is discussing the possible factors that may contribute to the balance of promiscuity and specificity. Our work on TIM's (see Chapter 5: Results) demonstrates the client specificity of the two chaperones and proposes the subtle difference in the biophysical properties of these chaperones to be the key in balancing the interactions leading to TIM's promiscuity but also specificity.

## Chapter 2

# Introduction 2.: Preparing chaperone-client protein complexes for biophysical and structural studies

The motivation for this section and a review on the experimental approaches for preparing chaperone–client complexes came from the challenges we were facing in our experiments when studying TIM chaperones bound to an unfolded client. In our biological system, the chaperone is a folded protein complex stabilized by intramolecular disulphide bonds and interactions between the individual protein subunits, while the client is a highly hydrophobic membrane protein, unstable and aggregation-prone in the absence of membrane, detergent or a specific molecular chaperone. The challenge in making these complexes *in vitro* is to find the experimental condition in which the process of client aggregation is slow enough to enable the correctly folded TIM chaperone to bind. Formation of the chaperone-client complex then stabilizes the highly hydrophobic client and prevents its aggregation in the aqueous environment. Unfortunately, experimental conditions under which the hydrophobic, unfolded client is stable (detergent, denaturant, high temperature) are not compatible with an active, folded chaperone protein. Optimizing our experimental approaches for TIM chaperone–client formation was an essential step for structural and biophysical studies of these complexes that enabled characterization of interactions driving chaperone–client specificity (Chapter 5 Results).

This introductory chapter gives a literature overview of several experimental approaches that were tested and shown successful for formation of different chaperone–client complexes used for structural and biophysical studies. We discuss these different approaches, the reason of their success in specific cases, and summarize what should be taken into consideration when preparing such complexes.

The experimental approaches of preparing chaperone–client complexes reviewed here are about to be published as a book chapter: I. Sućec, P. Schanda, **Preparing chaperone–client protein complexes for biophysical and structural studies**, a chapter in *Biophysics of Molecular Chaperones: Function, Mechanisms and Client Protein Interactions* – in press.

Protein folding is a process in which a polypeptide chain of a certain protein assembles in its biologically functional (native) tertiary structure. It is believed that during the folding of majority of proteins, in the absence of chemical denaturants, the first step is the hydrophobic collapse of the polypeptide chain followed by the secondary structure formation (Sadqi, Lapidus, and Muñoz, 2003). Hydrophobic collapse limits the number of the potential protein conformations that could be adopted by the polypeptide chain however,



most of the proteins fold through multiple transient intermediate folding states (review on protein folding mechanisms (Bartlett and Radford, 2009)). This folding mechanism works well for small soluble proteins where the hydrophobic residues are buried during the folding into the protein core and hydrophobic side chains remain exposed to the aqueous environment.

Many proteins are prone to misfolding or unfolding and aggregation in the aqueous environment of the cell, either due to their complex ternary structure, cellular localization, molecular crowding, cellular stress, or their intrinsic hydrophobic nature (Díaz-Villanueva, Díaz-Molina, and García-González, 2015; Ellis and Minton, 2006). Molecular chaperones are involved in biogenesis of such proteins, and in ensuring their intactness throughout their life cycle (Hartl, Bracher, and Hayer-Hartl, 2011; Kim et al., 2013; Sinnige, Yu, and Morimoto, 2020). While there are many different types of chaperones with various cellular roles, essentially all of them are characterized by a basic "holdase" function, i.e. the ability to bind to proteins that are generally in some non-native conformation – often these are called "client" proteins. Chaperones may have more "active" roles in addition to the basic function: "foldases" assist their client protein in folding to its native state (reviewed in (Balchin, Hayer-Hartl, and Hartl, 2020)), and "disaggregases" help in dissociating protein aggregates (reviewed in (Mattoo and Goloubinoff, 2014)). Understanding how these complexes form, what (residual) structure the client proteins have when bound, and how the client protein may eventually be released again from the chaperone is essential for resolving questions in biogenesis of cells and organelles.

Detailed study of chaperone—client complexes on the molecular level is challenging for several reasons. The interactions involved in the complex formation are often highly dynamic, i.e. binding often relies on avidity of multiple interactions (see reviews (Burmam and Hiller, 2015; He and Hiller, 2019; Hiller, 2019; Macošek, Mas, and Hiller, 2021; Sučec et al., 2021)). Furthermore, chaperones themselves are often large, up to the megadalton range, and can adopt multiple states (Bracher and Hartl, 2013; Skjærven et al., 2015). All these factors make their experimental characterization difficult. Moreover – and this is the focus of the present chapter – chaperone—client complex formation for any experimental study turns out to be challenging. The proteins to be bound to the chaperone are in many cases aggregation-prone. This is particularly the case for membrane proteins in aqueous solutions (Heyden et al., 2012; Schiffrin, Brockwell, and Radford, 2017), but it applies also to many soluble proteins in the state in which they interact with chaperones, i.e. in partially unfolded states (Capaldi, Kleanthous, and Radford, 2002; Espargaró et al., 2008; Tsigotaki et al., 2018). An experimental way to keep aggregation-prone proteins in solution is to add denaturant, but under such conditions many chaperones are also destabilized, which impairs complex formation. Moreover, the shielding of the client by the denaturant also hampers complex formation. Any experimental work that aims to characterize chaperone—client complexes *in vitro* needs to find a set of conditions that enable the formation of the complex. In this chapter, we review the possibilities for such complex formation, starting from some general considerations, and highlighting the different possibilities with recent examples from structural and biophysical studies. The cases we present are to be considered as example cases, with the aim to show the range of methods that have been proposed, rather than an exhaustive review of all chaperone-client complexes that have been reported.

## 2.1 Client-chaperone complex formation

### 2.1.1 General considerations

*In vivo*, client—chaperone complexes form either at specific locations where a client protein emerges – for example at the exit of the ribosome tunnel, or the exit of translocation pores, or they form in a situation where the chaperone is rather abundant compared to the client. Examples for the former are the trigger

factor chaperone, associated to the ribosome (Ferbitz et al., 2004); Trigger factor's association with the ribosome provides a shielded environment for the newly synthesized peptides where a co-translational folding of client domains could occur, with possibly several chaperone molecules bound to the client polypeptide chain (Saio et al., 2014). Another example are the TIM chaperones that localize at the membrane translocation pore TOM in mitochondria, presumably in a position ready to bind membrane-protein precursor polypeptides, as they emerge in the intermembrane space (Shiota et al., 2015). Some chaperones are abundant in the cell; for example, the family of 70-kDa heat-shock proteins (Hsp70) on its own is estimated to correspond to up to 3% of the total protein mass in eukaryotic cells under non-stress conditions Finka2011a, Finka2013. Under stress conditions, the cellular chaperone concentration or activity can be up-regulated, such as in the case of many of the heat-shock proteins (Bakthisaran, Tangirala, and Rao, 2015; Young et al., 2004) and the Spy chaperone (Quan et al., 2011). In cases other than binding nascent polypeptide chains, stress response chaperones are upregulated as a response to misfolding or aggregation of client proteins that become structurally unstable or denatured due to stress conditions like heat, pH changes, osmotic- and oxidative stress, and nutrient starvation (Alam et al., 2021), leading to increased chaperone cellular levels or their activation.

These cellular conditions and environments facilitate the encounter of chaperones and client proteins, and the formation of complexes with affinities that range all the way from low-affinity transient interactions to tight binding with nanomolar affinity.

For an *in vitro* preparation procedure, the complex formation starts from a very different point, namely, in general, from purified proteins (chaperone, client protein). The question is, thus, how one can find suitable conditions for the encounter of the two proteins and complex formation under native conditions.

Formation of a complex between a chaperone (C) and a client protein (P) requires that, under the chosen conditions, the P-C complex is thermodynamically the most stable state, or that it is kinetically trapped state. The desired outcome, namely formation of the P-C complex, is challenged by other reactions that the protein P may undergo with itself (intramolecular processes such as folding and unfolding), with other copies of itself (such as oligomerization or, quite commonly, aggregation) or with other components in the system (e.g. other protein binding partners, a lipid bilayer, small molecules). In order to understand how the chaperone—client-protein complex may form, it is instructive to understand these possible reactions. A simplified scheme of this set of reactions is shown in Figure 2.1. The rate of these reactions differs from one chaperone—client pair to another. Therefore, understanding this set of reactions provides a way of seeing the possibilities that the experimentalist has to prepare P-C complexes.

Any protein P exists in aqueous solution in an equilibrium between folded conformations ( $P_{\text{folded}}$ ) and possibly a multitude of different partially unfolded intermediate states ( $P_{\text{int}}$ ) and an ensemble of unfolded states ( $P_{\text{unfolded}}$ ; Figure 2.1). The distinction we make here between folded, partially unfolded and fully unfolded is somewhat arbitrary, as this is rather a continuum of states, and even the "folded state" is to be seen as a dynamic ensemble of conformations. For the sake of the discussion, we keep this simplified scheme here. The kinetics and thermodynamics of this equilibrium on the continuum from folded to unfolded is protein-dependent. Many proteins are able to fold only in a proper environment or in the presence of proper binding partners – this is particularly true for the vast majority of membrane proteins, which are unable to fold in aqueous solution.

It is primarily the partially or fully unfolded states that are prone to aggregation: by the encounter of several copies of the protein ( $(P)_n$ ) an aggregated state,  $P_{\text{aggregated}}$ , may form. The (partially or fully) unfolded states are generally more prone to aggregation than the folded states. The reason for this finding is that in the folded states of globular proteins, hydrophobic parts are mostly buried within the core, and they may

become exposed upon unfolding. Likewise, the polypeptides that finally get inserted into a lipid bilayer as membrane proteins are most often highly hydrophobic. The aggregated state that brings together the hydrophobic parts is often highly stable.

Molecular chaperones enter this picture of protein folding, unfolding, aggregation, and possibly insertion of the protein (into a membrane or a complex) because they can interact with one or several of these states and protect them from aggregation. They may also play an active role by lowering energy barriers in folding processes (Chakraborty et al., 2010; Hartl, Bracher, and Hayer-Hartl, 2011). From the practical standpoint of this manuscript, namely for the experimental preparation of chaperone—client complexes, it is important to understand where the chaperones intervene, and what the relevant kinetic rate constants are, compared to the competing processes, and how the chaperone—client complexes can be thermodynamically favored. Figure 2.1 schematizes the possible reactions generally, and the following sections then discuss the situation for several chaperone—client complexes, highlighting which practical experimental approach was successful for obtaining complexes.

Most chaperones bind to parts of proteins that are (partly) unfolded and expose hydrophobic residues (Rüdiger et al., 1997); this realization is linked to reports of chaperones that bind to locally "frustrated" sites in proteins (He et al., 2016a; He and Hiller, 2018), which are prone to local unfolding. P–C complex formation, thus, tends to be favored when the  $P_{\text{folded}}$  state is disfavored over the (partially) unfolded ones. For the preparation of P–C complexes that start with purified soluble protein P, it is, thus, often helpful to increase the population of (partially) unfolded states. *In vitro*, destabilization of  $P_{\text{folded}}$  can be achieved by different means. One way is by acting on the chemical properties of the protein sequence itself, namely through destabilizing mutations or by truncation of the sequence. Such changes disrupt interactions that form in the folded state. Similarly, the disruption of intramolecular disulfide bonds can equally favor the unfolded states. Other factors that can be utilized to act on the  $P_{\text{folded}}/P_{\text{int}}/P_{\text{unfolded}}$  equilibrium are the addition of chemical denaturant or application of high or low temperature and pH or high hydrostatic pressure.

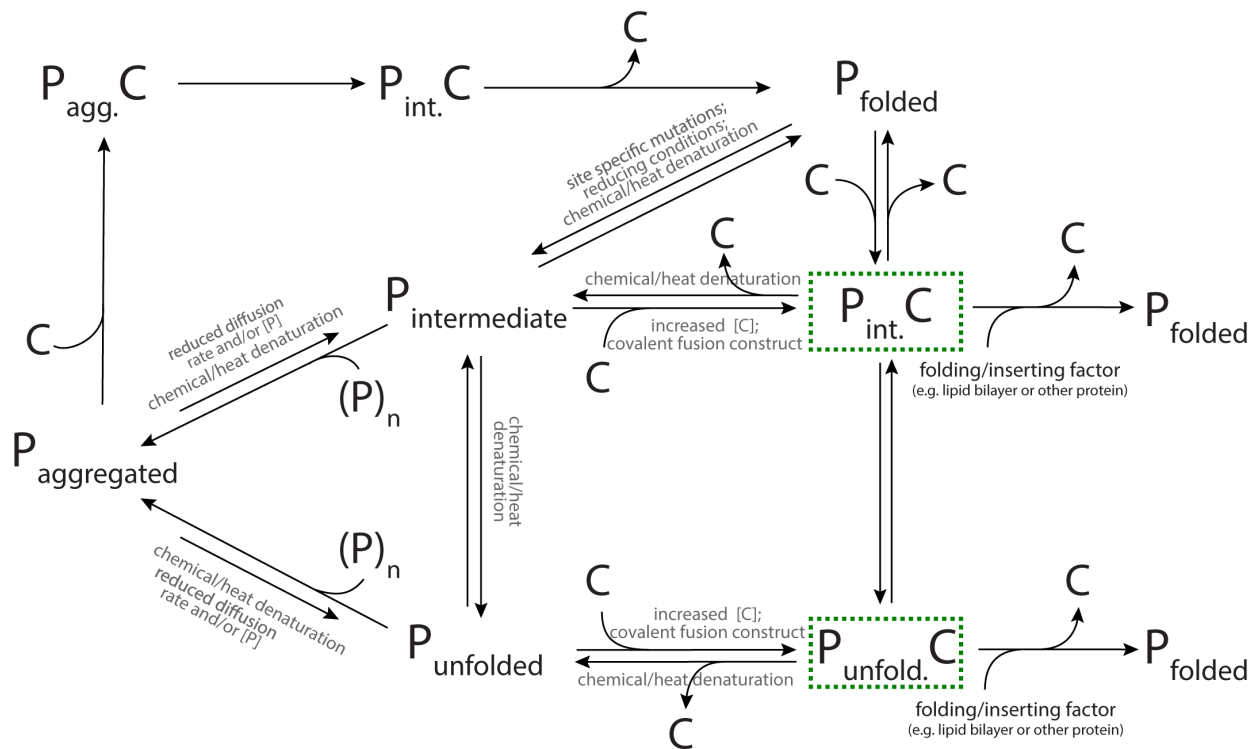


FIGURE 2.1: **Schematic overview of the reactions that may occur in a system composed of a chaperone (C) and a client protein (P).** The preparation of complexes of chaperones and client proteins may proceed by direct binding of a chaperone to a folded state or to (partially or fully) unfolded states. In the formed complexes the protein P differs generally from the folded state, and the complexes are depicted by green dashed boxes. The distinction between an "unfolded" client and a partially folded one ("intermediate" or "int.") are arbitrary, and shown here to denote that there is a range of possible residual structures in chaperone-bound clients. Several other reactions involving P can occur and compete with P-C complex formation. These are the folding of P to a state that is not folding-competent, the aggregation of P or, if available, the interaction of P with binding partners or a membrane (right in this scheme). Note that also the apo protein in its folded or partially unfolded state may bind to such an additional binding partner (not shown). The various reaction equilibria and their kinetics can be modulated by changing the experimental conditions or by modifying the client protein. The effect of some of these conditions on the reaction equilibria is indicated near the arrow. The condition shifts the equilibrium in the direction of the arrow. For example, the rate of aggregate formation can be reduced by immobilizing P (section 2.2.5) or reducing the concentration of P; the folding/unfolding equilibrium can be shifted towards unfolded states by e.g. denaturing conditions, reducing conditions or mutations of P (section 2.2.3). From one P/C system to another, the relative importance of these reactions, i.e. the relative rate constants and thermodynamic constants, can vary largely.

These different approaches – either the mutation/truncation of P or the change in buffer conditions, temperature or pressure – differ in an important aspect: all the changes in solvent, temperature or pressure can also affect the chaperone (e.g. unfold it) and its ability to bind client proteins. Thus, while such changes in solvent conditions may achieve the desired increase in  $[P_{\text{unfolded}}]$  or  $[P_{\text{int}}]$ , the formation of the complex ( $P_{\text{int}}-C$ ,  $P_{\text{unfolded}}-C$ ) may be disfavored, i.e. the reaction  $P_{\text{unfolded}} + C \rightarrow P_{\text{unfolded}}-C$  or  $P_{\text{int}} + C \rightarrow P_{\text{int}}-C$  either has a very low reaction rate or its equilibrium is on the left-hand side of these equations. Forming the P-C complex, thus, requires the removal of the conditions that denature/destabilize the chaperone (such as denaturant or high temperature).

A practically very important reaction is the one that converts soluble protein into insoluble aggregates. For the formation of soluble P-C complex, this reaction is a major threat. The essential question for *in vitro* complex formation is whether the formation of aggregates is thermodynamically and/or kinetically favored over the formation of the P-C complex. The coupled equilibrium between aggregated, free and chaperone-bound protein, thus has a few important parameters: the kinetics and thermodynamics of the

aggregation of soluble P, and the kinetics and thermodynamics of P–C complex formation. The thermodynamic stabilization of the soluble form over the aggregated form can be achieved by the addition of denaturant or high hydrostatic pressure. However, as discussed in the last paragraph, if such conditions are incompatible with keeping the chaperone intact, this strategy does not lead to productive formation of P–C complex. Therefore, in most cases, one can rather act upon the kinetics of aggregate formation vs. the kinetics of complex formation, as discussed in the following.

As aggregation is an intermolecular process involving many copies of P, the kinetics of formation of aggregates is concentration dependent. It can, therefore, be slowed down when the concentration, [P], is reduced, or the rate of diffusion of P is reduced, such that encounter of copies of P is slowed down. As the kinetics of P–C formation is also concentration dependent (namely on both [P] and [C]), a rather high [C] and low [P] helps favoring the reaction  $P + C \rightarrow P-C$  over the aggregation reaction  $P + P_n \rightarrow P_{n+1}$ . The experimentalist's options in this sense include drip-diluting the protein from a denatured state (e.g. from a solution containing urea or other chaotropes) to a buffer that contains the chaperone and sustains the chaperone's integrity. Another possibility is to have the protein synthesized in the presence of chaperone, such that a nascent protein would encounter chaperones before having the chance to encounter other copies of P and aggregate. This strategy is exploited in *in vitro* protein synthesis systems containing chaperones (Rampelt et al., 2020; Schneider et al., 2010; Weinhäupl et al., 2018).

The other parameter that the experimentalist can influence in order to slow down the aggregation process is by limiting its diffusion coefficient and thus limit the probability of encounter of copies of P. A useful approach that exploits this principle is based on immobilizing the protein to e.g. a resin via an affinity tag. Doing so, the encounter of proteins is strongly reduced (the diffusion coefficient is essentially infinitely small), and added chaperone C can bind with P, without having the competing process of aggregation of P.

The P–C complexes have often a substantial degree of flexibility, as revealed by numerous atomic-level studies reviewed in (Hiller and Burmann, 2018; Sućec et al., 2021). The client protein can exist in a range of conformational states when bound to C, from fully unfolded to largely folded (see the equilibrium of  $P_{\text{int}}-C$  and  $P_{\text{unfolded}}-C$  in Figure 2.1).

The release of P from the P–C complex is an important parameter that determines if biochemical, biophysical and structural studies of P–C can be performed. The different states of P have different affinity to C. Although not strictly necessary, it can be assumed that less native-like states of P have a higher propensity to bind C. Accordingly, the off-rate,  $P-C \rightarrow P_{\text{int}} + C$  or  $P-C \rightarrow P_{\text{unfolded}} + C$ , depends on the conformation of the client protein. Rapid off-rate and slow on rate is obviously counter-productive for obtaining large quantities of stable P–C complex. As discussed above, the off-rate is enhanced and the on-rate reduced in the presence of conditions that destabilize the chaperone (e.g. with denaturants).

One way of increasing the rate of P–C complex formation ( $P + C \rightarrow P-C$ ) is by covalently binding/fusing the client protein to the chaperone; this is best done by introducing a flexible linker between the sequences of P and C, such that the interactions between chaperone and client protein are minimally perturbed. In such fusion constructs, the off-rate is not necessarily altered as compared to a non-covalent protein complex. However, as the distance between C and P is short, the probability of encounter is enhanced. As an additional effect, the covalent fusion can hamper aggregation of P, even in its detached state, due to steric hindrance.

For completeness, Figure 2.1 (right) also sketches an additional reaction that can occur, particularly in the cell. The protein may become inserted into a complex with a cognate binding partner, or inserted into a membrane, and this reaction may strongly stabilize its folded state, and thus essentially remove it from

the equilibrium. For example, a membrane-protein precursor that gets inserted into the membrane is generally not bound by soluble chaperones any more. While this may be an important reaction in the cell, experimentally it is, of course, simply avoided by not having the respective binding partner or lipid bilayer present.

### 2.1.2 Overview of the possibilities to prepare chaperone-client complexes

Following these basic considerations, we can enumerate several strategies which, in principle, can be successful for generating P–C complexes. In the simplest case, mixing of a purified chaperone sample with a purified client protein may lead to spontaneous formation of a complex; this approach may need destabilizing mutations in the client protein, and/or covalently linking the client to its chaperone. Reducing conditions may also be considered as a similar strategy that impacts the client protein by disrupting disulfide bonds, but not the chaperone (unless it contains disulfide bonds). Alternatively, heating the mixed sample or applying pressure may allow the formation of complex, provided that the chaperone remains intact at the elevated temperature/pressure. Another possibility is to choose the solvent conditions to either destabilize the (soluble) protein, or to extract it from the aggregate state in which it may have been produced. The denatured protein can then be mixed with the chaperone, and simultaneous or subsequent removal of the denaturing conditions may allow the complex formation. This removal can be achieved in different ways (e.g. sudden dilution, dialysis; discussed below). Lastly, complex formation may be achieved directly in the host organism in which the protein is produced, and the intact chaperone—client complex may be extracted. Somewhat similarly, the client protein may be captured upon production, using a cell-free (*in vitro*) protein synthesis system.

Rationally designing the most appropriate complex-formation strategy may, in principle, work, based on the knowledge of the kinetics and thermodynamics of the reactions outlined in Figure 2.1. However, it is not often the case that these details are known, and trying a few different schemes is common. The following sections discuss these different strategies with reported examples from the literature.

## 2.2 Different complex-formation approaches in practice

### 2.2.1 Forming complexes in solution by mixing chaperones and their soluble client proteins

#### DnaK complexes

In some cases, the preparation of P–C complexes can be achieved in a straightforward manner by mixing chaperone and client. An example for such a case is the complex formed by DnaK (bacterial Hsp70) and the human telomere repeat binding factor 1 (hTRF1) (Sekhar et al., 2015). In solution, hTRF1 exists in an equilibrium of folded and unfolded states, but the folded state largely dominates; at 35 °C, ca. 4% are unfolded (Gianni et al., 2003) and at 25 °C only ca. 0.3% of the population is unfolded (Sekhar et al., 2015). There are two equilibria present in solution under the conditions chosen by the authors (25 °C, concentrations of ca. 0.5–0.6 mM): (i) folding/unfolding is heavily skewed towards the folded state with a population of 99.7%, and rate constants of ca.  $1\text{ s}^{-1}$  for the unfolding process and  $288\text{ s}^{-1}$  for the folding process; (ii) binding/release to the DnaK is with a dissociation constant in the low micromolar range ( $K_d \sim 18\ \mu\text{M}$ ,  $k_{\text{on}} = (1.1 \pm 0.2) \times 10^6\ \text{M}^{-1}\cdot\text{s}^{-1}$ ,  $k_{\text{off}} = 20.4 \pm 0.2\ \text{s}^{-1}$ ). The binding is proposed to proceed from the native unbound state (N) via the unfolded free state (U), to the bound state (Sekhar et al., 2015). In the complex hTRF1 has some residual secondary structure, and it seems plausible that DnaK can bind to hTRF1 when the latter has some residual helicity. In this particular case, the sample preparation involved mixing of hTRF1 with an

excess of DnaK (for the experiments in which hTRF1 was isotope-labeled for NMR) or DnaK with an excess of hTRF1 (for NMR studies of DnaK).

Sekhar, Kay and co-workers used similar sample-preparation strategies – namely mixing of chaperone and client – to resolve an important mechanistic question: does complex formation proceed via selection of the unfolded state that pre-exist in solution, or does the chaperone rather bind to the folded state, and then induce unfolding of the client in the bound state? The two schemes are sketched in Figure 2.2, following the general scheme of Figure 2.1. (Hereby, we considered that the aggregation-propensity can be ignored, i.e. the client is largely soluble.) Sekhar *et al.* used two different proteins, one with a native  $\alpha$ -helical fold and one with a native  $\beta$ -sheet fold, to resolve this question for both classes of proteins. They used marginally stable and slowly folding proteins, which allows dissecting the folding/unfolding equilibrium and the binding/release process. In an elegant combination of NMR dynamics experiment, such as ZZ-exchange and CEST, they demonstrated that conformational selection (left in Figure 2) is dominant (Sekhar *et al.*, 2018). Thus, the chaperone preferentially binds to pre-existing unfolded state(s).

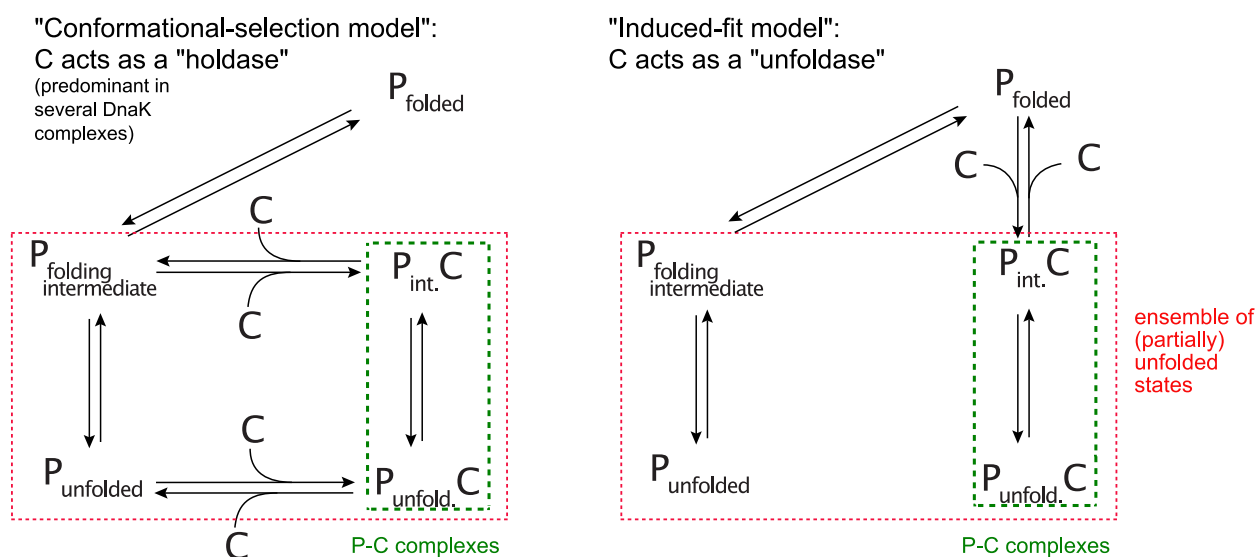


FIGURE 2.2: Schematic overview of the relevant reactions that can lead to the formation of a complex of a chaperone (C) with a soluble client protein (P) that has negligible aggregation propensity. P and C are mixed and the inherent property of P to expose (partially unfolded) parts, and the ability of C to recognize and bind these parts, leads to complex formation. For the case of DnaK, discussed in section 2.1., the left mechanism, namely selection of (partially or fully) unfolded states by the chaperone, is dominant (Sekhar *et al.*, 2018).

Similarly to working with full-length clients, another option is to use fragments thereof, which has been done by the same group (Rosenzweig *et al.*, 2017): using soluble peptides of the hTRF1 client and several NMR labeling schemes, multiple chaperone binding sites on the client were detected. Comparison of DnaK-bound state populations of these different binding sites (peptides) indicated that multiple DnaK chaperones bind to the hTRF1 client. Conceptually, what happens when working with fragments of clients is that the folding/unfolding equilibrium is heavily skewed towards the unfolded state – the fragment is often unable to fold – and hence the equilibrium NU is eliminated from the reaction scheme.

### Complexes of the chaperone Spy with Im7

The client-chaperone recognition and interaction of the stress-induced Spy foldase (Quan *et al.*, 2011) has been studied with the small soluble protein Im7. Im7 functions by binding to its cognate partner colicin E7, but in isolation Im7 tends to (partially) unfold, due to a number of locally frustrated sites. For this reason, Im7 has been used as a model protein in a number of protein-folding studies. Conditions and mutants of

Im7 have been reported in which Im7 remains soluble, either in partially folded or unfolded states (Capaldi et al., 2001; Friel et al., 2009; Whittaker et al., 2007). Complexes with Spy can be formed by mixing the protein with the chaperone (He et al., 2016b; Horowitz et al., 2016; Stull et al., 2016). Wild-type Im7 is in a partially unfolded state when it interacts with Spy, where the residues experiencing local frustrations are driving the sequence-unspecific binding to Spy (He et al., 2016b; He and Hiller, 2018). In addition to the wild-type Im7 client, site specific mutations were introduced that were shown to prevent Im7 folding (Pashley et al., 2012), while preserving its solubility during the production and purification. Mixing of this soluble but folding-incompetent Im7 variant with the chaperone was successful to prepare samples suitable for crystallography (Horowitz et al., 2016) and for NMR studies (He et al., 2016b). Although crystallography is challenging to interpret in this case of a highly disordered client (Wang et al., 2018), NMR showed that a combination of hydrophobic and charged residues appears to be required for binding to the unfolded client. The ITC and stopped-flow experiments showed that Spy binds to both folded and partially folded clients with similar affinities (3-10  $\mu\text{M}$ ) (Stull et al., 2016). In these examples, the concentration of the complex was high enough for NMR (several hundred  $\mu\text{M}$ ) and ITC studies.

In kinetic measurements, chemically denatured wild-type Im7 was used and its chaperone-bound folding was followed upon dilution into different concentrations of the Spy chaperone (Stull et al., 2016). In other assays, chemically denatured client malate-dehydrogenase (MDH) and DTT-reduced  $\alpha$ -lactalbumin ( $\alpha$ -LA) were used as fast-aggregating clients to probe the anti-aggregating activity of wild-type Spy (Stull et al., 2016) and its variants (He et al., 2020). The aggregation was followed by light scattering upon diluting denatured clients in the presence or absence of wild-type Spy or Spy variants.

### Hsp90 and Hsp110 complexes

The 90-kDa heat-shock protein (Hsp90) is another well-studied system of which complexes could be formed by mixing with clients. Street *et al.* probed binding and complex formation of the protein  $\Delta 131\Delta$  (131-residue fragment of staphylococcal nuclease; globally unfolded protein fragment) to Hsp90. Fluorescently labeled  $\Delta 131\Delta$  was mixed with Hsp90, and the binding was followed by the fluorescence polarization anisotropy titration experiments, with the increasing concentrations of the binding partners (up to 50  $\mu\text{M}$ ). Such straightforward titration experiments are possible due to the solubility of both the client and the chaperone. The dissociation constant was found to be in the low  $\mu\text{M}$  range (comparable to Spy-Im7 above), with a 1:1 stoichiometry; binding involves a partially structured region within the unfolded client (Street, Lavery, and Agard, 2011). Rüdiger and co-workers studied the complex formed by Hsp90 and the intrinsically disordered protein Tau (Karagöz et al., 2014) using NMR methods and electron paramagnetic resonance (EPR) (Weickert et al., 2020). Hsp90 and Tau were mixed at concentrations of tens to hundreds of micromolar, which is possible due to the relatively high solubility and comparably low aggregation-propensity of Tau (as compared to other client proteins described below).

A recent example investigated how the Hsp100 chaperone ClpB, a disaggregase, interacts with client protein. A 20-fold excess of  $\alpha$ -casein, an intrinsically disordered client protein, was mixed and incubated with an ATPase-deficient variant of ClpB (0.7 mg/mL) and structurally investigated by cryo-EM (Deville et al., 2017). Deville *et al.* prepared different mutants of ClpB at concentrations in the range 1.6 mg/mL to 8 mg/mL and incubated them with casein (in excess). By comparing the structural features of the mutant complexes with client protein, they proposed a mechanism of ATPase initiation and client binding by the disaggregase (Deville et al., 2019). Recent work by Yin *et al.* investigated the interactions of ClpB, DnaK and the client casein, i.e., a bi-chaperoning system, using cryo-EM (Yin et al., 2021). For preparing the complexes, a simple incubation of the components proved successful.



Taken together, for many soluble client proteins, generating complexes with chaperones is achieved by mixing solutions of the two binding partners. The possibility of obtaining complexes spontaneously provides opportunities for measuring binding affinity. The complex formation is facilitated by an inherent propensity of the client protein to unfold without aggregate formation (in the client's wild-type form or by mutations). At least in some documented cases (e.g. DnaK discussed above), the unfolded state is selected by the chaperone. If the binding affinity for the unfolded client is sufficiently high, then even a small proportion of unfolded client suffices to generate large quantities of P-C complex in the coupled set of reactions of unfolding and binding.

### 2.2.2 Making a client protein bind by adjusting sample conditions

#### Hsp60 (thermosome) complexes formed at high temperature

If the population of unfolded client is very low under native conditions, then it may be feasible to change the sample conditions to promote unfolding. For example, if the chaperone has a high tolerance to heat, then it may be straightforward to identify client proteins that become unfolded – and prone to chaperone-binding – at temperatures where the chaperone is still functional. This strategy has been chosen for making complexes of the Hsp60 chaperone from the hyperthermophilic archaeon *P. horikoshii* and two mesophilic client proteins (lysozyme and malate synthase G). *P. horikoshii*, an organism that is found in deep-ocean hydrothermal vents, has a growth optimum beyond 90 °C. Its Hsp60, also called thermosome, forms a ca. 1 MDa large assembly from 16 copies of the 60 kDa large subunits, which assemble into two large barrel-like chambers. Mas *et al.* have studied binding of lysozyme to thermosome (Mas *et al.*, 2018). In the absence of chaperone, heat unfolds lysozyme with a midpoint unfolding temperature of ca. 70 °C. Folding and unfolding occurs on a time scale of milliseconds in this temperature range. The unfolding is reversible: cooling the sample restored the folded state, without significant accumulation of aggregates. Mas *et al.* then mixed thermosome with lysozyme in an NMR tube at concentrations in the tens-of-micromolar range. At ambient temperature the interaction is negligible whereas at high temperature, lysozyme binds to the chaperone, evidenced by e.g. translational diffusion measurements that probe the size of the client protein. This finding shows that the chaperone binds preferentially (or exclusively) to the unfolded state. Interestingly, NMR showed, that the midpoint temperature of unfolding decreases by ca. 5 °C in presence of chaperone. This suggests that the presence of the chaperone shifts the equilibrium ( $P_{\text{folded}}/P_{\text{unfolded}}$ ) towards the unfolded state. The Hsp60-lysozyme<sub>unfolded</sub> interaction at high temperature is dynamic: NMR data show that binding and release are fast on the NMR time scale, i.e. in the sub-millisecond range. Cooling down the sample shifts the equilibrium of lysozyme to its folded state, which does not interact with the chaperone any more. Repeating the experiment at different ratios of lysozyme and thermosome (i.e., titration) allows measuring the proportion of lysozyme bound to thermosome, through NMR chemical shifts. This allowed the authors to determine the dissociation constant ( $K_d \sim 1.6 \mu\text{M}$ ); NMR measurements of isotope-labeled thermosome allowed mapping the interaction site. The short lifetime of this dynamic complex strongly contrasts with other cases discussed in sections 2.2.4 and 2.2.5, whose lifetimes are more than six orders of magnitude longer.

When applied to a different client protein, the 83-kDa large malate synthase G (MSG), the same strategy with the same chaperone resulted in a somewhat different behavior (Mas *et al.*, 2018). First, when MSG alone is heated, it aggregates at elevated temperature. When a mixture of MSG and *P. horikoshii* Hsp60 is heated, the chaperone protects MSG from aggregation (at least in part). The complex, once formed at high temperature can be brought to low temperature without dissociation, i.e. the complex formation is irreversible, or, more precisely, the off-rate of the client is very low. This might at least in part be explained

by the size of the client (83 kDa), which essentially fills the cavity of the chaperone cage. A further difference between these two cases is that MSG is much more aggregation-prone than lysozyme.

### Trigger factor complexes obtained via temperature or disulfide-bond breakage

Another example of applying heat for client-selective denaturing is the study of trigger factors (TF) holdase function where a temperature sensitive variant of maltose binding protein (MBP) showed reduced aggregation at high temperatures (50°C) in the presence of TF (Saio et al., 2014). The TF chaperone does not bind to the folded client (MBP variant at lower temperatures); however, by increasing the temperature partially unfolded states of the client protein become more populated, leading to complex formation. In the same study, with a different client protein, reducing conditions were used to keep the client unfolded and chaperone-binding prone: the client protein, wild-type PhoA protein, contains four cysteine residues and, *in vivo*, folds in oxidizing periplasmic environment. NMR was used to study the state of PhoA under reducing conditions. It was shown that several regions had only low (20-60%) secondary structure propensities (Saio et al., 2014). Isotopically (<sup>15</sup>N) labeled full-length PhoA, as well as the PhoA fragments, were titrated in reducing conditions with the unlabeled TF. This approach allowed identifying the binding site(s) of the chaperone (saio2014). These studies were successful because the TF chaperone is insensitive to the reducing agent because it lacks cysteine residues. This approach has a drawback, however, as the resulting sample may not be homogeneously reduced (see discussion in section 2.3). In addition to NMR experiments, isothermal titration calorimetry (ITC) with the fragments of the client protein and multiangle light scattering (MALS) experiments of the complex were performed to determine the stoichiometry of the complex and the binding affinity. The chaperone appeared monomeric upon interaction with the client and the dissociation constant was between 25 and 200 μM (Saio et al., 2014). The same two client proteins (PhoA and MBP) were used for characterization of client-protein binding by the chaperone SecB in similar conditions (reducing buffer or introducing mutations to make client soluble and un-foldable). The resulting complexes were studied by NMR, surface plasmon resonance (SPR) and bio-layer interferometry (BLI) (Huang et al., 2016).

Taken together, in cases where the client does not spontaneously bind (unlike those presented in section 2.2.1), the use of harsher conditions for the client protein (high temperature, reducing conditions). By choosing chaperones from hyperthermophilic organisms or chaperones that do not rely on disulfide bonds, such conditions may be found.

### 2.2.3 Mutating client proteins to make them chaperone-binding prone (or reduce their aggregation propensity)

As already mentioned above in some of the examples (Im7 in section 2.2.1 and MBP in section 2.2.2), site specific mutations are used to stabilize unfolded conformations of the client in order to increase the population of chaperone-binding competent states. The phosphatase A (PhoA) protein, encountered in section 2.2.2 contains native disulfides which are important for folding. The binding-competent unfolded state can be enhanced under reducing conditions (Huang et al., 2016; Saio et al., 2014), or by mutating the cysteines to, e.g. serines. The cysteine-less variant of full-length PhoA (proPhoAS4) shows no indication of residual structure based on its NMR spectrum, in contrast to the reduced form of a wild-type protein, for which some residual secondary structure propensity has been observed (Clerico et al., 2021). The reduction of the cysteines in a wild-type PhoA may not be complete, and it may change over time as the reducing agent gets consumed (oxidized). Consequently, a mixture of unfolded and partially folded species of the client protein are present in solution, and the unfolded portion of the client has a tendency decrease over time (Clerico et al., 2021). The cysteine-less variant has been used for studies of binding to the substrate-binding domain

(SBD) of DnaK (Clerico et al., 2021). The practical drawbacks of a cysteine-less variant, however, can be that (i) its production requires solubilization and purification from inclusion bodies and (ii) it may allow only for short time storing at low concentrations, because aggregation from the unfolded state is more rapid than if the protein was folded during storage (Clerico et al., 2021).

Mutating cysteines may also be of use in a different context: the formation of intermolecular disulfide bonds may enhance aggregation of client proteins. For forming complexes of membrane proteins, i.e. highly aggregation-prone polypeptides, Weinhäupl *et al.* developed a pull-down approach with the chaperone TIM9-10 described in section 2.2.5. The chaperone contains essential disulfide bonds while the two client proteins chosen by the authors (ADP/ATP carrier; GTP/GDP carrier) contain four and one cysteines, respectively (Weinhäupl et al., 2018). The authors found that the presence of the native cysteines enhances the propensity of the client protein to aggregate. Employing reducing conditions is not an option because the chaperone unfolds when its disulfides are broken. Hence, in this case, mutating the native cysteines in the client proteins to Ala or Ser was a successful solution for complex formation without accumulating aggregates (see section 2.2.5).

## 2.2.4 Complex formation upon removal of a denaturant

Some client proteins may not be kept in aqueous solution at all, i.e. the aggregation reaction may be predominant. For such proteins one needs to find conditions in which they can encounter (in solution) the chaperone proteins; of course, these conditions shall not be denaturing, so as to maintain the chaperone in a folded and active state. Membrane-protein (MP) precursors on the way to their membranes (e.g. bacterial outer-membrane proteins or MPs of organelles) are generally such highly-insoluble polypeptides. Bacterial outer membrane proteins (OMPs) with a native  $\beta$ -barrel fold require purification in chemically denaturing conditions from inclusion bodies. Chemical denaturation at high concentrations of chaotropic agents such as 8 M urea or 6 M guanidine-hydrochloride implies denaturing of all proteins present in solution and shielding of all (hydrophobic) intra- and intermolecular interactions. Consequently, the denaturing agent needs to be removed for client-chaperone complex formation. It has been shown that by drop-wise dilution (drip dilution) of denatured Omp into a solution containing a native chaperone of OMPs, Skp, leads to successful complex formation (Burmam, Wang, and Hiller, 2013). Drip dilution is a very fast dilution of the denaturant by dropwise addition of denatured sample into a typically 100 times bigger volume of the refolding buffer over a period of 1 h, keeping unfolded protein at minimal concentration during this procedure (Burgess, 2009). In the denaturing buffer, the OMP is unfolded, but as soon as the OMP is rapidly dissolved in chaperone-containing buffer at very low denaturant concentration, the OMP may bind to the chaperone, but it may also aggregate. (The folding reaction is negligible, as membrane proteins generally do not fold to a defined structure outside a membrane.) The relative rate constants of aggregation vs. binding are, thus, determining whether a complex forms. The aggregation rate constant is temperature-dependent, and aggregation is slowed at low temperature (Xie and Wetlaufer, 1996). The aggregation furthermore depends on the concentration of the client protein; diluting into sufficiently large volume and good mixing during dilution is, therefore, helpful to suppress aggregation. The rate constant of P-C complex formation is dependent on the concentrations of both the client protein and the chaperone. The drip dilution method implies high dilution of the client protein concentration too, causing additional decrease of aggregation rate due to reduced diffusion-collision probability of the unfolded client molecules. However, to ensure that most of the chaperone is in the client bound state, excess of denatured client needs to be added in a drop-wise manner until precipitation due to lack of free chaperone is observed (Burmam, Wang, and Hiller, 2013). In the preparation of Skp-OMP complexes, the Hiller group has added an excess of OMP (solubilized in 6 M guanidine) in a drop-wise manner to Skp in native buffer (such as 25 mM MES, 150 mM NaCl, pH

6.5), under stirring, until precipitation appeared (Burmam, Wang, and Hiller, 2013; Callon, Burmann, and Hiller, 2014). The solution was then centrifuged before analysis by structural and biophysical methods. In some special cases, the assembly buffer in which client is drip-diluted can contain reducing agent such as DTT (see section 2.2.2 for keeping client binding-prone in chaperone compatible conditions) if working with disulphide bond-free chaperone. Once the chaperone-client complex is formed, the assembly buffer can be exchanged for the analysis compatible buffer without chemical reductant (Thoma et al., 2015).

Essentially the same method was applied for forming complexes of trigger factor (TF) with OmpA. Here, the purification of the client protein was performed in denaturing conditions and complexes were formed by diluting the client sample 20-fold in the presence of chaperone in reducing conditions to keep the OmpA unfolded and TF-binding prone (Saio et al., 2014).

It is noteworthy that the kinetics and thermodynamics of complex formation between chaperones and highly aggregation-prone polypeptides are very different from those of the soluble clients (described in sections 2.2.1, 2.2.4, 2.2.3). For Skp complexes with several  $\beta$ -barrel client proteins of the OMP family, the dissociation constant was consistently in the one- to two-digit nanomolar range (Qu et al., 2007). The life time of the Skp-OmpA complex is ca. 2.6 hours (Burmam, Wang, and Hiller, 2013), which is more than six orders of magnitude longer than complexes formed by e.g. thermosome and lysozyme, and more than three orders of magnitude longer than the DnaK complexes (section 2.2.1).

In a similar way, a GroEL/ES-client complex was obtained for cryo-EM studies capturing chaperone 'in action' with an encapsulated client. The clients used in this study were denatured in acidic urea buffer and mixed with an open-state GroEL/ES chaperone solution in reducing conditions, by drip-diluting the denaturant 50-fold, after which ATP was added to ensure the encapsulation of the client (D. H. (Chen et al., 2013).

Complex formation by the drip dilution method requires that the binding is faster than aggregation and that the complexes are stable. The method works for forming complexes of bacterial outer-membrane proteins with different chaperones (Skp, SurA, trigger factor). We found that in another case of membrane-protein precursors as client proteins, the drip dilution method fails: the case of mitochondrial-carrier membrane proteins and their native chaperones, the hetero-hexameric TIM9·10 chaperone. Drip dilution of different guanidine-solubilized mitochondrial carriers into a solution of TIM9·10 (testing a range of optimization parameters), has not produced complex; rather, the client protein precipitated (not published). A reason for this failure may be that mitochondrial carrier proteins ( $\alpha$ -helical in their native state) are more hydrophobic than OMPs. This higher hydrophobicity may lead to faster aggregation compared to OMPs. Additionally, TIM chaperones disassemble above approx. 100 mM guanidine-hydrochloride (unpublished data). The guanidine-sensitivity of the chaperone might result in destabilization of the chaperone at the locally high denaturant concentration as the drop gets diluted into chaperone-containing buffer.

In a Förster resonance energy transfer (FRET) study, DnaK, introduced in section 2.2.1, has been studied in complex with rhodanese (Kellner et al., 2014). Of interest is that the rhodanese–DnaK complex is strictly dependent on the presence of the J-domain protein DnaJ (which is essential for ATP hydrolysis of Hsp70 chaperones) and ATP. Kellner *et al.* prepared samples from rhodanese that was denatured in 4 M guanidinium buffer, and then diluted the protein into a buffer without denaturant, but with DnaK, DnaJ, GrpE (a nucleotide exchange factor) and ATP. The particularity of single-molecule FRET is its ability to work at very low concentration (25 to 75 pM in this study).

## Complex formation by dialysis

Another method towards formation of complexes from a denaturant solution is dialysis. In such a procedure, one would mix the client protein and the chaperone in the presence of denaturant, and dialyze out the denaturant. At the initially high denaturant concentration, not only the client protein is solubilized but, most often, the chaperone is also perturbed; moreover, the shielding by the chaotropic agent likely hampers the interaction between client and chaperone. Dialysis may, thus, work if during the time-dependent reduction of denaturant concentration, the chaperone's integrity and binding capacity are restored before the aggregation of the client protein becomes dominant. This approach has been reported to be successful e.g., for the DegQ chaperone/protease, (Malet et al., 2012; Mauldin and Sauer, 2013). In the case of the above-described TIM9·10 chaperone with mitochondrial carriers, dialysis approaches proved unsuccessful: after dialysis, we could only obtain aggregated client protein and free chaperone (unpublished). It is likely that this is due to the fact that the hetero-hexameric chaperone is disassembled at a concentration of ca. 100 mM guanidine, likely much less than what is needed to keep mitochondrial carriers in solution.

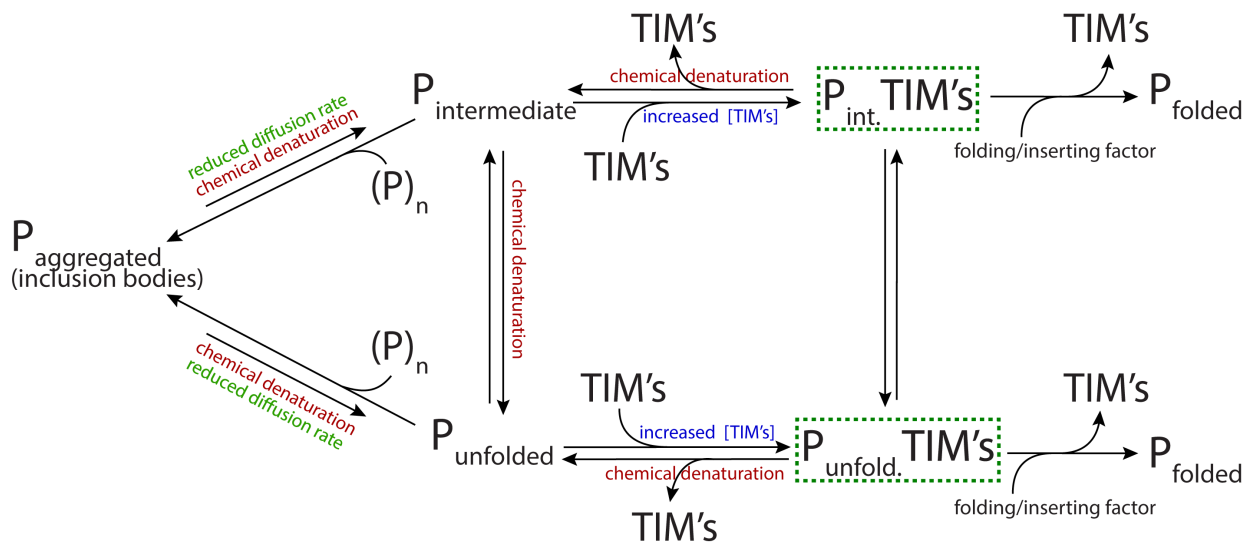
### 2.2.5 Hampering aggregation by client immobilization: complex formation with a pull-down approach

As outlined above, drip dilution and dialysis were unsuccessful for generating complexes of TIM9·10 and mitochondrial carriers (such as the ADP/ATP carrier Aac1 and the GDP/GTP carrier Ggc1). Weinhäupl *et al.* have developed an approach for complex formation that strongly reduces aggregation and allows obtaining homogeneous and long-lived complexes, stable over days to weeks (Weinhäupl et al., 2018; Sučec et al., 2020) (Figure 2.3). The method comprises first the binding of unfolded client protein in denaturant (such as 6 M guanidine) on an affinity column, via an affinity tag on the client-protein construct (His<sub>6</sub>-tag). As the individual molecules are attached to the beads at fixed location, they are unable to diffuse and encounter each other and thus to form multimeric aggregates. Then, the buffer is exchanged to a more native buffer without denaturant ("binding buffer"). The concentration of denatured client loaded on the column is kept low: on a 5 ml affinity resin approximately 2.5 mg of unfolded client is loaded. This helps to reduce the aggregation rate. Immediately after removal of denaturant, an excess of chaperone, typically 2 times higher compared to the client protein, is passed through the column. The non-bound chaperone is washed off the column, and the complex then eluted (e.g. with imidazole). In a slightly modified version of this protocol, the denaturant was not removed before addition of the chaperone. Instead, a 2-fold excess of chaperone (compared to loaded client protein) was added in one fifth of the column volume and the denaturant concentration gradually decreased by diluting the flow-through solution 1:1 with binding buffer before passing it onto the column again. This process was repeated until denaturant concentration was below 0.05 M (roughly in a total volume of 10 times the column volume). In this latter approach, the chaperone is presumably disassembled in the beginning and when the denaturant concentration is below a critical level (for TIM9·10 this is of the order of 100 mM), the chaperone assembles and is able to bind the client protein.

For TIM chaperones and mitochondrial carriers, both variants work, and the former (namely removing the denaturant from the column before adding the chaperone) is experimentally somewhat simpler. After desalting and concentrating the eluted complex, the final concentration of chaperone complex was of the order of 150  $\mu$ M, well suited for structural studies. The sample was stable for days of measurement at up to 35 °C. The lifetime of the complex was ca. 4 hours, i.e. when apo-chaperone was added to P-C complex, it took 4 hours to transfer the client from one chaperone to another. No significant aggregation of the client was observed under these conditions. Interestingly, however, gel filtration experiments of the complex were unsuccessful: the TIM9·10 complex with mitochondrial carriers dissociates during passage on a size

exclusion column (unpublished data). In contrast, with another client protein, Tim23, the P–C complexes of both TIM9·10 and a homologous chaperone, TIM8·13, could be analyzed by gel filtration and coupled SEC-MALS (see supplemental figures S12 and S13 in (Sučec et al., 2020)). Characterization of these complexes has also been done by other biophysical/biochemical methods, such as analytical ultra-centrifugation, small-angle X-ray scattering, NMR-detected translational diffusion experiments or SDS-PAGE (Weinhäupl et al., 2018; Sučec et al., 2020). Note that with the drip-dilution method one can determine dissociation constants  $K_d$ , by following e.g. fluorescence spectra upon titrating the client protein into a chaperone solution (Qu et al., 2007). In the pull-down method describe here, determining affinities or stoichiometries is not possible.

It is worth noting that this approach to complex formation was not successful for the TIM chaperones with a precursor of the outer membrane  $\beta$ -barrels. None of the above methods (drip dilution, dialysis, pull-down) resulted in complex formation. The  $\beta$ -barrel client protein was found to aggregate in all these conditions (reported in (Weinhäupl et al., 2018)). The authors were able to generate complexes of weak affinity ( $K_d$  in the low mM range) only with a  $\beta$ -turn fragment of the  $\beta$ -barrel client, but only if the  $\beta$ -turn was stabilized by cyclization of the peptide. A likely explanation for this observation is that in a cyclic  $\beta$ -turn element the side chains that are hydrophobic and hydrophilic point to the two faces of the  $\beta$ -turn element. In other words, such an element creates one hydrophilic and one hydrophobic face – and the latter presumably binds. In a totally disordered polypeptide, the hydrophilic/hydrophobic side chains point randomly in different directions, which likely impedes the chaperone binding.



**FIGURE 2.3: Scheme of different approaches for the mitochondrial intermembrane space chaperone complex formation.** Method developed for complex formation with mitochondrial membrane proteins (Weinhäupl et al., 2018) is comprised of three steps. In the first step (red) hydrophobic client protein is unfolded from the inclusion bodies using strong chemical denaturant that prevents formation of protein aggregates and keeps the client in an unfolded state. However, strong denaturant is preventing any possible interaction with the client due to chaperones instability under such conditions. In the second step (green) client is bound to the affinity column, which reduces the local concentration of the client and its diffusion rate and the rate of clients aggregation is reduced once the denaturant is removed. In the third step denaturant is removed and saturating concentration of the chaperone is added (blue) to the bound client increasing the rate of complex formation. In a second method, the client is produced by a cell-free (in vitro) expression system which contains the chaperone, in addition to the usual components required for in vitro translation. In a second method, complex formation in a cell-free system, client is produced in in vitro expression system while providing increased concentration of purified chaperone. In this approach there are no aggregation-prone client species (i.e. unfolded and free client) as long as there are free chaperones in the cell-free reaction solution.

### 2.2.6 Purifying P–C complexes from the cell

Direct purification of the P–C complex from the host organism would be biologically most relevant complex for biophysical studies and characterization. In order for this approach to be successful, the stability of the P–C complex must be sufficiently high such that the complex survives all steps from cell disruption to purification. This approach was shown to be successful in the case of the Hsp60 chaperone GroEL/ES (Fei et al., 2014) and the bacterial periplasmic chaperone/protease DegP with outer-membrane proteins (Krojer et al., 2008). The majority of DegP purified from *E. coli* is co-purified as 12-mer and 24-mer with bound OMPs (OmpA, OmpC, OmpF and LamB). These complexes could be separated from the apo-chaperone by ion exchange chromatography. The yield of purified complex was sufficient for electron microscopy structural studies to determine the active chaperone state (Krojer et al., 2008). However, this approach is often discarded for client-chaperone sample preparation due to the heterogeneity of the sample, where chaperone with different client proteins can be present. In a study by Fei et al. a subsequent refolding of GroEL was applied after chaperones purification to reduce unwanted client-bound particles Fei et al., 2014. However, co-purification of chaperone-client particles may in fact be highly desirable, in particular if coupled to cryo-EM, where the separation into free and client-loaded classes can be done post acquisition.

Co-expression of the client and substrate protein from the same expression plasmid in the bacterial cells was applied to obtain the trigger factor in a complex with the S7 (ribosomal subunit protein) and OmpA client proteins for crystallography studies (Martinez-Hackert and Hendrickson, 2009). This was done after identifying the TF-substrate proteome in *E. coli* by co-purification with TF and liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. The complexes of trigger factor formed during co-expression could be co-purified on the affinity column with only one binding partner carrying the affinity tag, and its stability is sufficient for characterization by the size-exclusion chromatography and analytical ultracentrifugation with a yield high enough for crystallographic studies (Martinez-Hackert and Hendrickson, 2009).

Baculovirus co-expression of human Hsp90 complexes in the insect cells was used for production of chaperone-cochaperone-client complex for (cryo-) electro-microscopy studies. With addition of a mimetic of the ATP  $\gamma$ -phosphate (molybdate) during the complex co-purification, the closed state of an Hsp90-Cdc37-Cdk4 ternary complex was captured with the client in the semi-folded state (Vaughan et al., 2006; Verba et al., 2016). Interestingly, Verba et al. also reported that simple mixing of proteins individually produced and purified did not result in successful complex formation. This may be due to modifications that occur after complex formation or additional factors required for the complex formation (Verba et al., 2016).

### 2.2.7 Capturing emerging client proteins in a cell-free system

Another method of chaperone-client complex formation, with a client being highly aggregation prone, is during *in vitro* protein synthesis (cell-free expression) of a client Schneider et al., 2010, in the presence of purified chaperone Rampelt et al., 2020; Weinhäupl et al., 2018. *In vitro* protein synthesis using bacterial cell extract in the presence of recombinantly produced and purified TIM9-10 chaperone significantly improved the solubility of mitochondrial carrier precursors (Mpc3, Mpc2, Ggc1) in a chaperone concentration-dependent manner. The majority of the hydrophobic carrier precursor aggregate during the cell-free expression in the absence of detergent or chaperone. From these experiments it was concluded that the mitochondrial pyruvate carriers (Mpc) depend on the TIM9-10 chaperone during their import. In addition, binding to the previously established hydrophobic binding site on the chaperone could be confirmed (Rampelt et al., 2020). High local concentration of chaperone in the cell-free reaction mixture is enabling clients capture on its exit of the ribosome, reducing the probability of client aggregation.

The protein yield in cell-free expression systems is often of the order of 0.5 mg per milliliter of reaction mixture, and can reach up to 5 mg per ml. Constant improvements in the methodology enable the production of 'difficult-to-obtain' proteins (some eukaryotic, post-translationally modified, membrane or 'toxic-for-the-cell' proteins; (Jin and Hong, 2018)), as well as specific labeling schemes for NMR studies, including deuteration (Imbert et al., 2021). In parallel with the reduction of costs, this method is becoming used more often to obtain samples for structural biology studies (Terada et al., 2016). However, it often requires extensive optimization of the reaction conditions that is protein dependent, time consuming and often expensive.

### 2.2.8 Fusing chaperone and client into a single polypeptide chain

An additional approach to generate complexes is to fuse a client protein (or a fragment) to a chaperone, i.e. to generate a single polypeptide containing both proteins, separated by a suitable linker. Conceptually, this approach shifts the equilibrium (P / P-C) towards the bound form because of the proximity; furthermore, the presence of the chaperone may hamper the aggregation reaction, and may also hamper folding of the client protein. This approach was used, for example, for stabilizing a predominantly unfolded variant of maltose binding protein (MBP) to trigger factor separated by a 25-residue linker for a binding site(s) verification (Saio et al., 2014).

### 2.2.9 Chaperones bound to protein aggregates

A particularly interesting process is how chaperones bind to pre-existing protein aggregates. Chaperones have been shown to play key roles in suppressing protein aggregation by interacting with aggregation-prone proteins at different stages, including in oligomeric and fibrillized forms. In one recent example, the interaction of J-domain proteins (Hsp40) with amyloid fibrils formed by the protein Tau has been monitored. From a preparation standpoint such samples are, arguably, simpler than the above examples that aim to generate a soluble sample: the pre-formed amyloid fibers were incubated with the chaperone and monitored by negative-stain EM (Irwin et al., 2021).



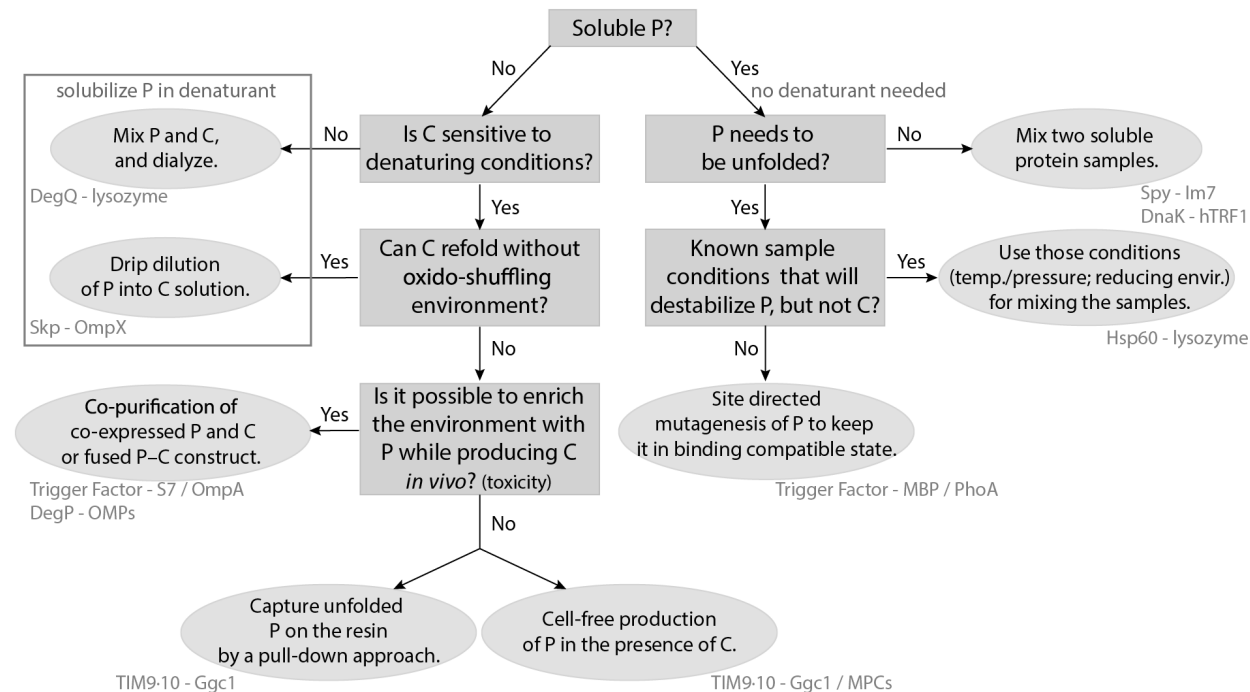


FIGURE 2.4: Overview of the main methods used for preparing chaperone-client complexes. Examples for these different cases are given by the name of the involved proteins. See main text for details.

### 2.3 Concluding remarks

We have outlined here a number of strategies that can be successful for preparing chaperone complexes with client proteins, with the aim to characterize these complexes structurally and functionally. These strategies are summarized in the scheme of Figure 2.4. The properties of the various client proteins and chaperones vastly differ, and so do the preparation schemes, affinities and complex life times. In our view, it is not possible to make precise predictions which of the protocols will be successful, but general trends nonetheless emerge. If the client protein is largely soluble, then not aggregation of P, but rather the stability of the folded state of P may hamper successful P–C complex formation. In such cases, destabilizing P by mutations or appropriate sample conditions (denaturing, heat, possibly also pressure) is a promising route, provided the chaperone resists. The studied cases suggest that the binding and release reaction in such complexes are fast, often on millisecond time scales.

If the client protein is highly aggregation-prone, then protocols that start with denatured state are the way to go. Although the number of reported cases is small, we propose that the most aggregation-prone client proteins may not be amenable to the drip-dilution approach, and they may rather need e.g. a pull-down approach (Weinhäupl et al., 2018) or a cell-free production or isolation from the native environment.

The steady improvement of methods for the structural and dynamical study of chaperone complexes, including cryo-EM, solution- and solid-state NMR and optical spectroscopies such as single-molecule FRET comes jointly with a better understanding of the sample preparation protocols. Biochemistry, biophysics and structural-biology tools as well as *in vivo* experiments have allowed understanding the fascinating mechanisms of chaperone function at the atomic level, which is not only of fundamental interest but also of biomedical importance.

## Chapter 3

# Open questions addressed in this thesis: molecular mechanisms of chaperones and receptor domains

More than three decades of intense research on the mechanisms of mitochondrial import have provided a wealth of data and (at least) low-resolution views of the series of events taking place when a protein gets imported into mitochondria and sorted into the right compartment. We now know the main pathways of import (and it seems plausible that no entirely new pathway will be identified), sketched in Fig. 1.1. The key players involved in the respective import machineries have been identified, and for many of them there are even NMR structures or crystal structures (e.g. Tom20, Tom70). The cryo-EM structures of the complexes have been obtained during the period of my PhD thesis, such as for the TOM core complex (Fig. 1.2) or the SAM insertase machinery, in its apo state (Takeda et al., 2020) or in complex with the (already membrane-inserted) Tom40 pore (Wang et al., 2021), and have provided very useful insights, compared to the status of the field at the onset of my thesis.

While all these results, and particularly the most recent cryo-EM data, are phenomenal and have provided much mechanistic insights, some important mechanistic questions remain. For the case of the TOM translocase, the events occurring at the reception of precursor proteins are still not clear at the atomic level. Tom20 and Tom70 are absent in the cryo-EM structures of the TOMcc complex, and also the cytosolic part of Tom22 is unobserved in these structures, presumably due to large-amplitude flexibility of this part.

Moreover, most of the available structures are without the client proteins. It is, thus, not clear how the client proteins interact with the respective machineries, i.e. chaperones, receptor domains and translocases. It is still not clear how, after the initial recognition, the receptors handle the precursor protein and direct it towards the import pore. Is there receptor–receptor interaction that in the presence of the client anchors the receptor to the TOMcc, is the receptor–receptor interaction out-competing the precursor, enabling its entry to the pore? What are the molecular drivers of these (transient) binding events? How do chaperones fulfill their requirements of promiscuity (handling many different proteins) and some degree of specificity?

In the main part of my thesis, I have focused on understanding how the receptor domains interact with each other, and with precursor proteins and (co-)chaperones. In a second part, I have studied how the small TIM chaperones interact with different preproteins, and what confers binding specificity. Lastly, I have contributed to identifying how a rather recently discovered class of mitochondrial inner-membrane proteins, the mitochondrial pyruvate carriers, are imported.

The results of this manuscript are divided in three parts.

**1<sup>st</sup> Part.** Chapter 4 investigates the structural basis of receptor interactions and precursor protein handling at the mitochondrial import gate. This part is being prepared for publication.

**2<sup>nd</sup> Part.** Chapter 5 investigates the substrate specificity of the small TIM chaperones.

Published in:

I. Sućec, Y. Wang, O. Dakhlaoui, K. Weinhäupl, T. Jores, D. Costa, A. Hessel, M. Brennich, D. Rapaport, K. Lindorff-Larsen, B. Bersch, P. Schanda, **Structural basis of client specificity in mitochondrial membrane-protein chaperones**, *Sci. Adv.* (2020) 6: eabd0263

**3<sup>rd</sup> Part.** Chapter 6 investigates new protein clients of TIM9.10 dependant carrier import pathway.

The results presented in this chapter is work that I performed as a part of the collaborative project with the group of Professor Nikolaus Pfanner (Rampelt et al., 2020). The results confirm requirement of the TIM9.10 chaperone complex in MPC proteins import pathway, and the role of the conserved hydrophobic motifs of the chaperone.

Published in (full publication available in AppendixB.1):

H. Rampelt, I. Sućec, B. Bersch, P. Horten, I. Perschil, J.-C. Martinou, M. van der Laan, N. Wiedemann, P. Schanda and N. Pfanner (2020) **The mitochondrial carrier pathway transports non-canonical substrates with an odd number of transmembrane segments**. *BMC Biology* 18.1, pp. 1–15

In the appendix B.1 of this manuscript an additional publication is included, where my contribution was in performing and analysing the peptidase cleavage assays.

## Chapter 4

# Results 1: receptor interactions and precursor protein handling at the mitochondrial import gate

The outer membrane translocase (TOM) is the import channel for most of the nuclear-encoded mitochondrial proteins. The core of the import pore contains the pore-forming protein Tom40, the essential receptor protein Tom22, and three small proteins, Tom5, Tom6 and Tom7. Precursor proteins are thought to be initially bound by the peripheral receptor proteins Tom20 and Tom70 before being imported by the TOM complex. It is deeply embedded in the field that the Tom22 receptor is taking over the substrates from the peripheral receptors and handing them over to the Tom40 pore (Kiebler et al., 1993; Bolliger et al., 1995; Brix, Dietmeier, and Pfanner, 1997; Hönlinger et al., 1995; Brix et al., 1999). Additionally, from the *in vivo* and *in organello* import studies it is often concluded that the three receptors have partially overlapping functions (Koh, Hájek, and Bedwell, 2001; Kondo-Okamoto, Shaw, and Okamoto, 2008). However, the molecular mechanism of the initial recognition and import steps at the outer mitochondrial membrane are poorly understood. The difficulties in determining the initial sequence of the recognition and import events by *in vivo* studies come from different reasons: (i) the transmembrane helix of Tom22 is essential for assembly of the TOM core complex and hence for the yeast cell growth, (ii) genetic deletion of Tom20 causes decrease in the amount of Tom22. Even more fundamentally, the structures of the cytosolic domains of the Tom22 and the yeast Tom20 receptors are not known. The recent cryo-EM structures of the TOM core complex from *Saccharomyces cerevisiae* (Tucker, 2019; Araiso et al., 2019) enabled a near-atomic view of this complex, but lacks the entire Tom20 and Tom70 proteins and the cytosolic domain of Tom22 receptor.

Here we investigated the structure and the interaction patterns of the Tom22 receptor in the first steps of mitochondrial precursor import. According to the proposed roles of Tom22 in presequence binding and/or recruiting Tom20 and Tom70 to the TOM complex, we studied presequence binding to Tom22 but also its interaction with Tom20 or Tom70. In addition, we also tried to shed light on the interaction of import signal sequences with the two receptor proteins of the Tom complex, Tom20 and Tom70. Our results propose the binding of the small cytosolic helix of Tom22 receptor to the substrate binding site of both Tom20 and Tom70 receptors. The mechanism where Tom22 is replacing the substrate in the receptors binding site could explain the observations reported so far in the literature, mainly that the deletion of either Tom22 or Tom20 cytosolic domains leads to similar effect of presequence import impairment (Yamano et al., 2008b).

We propose a molecular model for Tom22 interacting with the other two receptors of the TOM machinery, where Tom22 has a dual role: (i) docking the precursor-loaded Tom20/Tom70 (previously suggested

Yamano et al., 2008b) and (ii) helping in releasing the precursor towards the Tom40 pore by replacing the precursor in the Tom20/Tom70 binding sites.

## 4.1 Conformation and dynamics of Tom22

### 4.1.1 The cytosolic Tom22 domain is an intrinsically disordered protein with a transient helix

To determine the structure and interaction patterns of the cytosolic domain of the Tom22 receptor (Tom22<sub>cyt</sub>) we mainly relayed on studies by NMR spectroscopy. The two protein constructs we used comprised either the first 97 residues (Tom22(1-97)) or residues 1-74 (Tom22(1-74)). Tom22(1-97) was chosen since it represents the full cytosolic domain of the receptor; predictions (done before the cryo-EM structures were available), indicated that the TM helix starts roughly at residue 98. The shorter Tom22(1-74) construct was obtained since it comprises the predicted fully disordered N-terminus and stops after the conserved cytosolic region (residues 56-69) (see Figure 4.2). Both constructs were fused to a His<sub>6</sub>-tagged GB1 domain which helped in the production and purification of these constructs, and a TEV protease cleavage site between Tom22 and GB1 allowed obtaining the respective Tom22 construct only (see Methods7.4.2). I used both constructs for biochemical and NMR experiments. The longer Tom22(1-97) construct was obtained with a much lower yield (ca. 4 mg/L of culture) compared to the shorter Tom22(1-74) (about four times higher yield) which is why in most of the studies we used the shorter protein (see Table7.1). The solubility tag was kept in some of the cases since the cleaved Tom22 has no aromatic amino acids to facilitate measurements of protein sample concentration. When the tag was cleaved, the concentration was measured using the BCA protein assay (see Methods7.4.2).

The fingerprint spectrum of Tom22 <sup>1</sup>H<sup>15</sup>N BEST-TROSY shows the hallmark features of a disordered protein, with low dispersion of the resonances (peak positions) in the <sup>1</sup>H dimension, characteristic for an IDP. Backbone chemical shift assignment of cytosolic domain of Tom22(1-74) (BMRB ID: 51381; Figure 4.1A.) and Tom22(1-97) was performed using three-dimensional NMR experiments in which the resonances of an amide group (NH, H) of each amino acid are correlated via scalar couplings with the carbon resonances (C $\alpha$ , C $\beta$ , CO) of the same or the preceding residue or to the amide group (NH, H) of preceding or following residue (see Methods7.10.1). 83% of the backbone resonances of Tom22(1-74) sequence was successfully assigned (77% of peaks in the BEST-TROSY spectrum). Two regions (residues 15-19) and the C-terminus (residues 70-75) are left unassigned, due to difficulties in finding corresponding resonances and to the presence of proline residues.

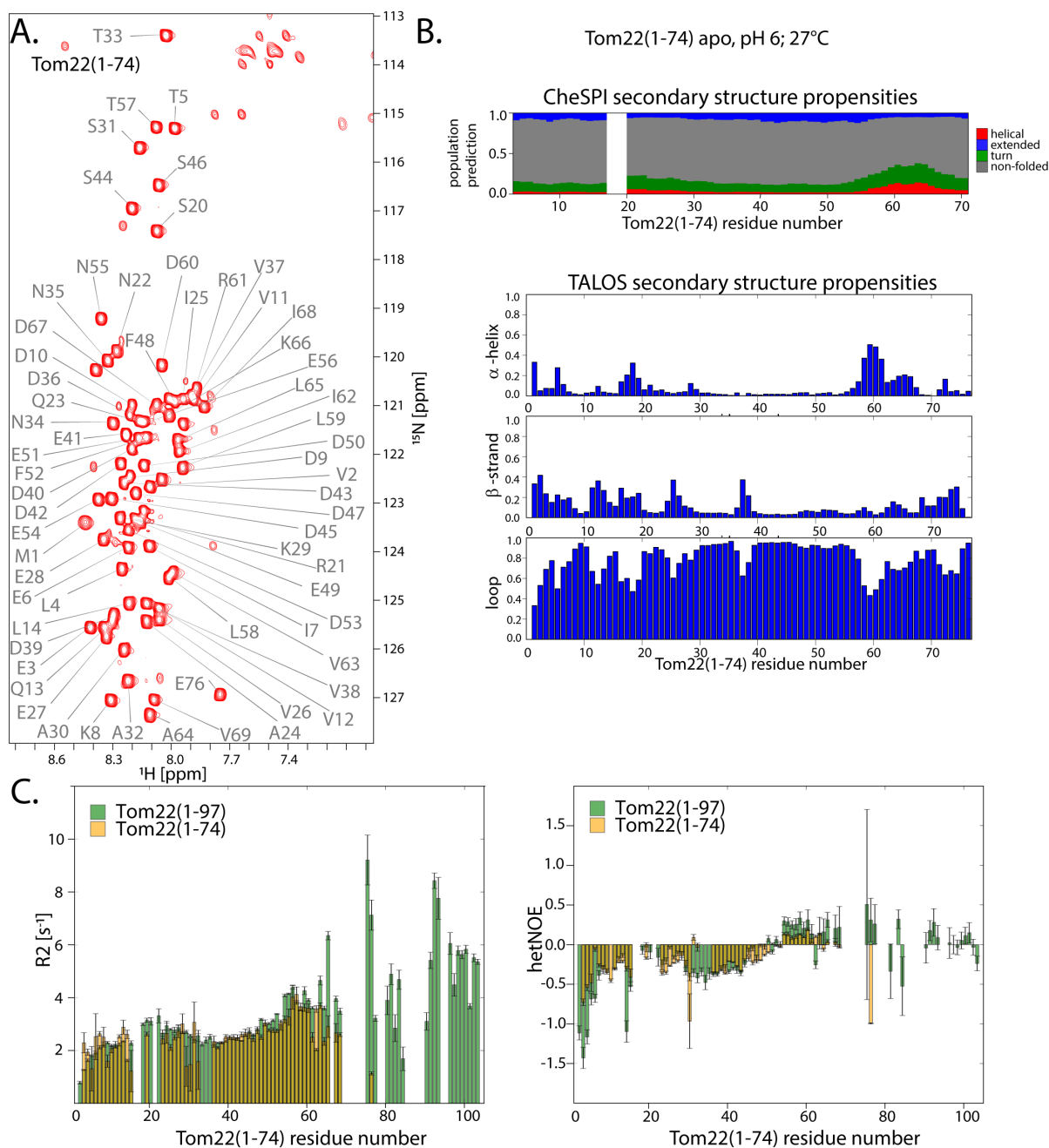


FIGURE 4.1: **Characterization of the intrinsically disordered domain of Tom22.** **A.**  $^{15}\text{N}$  $^1\text{H}$  spectrum and shift assignment of Tom22(1-74). **B.** Residue-wise secondary structure propensities for Tom22(1-74) obtained by CheSPI software (upper plot) and by TALOS-N software (lower plot) showing propensities for unstructured (loop),  $\beta$ -strand or  $\alpha$ -helical conformation. **C.** Relaxation properties of Tom22(1-74) in yellow and Tom22(1-97) in green. In left panel,  $R_2$  relaxation rate constants for two constructs are compared, showing similar relaxation properties, the faster relaxation of the flexible and unfolded N-terminus. In the panel on the right, the hetNOE values are plotted as a function of Tom22 residue number. Negative hetNOE values were observed for the N-terminus indicating highly flexible region. The increase in the hetNOE values after residue 55 indicate less psec time scale motion.

We assessed the secondary structure propensity and local backbone dynamics of the Tom22<sub>cyt</sub> using chemical shifts and  $^{15}\text{N}$ -relaxation data. Chemical shifts report on the time-averaged local backbone dihedral angles.  $\alpha$ -Helical and  $\beta$ -strand conformations show characteristic chemical shift deviations from values

expected for random coil proteins. Dynamic disorder of the protein, which averages all possible conformations, results in population and time averaged chemical-shifts. Analysis of the chemical shift and comparison to databases reporting chemical shifts for known protein structures (empirical and possibly assisted by neural networks), thus, allows assessing the secondary-structure content and order parameters for the protein backbone. The assigned Tom22 chemical shifts were used as input to two different approaches, the CheSPI software (Nielsen and Mulder, 2021), developed for IDPs, and the TALOS-N software (Shen and Bax, 2013).  $^{15}\text{N}$  heteronuclear Overhauser effect (hetNOE) and  $^{15}\text{N}$   $R_2$  relaxation experiments (see Figure 4.1B and C) provide information on protein backbone dynamics.

The TALOS-N software (Torsion Angle Likelihood Obtained from Shifts; Shen and Bax, 2013) predicts the protein backbone and sidechain torsion angles from the user provided chemical shifts. TALOS-N relies on neural network analysis of the input chemical shift data and uses the database (RefDB) of assigned chemical shifts of soluble proteins with known structure by comparing the small fragments of assigned peptides to the sequence of a target protein. Information about the secondary structure of a protein is possible to obtain because chemical shifts of amide proton, amide nitrogen,  $C\alpha$ ,  $C\beta$ ,  $C_O$  depend on torsion angles of the protein backbone. When the chemical shifts for certain residues are not provided by the user (in our case residues 15-19 and the C-ter of Tom22), the software determines propensities based only on the protein sequence and the matches in the database.

A second software for determining the local structure and disorder we explored was CheSPI (Chemical shift Secondary structure Population Inference) (Nielsen and Mulder, 2021). CheSPI was developed for detecting very small amounts of residual structure (in eight structural classes) from small chemical shift deviations. These deviations are calculated using pH, neighbor and temperature corrected random coil chemical shifts as a reference (POTENCI-Prediction Of Temperature, Neighbor and pH Corrected shifts for Intrinsically disordered proteins; Nielsen and Mulder, 2018). This allows detecting the presence of slightly populated stretches of residual secondary structure in intrinsically disordered protein sequences, such as the cytosolic domain of Tom22.

Chemical shifts of assigned backbone atoms of Tom22(1-74) (amide proton, amide nitrogen,  $C\alpha$ ,  $C\beta$  and carbonyl  $C_O$ ) were provided to both software tools. Secondary structure prediction performed by TALOS-N shows that most of the Tom22 residues are in an extended conformation and probably unfolded (loop conformation), with a small fragment from residue 55 to 68 that have an approximately 40% propensity for an  $\alpha$ -helix (Figure 4.1B, lower plot). At the C-terminus of the Tom22 receptor, the propensity for  $\beta$ -strand is approximately 30% however, for this region (residues 61-66) the score of TALOS-N prediction from chemical shifts is classified as dynamic and the prediction for the C-terminus (not assigned region 70-75) is based on sequence matches only. The same chemical shifts that were provided to TALOS-N were used for CheSPI. Similarly to the analysis by TALOS-N, most of the Tom22 sequence is predicted to be in non-folded conformation with only small fragment of residues 55-68 that show propensity of  $\sim 25\%$  for turn/helical conformation (Figure 4.1B, upper plot).

Residue-specific  $R_2$  relaxation rate constants report on motions occurring on time scales of  $[10^8\text{-}10^{12} \text{ s}^{-1}]$ , and thus give insight into overall molecular tumbling (nanosecond time scales) and local sub-nanosecond motion.

For the N-terminal part (first  $\sim 50$  residues) of the Tom22<sub>cyt</sub> the  $R_2$  rate constants are smaller (slower relaxation) suggesting that this part has a short effective correlation time, which we ascribe to large disorder (Figure 4.1C, plot on the left). For the C-terminus longer effective correlation time (slower tumbling) is observed, which is indicating that the C-terminus might be more structured.

The hetNOE values report on the residue wise motions on the picosecond time scale. Residues in folded structures generally have hetNOE values of around 0.8, while residues in flexible parts have values of around or below 0. The data confirm that a large part of the cytosolic domain of Tom22 receptor appears as intrinsically disordered. The sequence comprising ca. residues 55 to 68 has slightly positive values, pointing to less flexibility, in line with the  $\alpha$ -helix propensity revealed by the chemical shift and the elevated  $^{15}\text{N}$   $R_2$  values (Figure 4.1C, plot on the right).

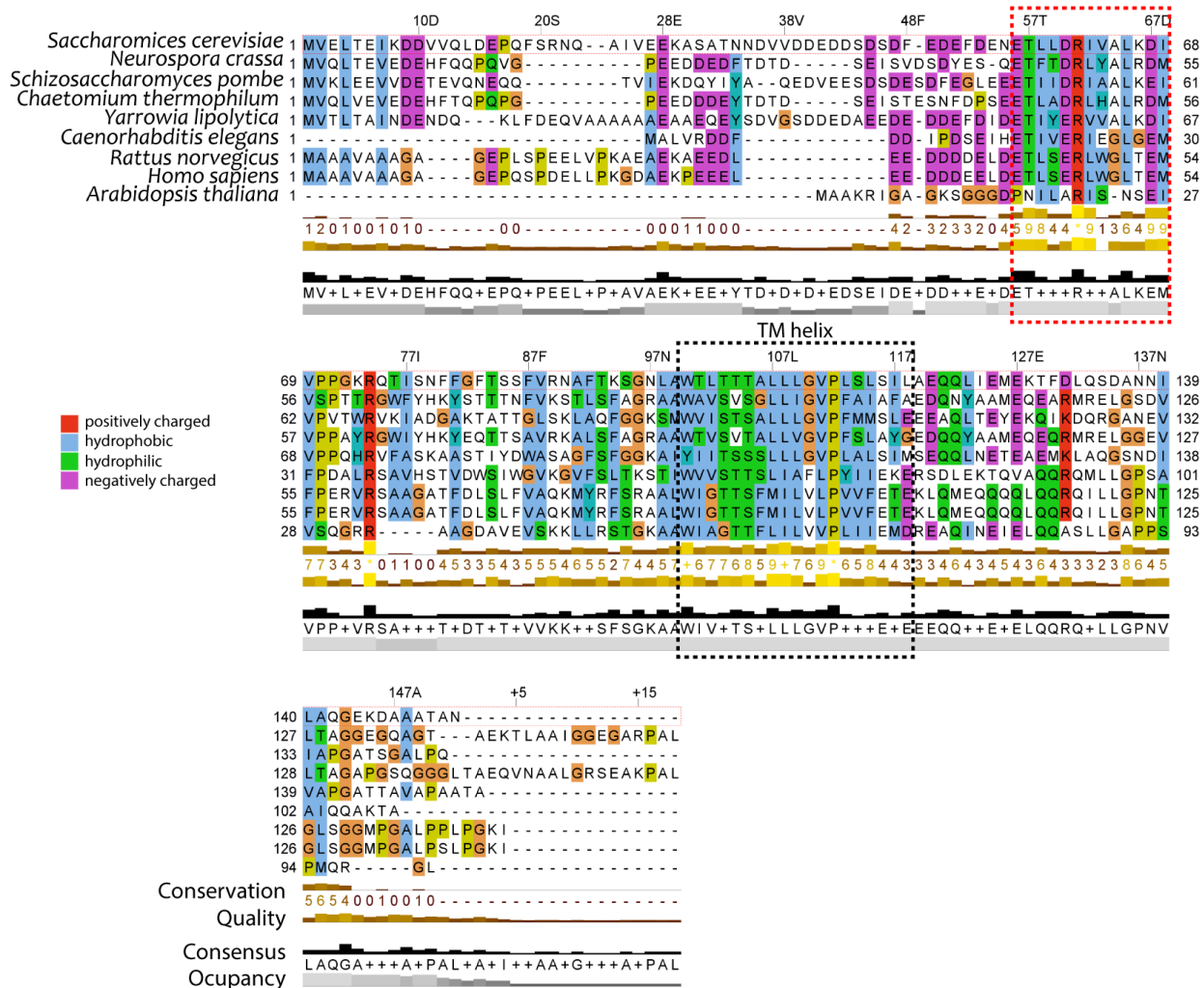
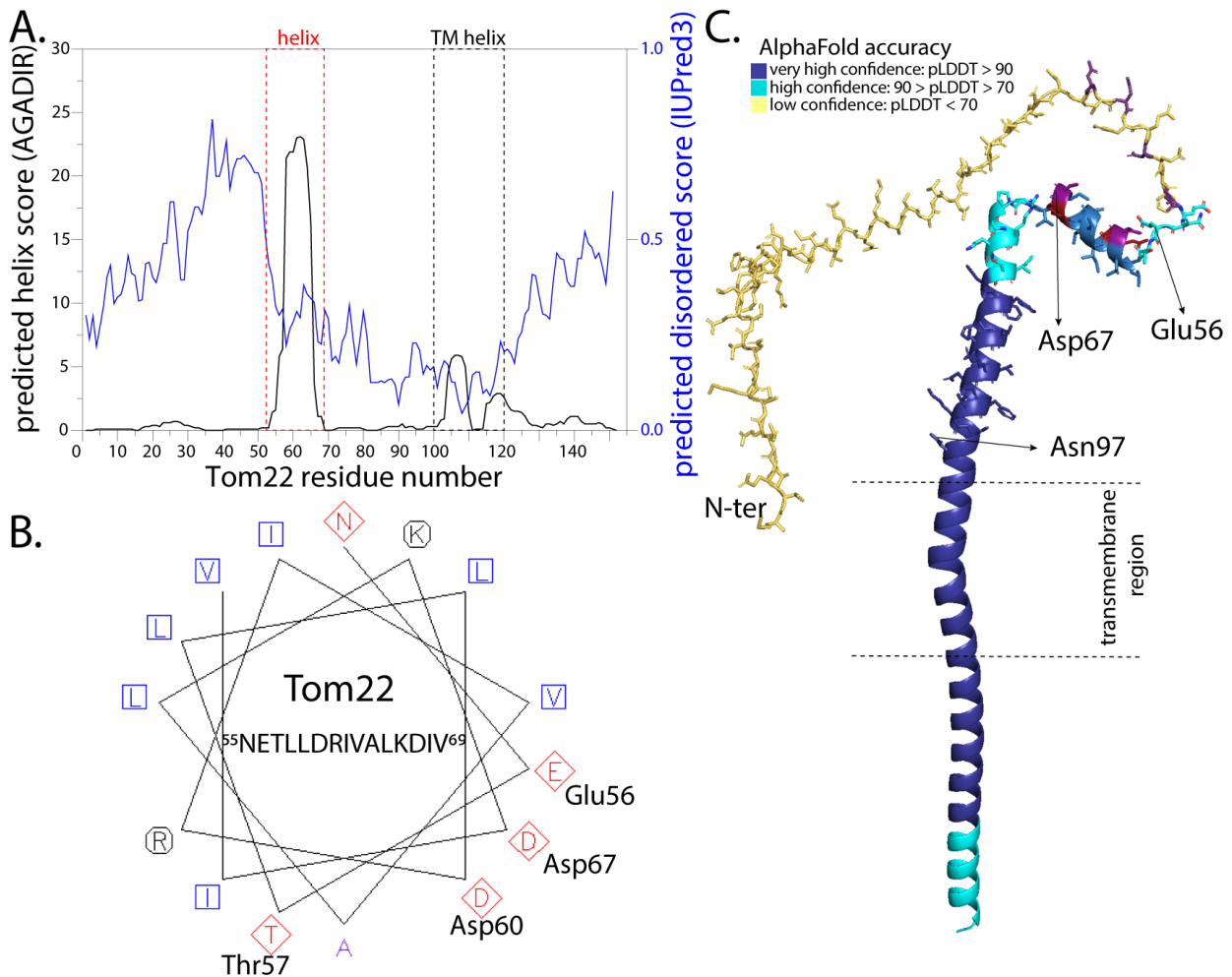


FIGURE 4.2: **Multiple sequence alignment and tertiary structure prediction of Tom22 receptor.** Protein sequence alignment of Tom22 receptor from selected organisms performed by UniProt Align software. Presence of conserved region from residue 56 to 70 with amino acid properties:  $-p\varphi\varphi-+\varphi x s\varphi p/+-\varphi\varphi$  after which the helix breaker (Pro) could be observed;  $\varphi$  stands for hydrophobic, p for polar, x for any and s for small residue. This region corresponds to the experimentally observed helical region of Tom22 (in red square). The Tom22 protein from *Arabidopsis thaliana* is included in this alignment, however it should be noted that the plant receptors arose by convergent evolution from a distinct ancestral gene and there is almost no sequence similarity with the rest of the Tom22 proteins. On the top of the alignment, numbering of residues of *S. cerevisiae* protein is shown.

Multiple sequence alignment of Tom22 protein from selected eukaryotic organisms reveals conservation of the region between residues 56 to 70 (in *S. cerevisiae* protein numbering) (Figure 4.2). The conservation of these residues cannot be shown with a strict amino acid consensus, but the physico-chemical properties of the residues are conserved. Residues from 56 to 70 have an amino acid properties:  $-p\varphi\varphi-+\varphi x s\varphi p/+-\varphi\varphi$  ( $\varphi$  hydrophobic;  $-$  negatively charged;  $+$  positively charged; p polar; s small; x any). Remarkably, these residues correspond to the region of Tom22 with experimentally observed helical propensities. Upstream



from this conserved helical region several negatively charged residues are conserved (E49, D50, E51, D53 in *S. cerevisiae*) but not their exact position. They form the conserved acidic region that was suggested to be involved in binding to the positively charged side of the presequences (Kiebler et al., 1993; Brix, Dietmeier, and Pfanner, 1997; Shiota et al., 2011). Finally, at the N-terminus of all yeast Tom22 proteins two negatively charged residues are present (D9, D10).



**FIGURE 4.3: Structure prediction of Tom22 receptor.** **A.** Sequence based prediction of the Tom22 disordered regions by IUPred3 (<https://iupred.elte.hu/>) by which the global structural disorder that encompasses at least 30 consecutive residues of the protein is predicted (plot and scale in blue). Prediction of the helical content of Tom22 obtained with Agadir algorithm (<http://agadir.crg.es/>; plot and scale in black). The two predictions complement each other, showing the disorder of the N-terminus of Tom22 and predicted helices in 55-70 (shown in red rectangle) and 105-125 (shown in black rectangle) corresponding to the transmembrane helix. **B.** Helix prediction of Tom22 residues 55-69 run by EMBOSS pepwheel software (<https://www.bioinformatics.nl/cgi-bin/emboss/pepwheel>) shows that this region of Tom22 could form an amphipathic helix with one hydrophobic side (residues in blue squares) and one negatively charged side (residues in red rhombus). **C.** Ternary structure prediction of Tom22 from the AlphaFold Protein Structure Database. The accuracy of the prediction is shown, with yellow being the not accurate prediction and the dark blue with the highest accuracy. Protein region shown with sticks is the 1-97 cytosolic Tom22 domain. Residues ~ 100-118 form the  $\alpha$ -helical transmembrane part (Tucker, 2019) of the receptor, residues within dashed lines.

Protein secondary structure and disorder can also be predicted from the protein sequence. The entire sequence of Tom22 was analyzed using two different algorithms. IUPred3 (<https://iupred.elte.hu/>) predicts the global structural disorder that encompasses at least 30 consecutive residues of the protein (Figure 4.3A, plot and scale in blue). The Agadir algorithm (<http://agadir.crg.es/>) predicts the  $\alpha$ -helical content of the

protein sequence (Figure 4.3A, plot and scale in black). Both algorithms indicate disordered N- and C-termini of Tom22 with a more ordered region that encompasses the predicted helices (residues 55-70) and the transmembrane region. We examined the properties of the first helix by drawing a helical wheel diagram for residues 56-68 of Tom22. These residues form an amphipathic helix with one hydrophobic and one negatively charged side (Figure 4.3B.). Remarkably, the presequences of the matrix targeted precursors, recognized by the Tom20 receptor, also form amphipathic helices but with one hydrophobic and one positively charged side (Schneider et al., 1998; Maduke and Roise, 1996).

AlphaFold is an algorithm that predicts three-dimensional structures of proteins with high precision (Jumper et al., 2021). The AlphaFold Structure Database provided a structural model for the full-length yeast Tom22 receptor. As expected for an intrinsically disordered region, most of the cytosolic domain of Tom22 could not be predicted with confidence (low accuracy score) (Figure 4.3C.). However, our experimentally determined helical region (56-67) was also predicted as a small helix by AlphaFold, with better accuracy score than the rest of the cytosolic domain but not as accurate as the transmembrane region. To be noted, the high accuracy score for the transmembrane region is expected considering that the structure of the transmembrane region is already in the database (from the cryo-EM structures of the TOM core complex).

Our experimental data from NMR relaxation studies and chemical shift measurements, show that the cytosolic domain of Tom22 receptor is mostly intrinsically disordered. This is in agreement with the secondary and tertiary structure predictions obtained by different prediction servers from the protein sequence as the only input. However, a short conserved sequence (residues 56-69) of Tom22 forms a partially populated, amphipathic  $\alpha$ -helix. Comparison of the relaxation parameters of the long Tom22(1-97) and short Tom22(1-74) suggests extension of the rigidity or foldedness of the residues 74-97. For this reason in most of the following interaction experiments the longer construct was also tested.

After characterizing the free Tom22 cytosolic domain, we proceed to the characterization of the interaction patterns of Tom22 with presequences of mitochondrial precursor proteins and with the cytosolic domains of the other receptors of the TOM complex, Tom20 and Tom70.

## 4.2 Interaction of the presequence and the Tom receptors

### 4.2.1 Tom22<sub>cyt</sub> does not bind the precursor presequences

It has been proposed that the cytosolic domain of Tom22 is involved in the binding of the mitochondrial precursor protein during import (Kiebler et al., 1993; Brix, Dietmeier, and Pfanner, 1997; Brix et al., 1999). This was based on the observation of a conserved charge complementarity between the negatively charged residues in the N-terminus of Tom22 and the positively-charged side of the precursor's amphipathic helix. Experimentally, it has been shown that the passage of the precursor towards Tom40 is blocked if antibodies against the cytosolic domain of Tom22 were added (Kiebler et al., 1993). Additionally, peptide-scan assays (Brix et al., 1999) and *in vitro* cross-linking experiments (Shiota et al., 2011) suggest Tom22-presequence interaction. However, confirmation of direct binding and the precise binding site of the presequence on Tom22 are still missing.

Having assigned the backbone NMR resonances and characterized the residual secondary structure of Tom22, we investigated the binding of the mitochondrial precursor presequences to Tom22 by NMR. We decided to test the interaction with the presequences using the longer Tom22 construct to ensure that all possible interaction sites on the cytosolic domain were available. The construct that we used for these experiments was Tom22(1-97) with the GB1 solubility tag. Keeping the GB1-tag enabled faster determination of protein sample concentration and reduced further losses in the protein yield, which is already quite low

for this construct, by avoiding additional purification steps. GB1, a well-folded small protein domain is not expected to interact by its own. In addition, such an interaction would be visible in the NMR spectra.

The first presequence construct that was tested was GB1-pSu9(1-69). This is a fusion protein construct comprising pSu9, the first 69 residues of a mitochondrial subunit 9 of the F<sub>0</sub> ATP synthase of *Neurospora crassa* that has been extensively used in *in vivo* and *in organello* studies (reviewed in Maduke and Roise, 1996; Pfanner, 2000), with an N-terminal solubility and purification tag His<sub>6</sub>-GB1. This construct precipitated immediately upon mixing with the Tom22 protein sample (see Methods 7.4.4 and Section 4.2.3), and we were not able to see any chemical-shift perturbation with a sample containing isotope-labeled Tom22 and pSu9.

We then tested different peptides that correspond to other known presequences: pF1 $\beta$ (21-33), pATP1(1-20) and pSu9(1-25) and one peptide corresponding to an internal matrix targeting signal peptide of ATP1-iMTS-L (residues 306-324) (full peptide sequences in Methods 7.5). These peptides were added to <sup>15</sup>N-labelled Tom22 and NMR spectra were acquired. From the superposition of the 2D <sup>15</sup>N<sup>1</sup>H BEST-TROSY NMR spectra of Tom22(1-97) alone or in the presence of the peptides, no changes in the peak positions and therefore in the chemical environment of Tom22 residues could be observed (Figure 4.4).

From these observations we can conclude that the cytosolic domain of Tom22 does not interact with the presequences of the mitochondrial precursor proteins under our experimental conditions.

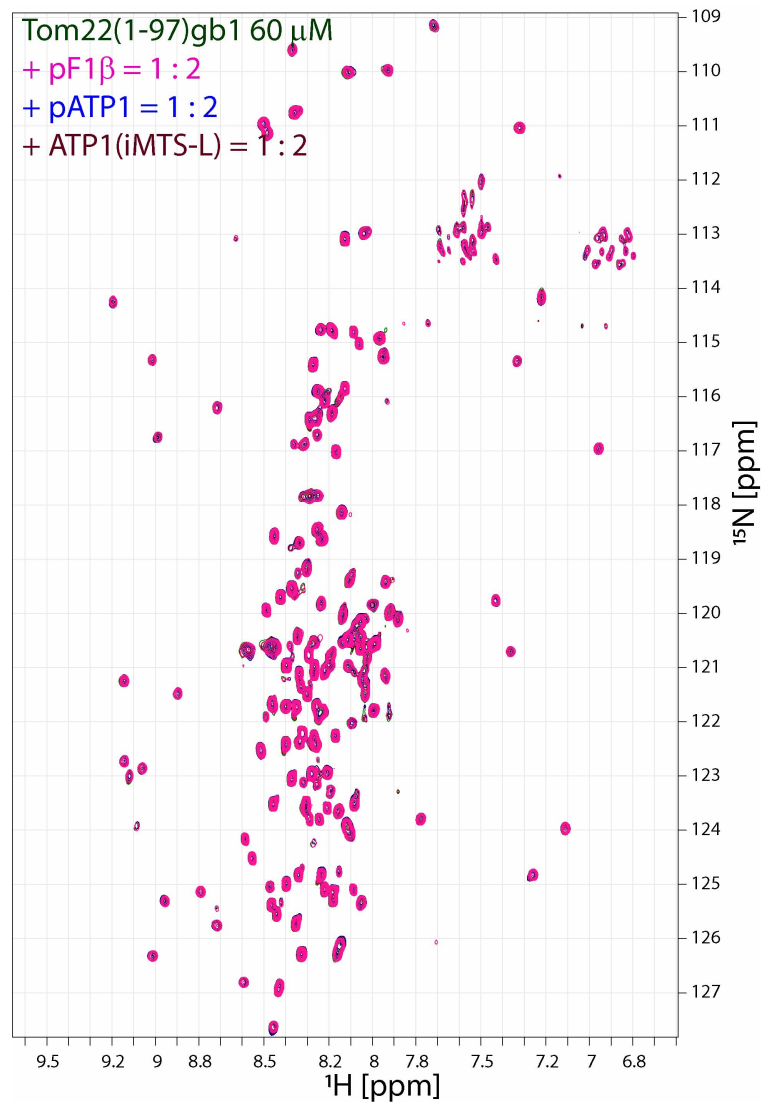


FIGURE 4.4: **Tom22 does not interact with the presequences of the mitochondrial precursor proteins.**  $^{15}\text{N}^1\text{H}$  BEST-TROSY spectra of Tom22(1-97)-GB1 alone shown in green, in the presence of pF1 $\beta$  in pink, pATP1 in blue or in the presence of ATP1-iMTS-L peptide in brown. All spectra are perfectly superimposing indicating no changes in the chemical environment of Tom22 residues in the presence or absence of the peptides. The more dispersed peaks, outside of the 7.8-8.7 ppm range, correspond to the GB1-solubility tag, a well-folded protein domain with known chemical shifts.

#### 4.2.2 Presequence binding site on the Tom20 receptor

To determine and characterize the binding site of the presequences of mitochondrial precursor proteins on the Tom20 receptor we initially used the cytosolic domain of Tom20 receptor from *S. cerevisiae* (ScTom20(37-183)), since the Tom22 protein used in our studies is also from yeast. Assignment of the backbone chemical shifts of the yeast Tom20 (ScTom20) was not straightforward and only 38% of the protein sequence could be assigned. The interpretation of the spectra was difficult as a much larger number of peaks than expected was observed. In addition, the peak intensity was highly variable, further complicating the assignment process (see blue spectrum in Appendix B, Figure S2). Using SEC and analytical ultracentrifugation (AUC) at different concentrations of Tom20, we could demonstrate that under our experimental conditions Tom20 is a mixture of monomer and dimer (see Appendix B, Figure S1).

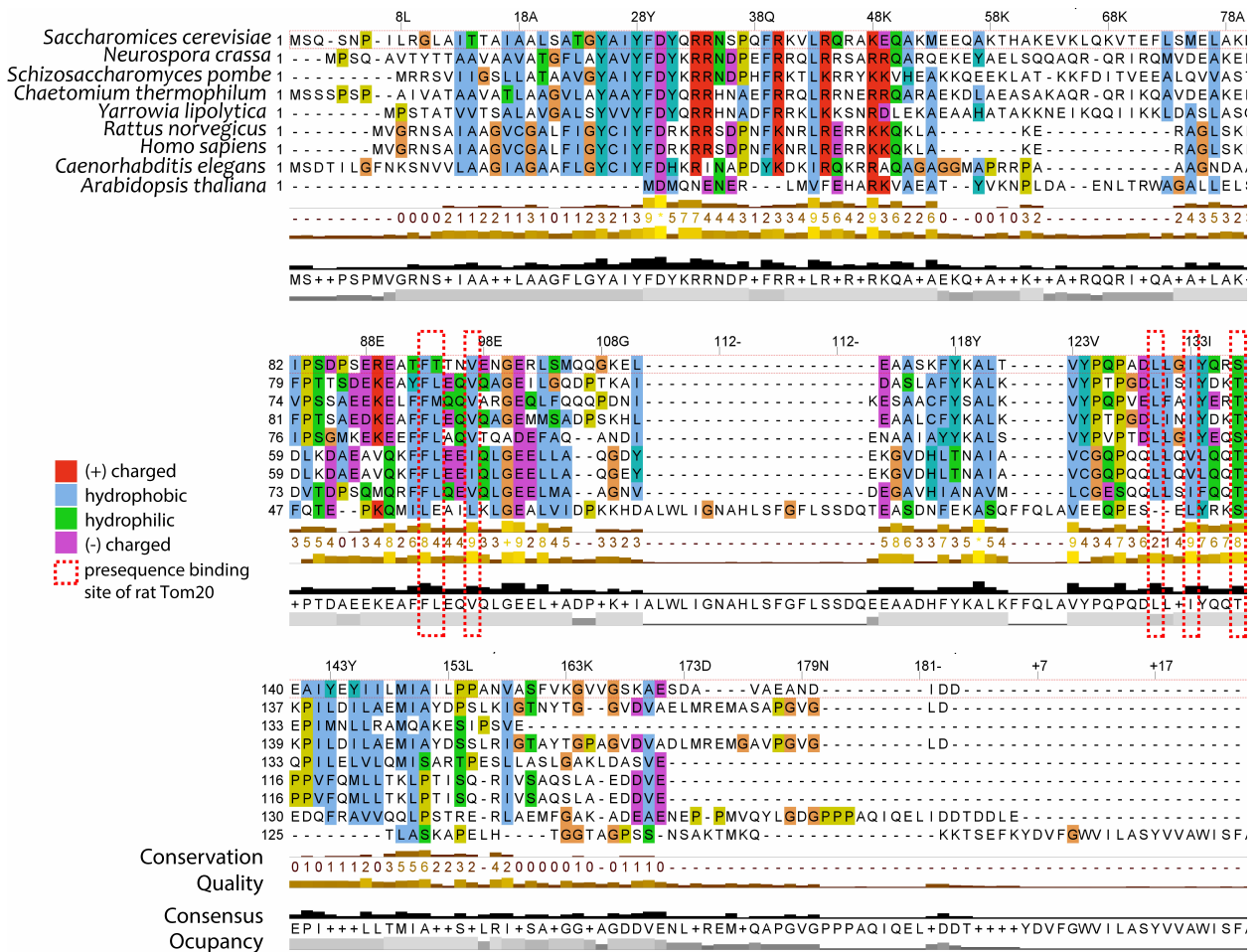


FIGURE 4.5: Multiple sequence alignment of Tom20 protein from different organisms. Protein sequence alignment of Tom20 receptor from selected organisms performed by UniProt Align software. The residues shown to be involved in the binding of the presequence in the Tom20 receptor from rat are shown in red dashed rectangles.

On the other hand, the cytosolic domain of the rat Tom20, RnTom20, was previously studied by NMR. The backbone and side chain chemical shifts have been assigned and the presequence binding site has been shown (Abe et al., 2000) (PDB:1om2). The rat protein shows a high sequence similarity with the yeast Tom20 proteins (Figure 4.5) and can complement the functional defects of Tom20-deficient yeast cells (Swie Goping, Millar, and Shore, 1995; Iwahashi et al., 1997). In order to verify observations made on the incompletely assigned yeast Tom20 (ScTom20) cytosolic domain, we also used the corresponding rat Tom20 (RnTom20) construct. The chemical shift assignments from Abe *et al.* were transferred to the RnTom20(51-145) spectra.

Since the structure of the yeast Tom20 receptor is not solved so far, we constructed structural models of yeast Tom20(37-183) with Robetta, a protein structure prediction service (<https://robetta.bakerlab.org/>), using a deep learning based modeling method RoseTTAFold (Baek et al., 2021). The presequence binding site on yeast Tom20 could then be predicted from the superposition of the crystal structure of the rat Tom20 in the presence of pALDH presequence peptide (PDB: 2v1s, (Saitoh et al., 2007)) and the yeast Tom20 model structures.

NMR was used in order to determine the presequence binding site on ScTom20(37-183) by titration experiments. Isotopically labeled (NMR visible) yeast Tom20(37-183) was titrated with different presequence peptides (pF1β(21-33), pSu9(1-25), and ATP1(306-324)iMTS). Five-fold molar excess of the presequence peptide were added to the labeled ScTom20 sample and the changes in the <sup>15</sup>N<sup>1</sup>H spectrum of each sample

were followed (see spectra overlay in Supp. Figure S2). Residue-wise plotting of the changes in the chemical shifts (CSPs) and peak intensity differences between the apo Tom20 sample and the samples in presence of different presequences, enables identification of the Tom20 residues affected by the presence of the presequence (Figure 4.6A.). After mapping the presequence-affected residues on the structural model of the yeast Tom20 we can conclude that the presequences (pSu9(1-25), pF1 $\beta$ , GB1-pSu9(1-69)) and the matrix targeting signal peptide (ATP1(306-324)iMTS-L) all bind to the same region of the yeast Tom20 receptor (Figure 4.6B.). This region corresponds to the presequence binding site that was predicted considering structural similarity of the ScTom20 model structure and the RnTom20 cytosolic domain (Figure 4.6B.). Yeast Tom20 residues involved in binding of the presequences are: Thr94, Asn96, Leu131, Ile133 and Tyr134. Note that the hydrophobic residues are conserved within the Tom20 receptors, further corroborating their functional importance in presequence binding (see Tom20 sequence alignment in Figure 4.5).

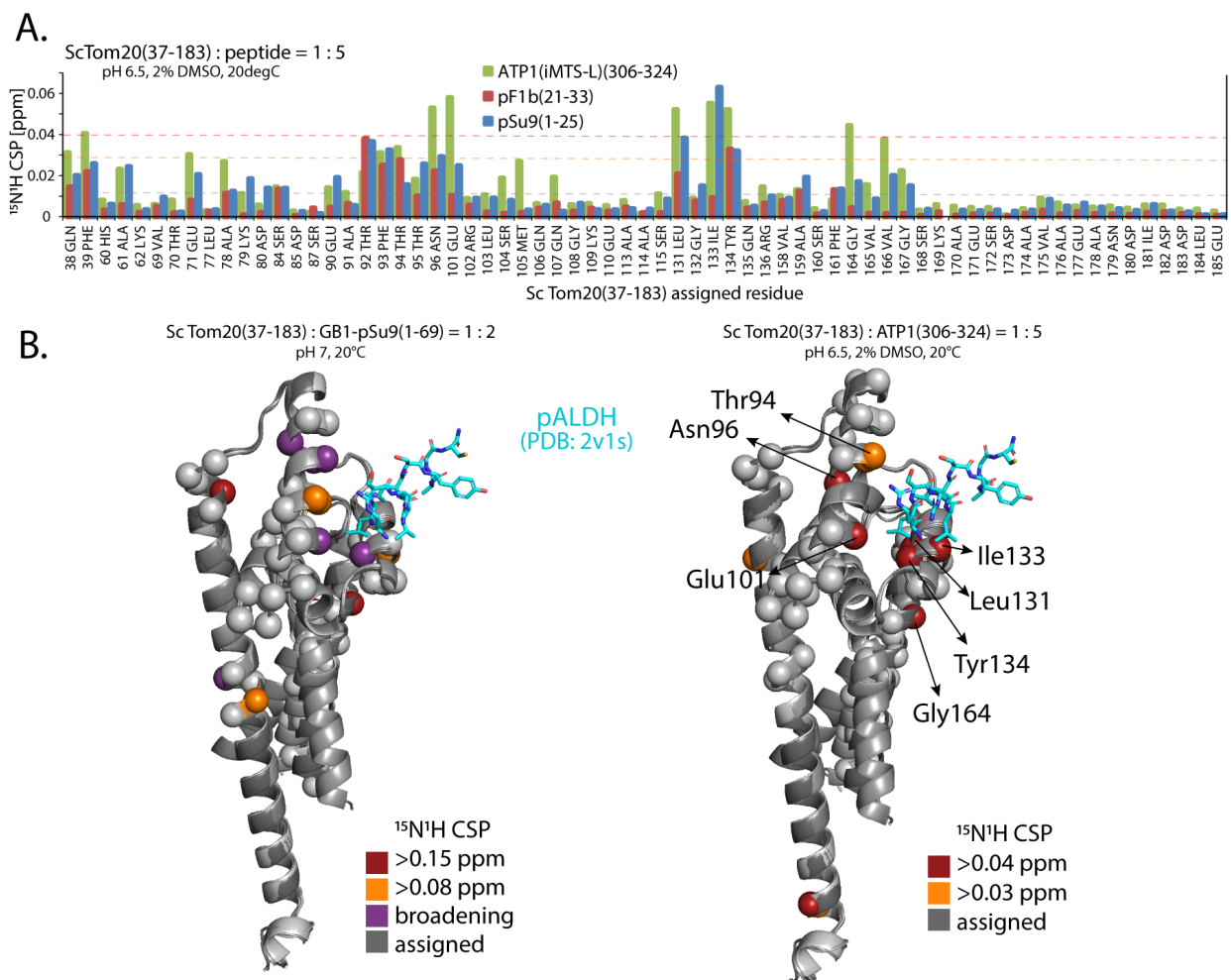


FIGURE 4.6: **Yeast Tom20 receptor binds the presequence in the predicted binding site.** **A.**  $^{15}\text{N}/^1\text{H}$  CSP of the assigned ScTom20(37-183) residues upon adding five fold excess of the presequence peptides: pF1 $\beta$ (21-33) in red, pSu9(1-25) in blue and ATP1(306-324)iMTS-L in green. The standard deviation over all CSP induced by adding ATP1(iMTS-L) (0.015 ppm) is indicated with the gray horizontal dashed line. Dashed red and orange horizontal lines indicate an arbitrary cut-off for CSPs shown on the ScTom20 model (see panel (B.) of this figure). For comparison, the standard deviation over all CSP induced by adding pSu9(1-25) peptide is 0.011 ppm. **B.** Projection of significant CSPs or line-broadening on five superimposed ScTom20(37-183) Robetta model structures upon adding five fold excess of ATP1(306-324)iMTS-L peptide (model on the right hand side) or upon adding two-fold excess of GB1-pSu9(1-69) (model on the left hand side). The coordinates of pALDH peptide from the crystal structure of RnTom20 ((Saitoh et al., 2007); PDB:2v1s) are shown in cyan stick representation to indicate presequence binding site on the rat protein.

The incomplete backbone chemical-shift assignment of the yeast Tom20(37-183) construct results in a lower



number of available probes for the NMR-detected interaction studies. Therefore, the binding site of the presequences was verified on the cytosolic domain of rat Tom20 for which a complete assignment is available. The presequence peptide binding site on RnTom20(51-145) was confirmed by residue-wise plotting of the chemical shift perturbations induced by addition of the following peptides: pATP1(1-20), pALDH(12-22), pF1 $\beta$ (21-33), pALDH(1-22) and ATP1(306-324)iMTS-L in a five-fold molar excess (Figure 4.7A.). The presequence peptide pATP1(1-20) induced the biggest CSPs. They were mapped on the NMR structure of rat Tom20 (Abe et al., 2000; PDB: 1om2) and confirmed the presequence binding site as well as involvement of the conserved hydrophobic residues: Leu71, Ile74, Val99, Leu107, Leu110 and Thr113 (Figure 4.7B. and alignment in Fig.4.5).

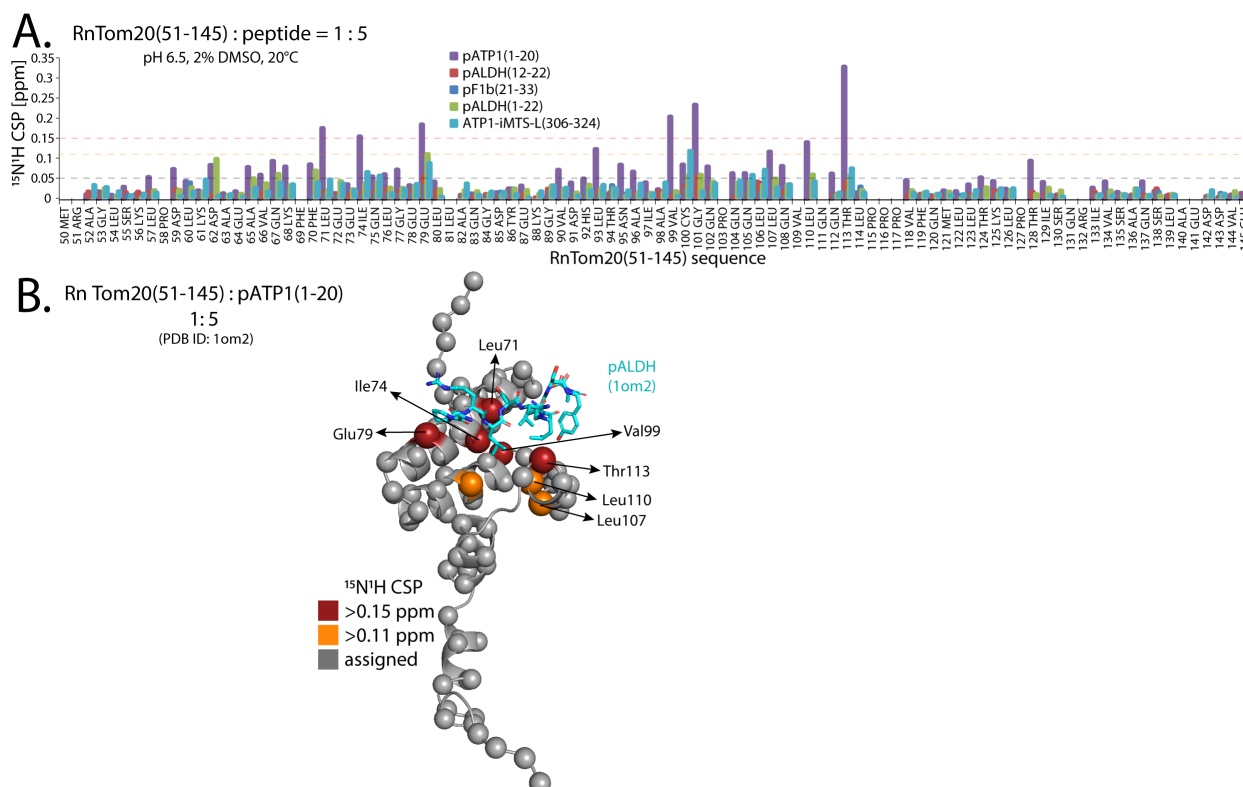


FIGURE 4.7: **Confirmation of the presequence binding site on the rat Tom20 receptor.** **A.** Combined  $^{15}\text{N}^1\text{H}$  CSP of the rat Tom20(51-145) residues upon adding a five fold excess of the following peptides: pATP1(1-20) (purple), pALDH(12-22) (red), pF1 $\beta$ (21-33) (dark blue), pALDH(1-22) (green) and ATP1(306-324)iMTS-L (light blue). The standard deviation over all CSPs induced by pATP1 peptide (purple plot) is indicated with a gray horizontal dashed line (0.057 ppm) - below this line the CSPs are considered non-significant. Dashed red and orange horizontal lines indicate an arbitrary cut-off for pATP1 induced CSPs, mapped on the RnTom20(51-145) structure in panel (B.) of this figure. **B.** Residues with significant CSPs or peak broadening upon adding five fold excess of pATP1(1-20) peptide (purple in panel A) are shown as colored spheres. The coordinates of the NMR structure of RnTom20 bound to the pALDH presequence peptide (in cyan) is shown here (Abe et al., 2000; PDB: 1om2).

### 4.2.3 Presequence-containing client protein and its interactions with the Tom receptors

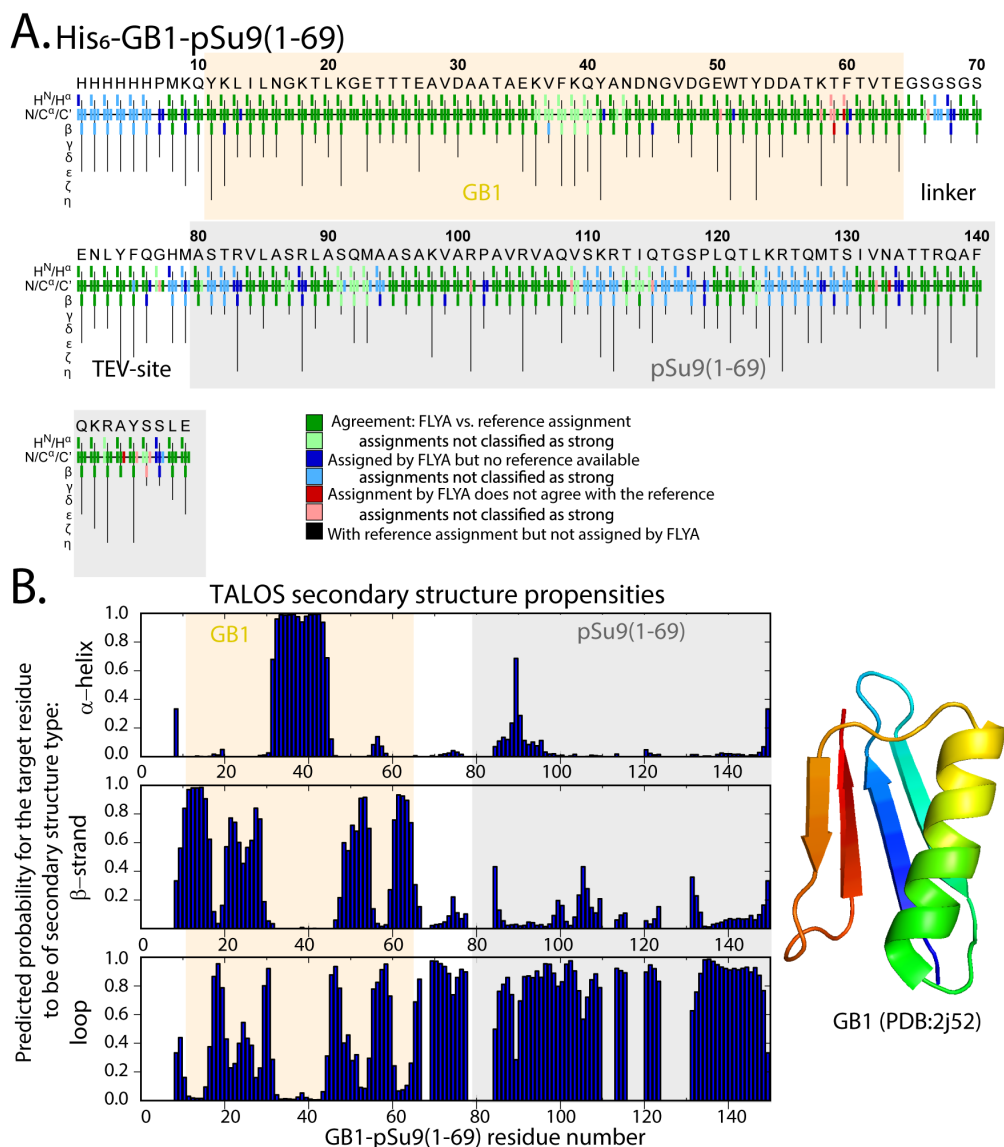
The previous NMR experiments allowed identification of the presequence binding site on ScTom20(37-183). In absence of Tom20, these peptides are believed to be unstructured in solution. It would be interesting to determine whether they form the amphipathic helix observed in the previous NMR structure, in which a six-residue long helix has been identified (Abe et al., 2000). As shown in the preceding chapters, systematic backbone chemical shift deviation, including  $^{15}\text{N}$ , and  $^{13}\text{C}$ , allow identification of even partially populated

secondary structure elements. This however requires isotopic labelling, which is not convenient for small peptides that cannot be obtained by recombinant expression in bacteria. Therefore, in order to investigate how the presequence behaves upon binding to Tom20, we used the fusion protein construct comprising pSu9(1-69), the first 69 residues of a mitochondrial subunit 9 of the F<sub>0</sub> ATP synthase of *Neurospora crassa* with the N-terminal solubility and purification tag, His<sub>6</sub>-GB1. This tag ensured stable bacterial expression, ease of purification and an increased solubility of our construct (see Methods 7.4.4). It was therefore possible to obtain an isotopically labelled, presequence-containing fusion protein. It is expected that the GB1 tag, a well-folded globular domain, interacts neither with the presequence pSu9(1-69) nor with Tom20. This assumption has been verified by comparing the NMR resonances to the known chemical shift assignments (BMRB ID: 27169).

The backbone <sup>15</sup>N, <sup>13</sup>C, <sup>1</sup>H resonance assignment of the GB1-pSu9(1-69) construct resulted in approximately 66% of the HN resonances of the pSu9(1-69) part being assigned (see Figure 4.8A.). The resonance assignment has been submitted to the BioMagResBank (BMRB: [www.bmrb.wisc.edu](http://www.bmrb.wisc.edu); entry number: 51093). The 2D <sup>15</sup>N<sup>1</sup>H fingerprint spectrum of the GB1-pSu9(1-69) construct (Figure S4 in Appendix B) revealed that the peaks corresponding to the pSu9(1-69) resonances have a low dispersion in the <sup>1</sup>H dimension, characteristic for an IDP. The peaks of GB1 are well dispersed, as expected for a well folded protein domain. To assess the secondary structure of our GB1-pSu9 construct we provided the H, N, C $\alpha$ , C $\beta$  and C $\text{O}$  chemical shifts to the TALOS-N software. The secondary structure propensity prediction performed by TALOS-N showed that the GB1 domain adopts both  $\beta$  and  $\alpha$  secondary structural motifs, in excellent agreement with its known structure (Figure 4.8B.). As expected, pSu9(1-69) residues appear as a mainly unstructured protein domain. However, residues 7-17 of pSu9(1-69) (corresponding to residues 85 and 95 of the full GB1-pSu9(1-69) construct) show an approximately 20% propensity for  $\alpha$ -helix formation. The observation that the pSu9 presequence is primarily disordered is in agreement with previous observations of the CoxIV presequence that was found to be disordered in aqueous solution (Maduke and Roise, 1996).

Both hetNOE and R<sub>2</sub> NMR relaxation experiments confirm that the GB1 domain of our construct is well folded, with hetNOE values around 0.8 (Figure 4.9 left panel) and higher R<sub>2</sub> rate constants compared to the pSu9(1-69) domain (Figure 4.10). In contrast, residues corresponding to the linker between GB1 and pSu9(1-69) as well as the pSu9(1-69) part of the construct show hetNOE values around 0.2 or below (Figure 4.9 left panel) and reduced R<sub>2</sub> relaxation rates compared to the GB1 portion (Figure 4.10). This further confirms that the pSu9(1-69) sequence is much more flexible than the GB1 part of the construct. Analysis of the backbone dynamics therefore confirms the disordered nature of the pSu9(1-69) sequence in solution as well as the propensity to form a transient  $\alpha$ -helical stretch between residues 7 and 17 (pSu9 numbering).





**FIGURE 4.8: Output of the automated resonance assignment with FLYA and secondary structure prediction for GB1-pSu9(1-69) construct.** **A.** The FLYA software, an automated resonance assignment algorithm for NMR chemical shift assignments based on the peak lists from a combination of multidimensional through-bond NMR experiments, was used to verify and complement the manual assignment of the GB1-pSu9(1-69) protein construct. The graphic representation of the sequence shows how well the manual and automated assignments agree. Only green, assignments verified by FLYA, were kept and submitted to the BMRB database. The GB1 protein sequence is marked with box in wheat color and the pSu9(1-69) protein sequence is indicated with gray box. **B.** Secondary structure propensities calculated with TALOS-N (Shen and Bax, 2013) (<https://spin.niddk.nih.gov/bax/nmrserver/talosn/>) for the GB1-pSu9(1-69) construct. The pSu9(1-69) protein sequence is indicated with gray box. The secondary structure propensity prediction for GB1 domain adopts both  $\beta$  and  $\alpha$  secondary structural motifs, in excellent agreement with its known structure shown on the left.

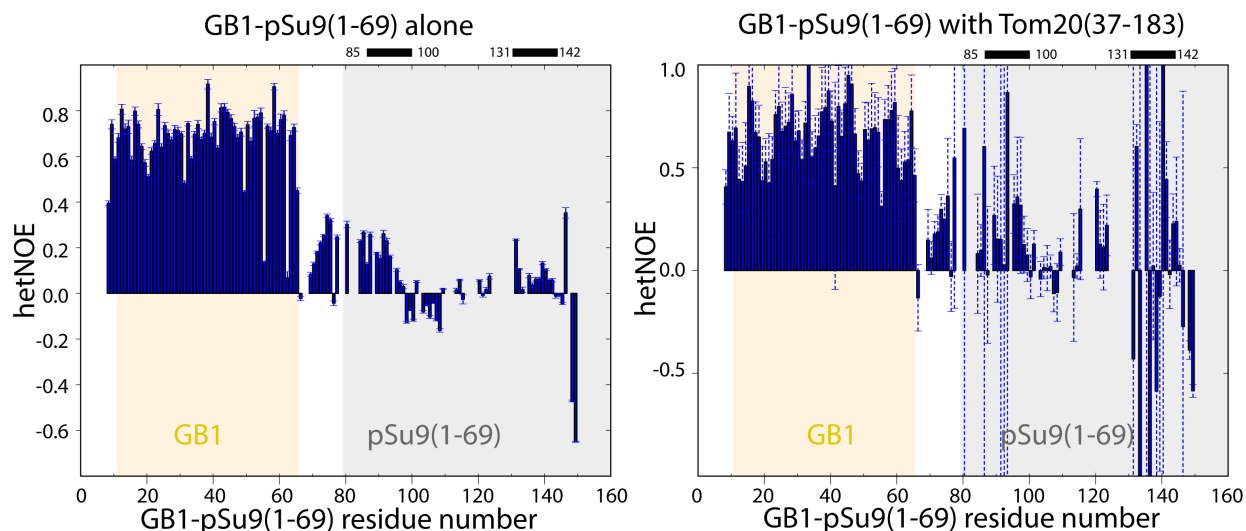


FIGURE 4.9: **hetNOE relaxation experiments on GB1-pSu9(1-69) construct alone and upon adding ScTom20.** Plotted hetNOE values of apo-GB1-pSu9(1-69) (shown in the left panel), and GB1-pSu9(1-69) residues upon addition of Tom20(37-183) (shown in the right panel). The big error bars are due to peak broadening (low intensities) observed for many pSu9 peaks upon interaction with Tom20. The residues corresponding to the pSu9(1-69) are inside the gray box. Black bars indicate two ScTom20 interacting regions of pSu9(1-69), observed from the CSP and intensity ratios upon adding ScTom20 in the sample (from Figure 4.11)

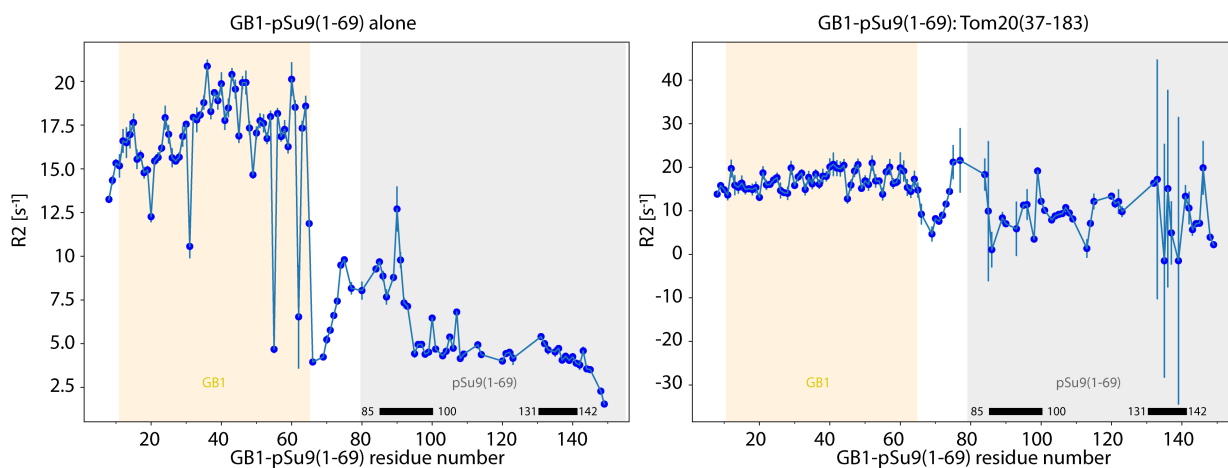


FIGURE 4.10: **Interaction of GB1-pSu9(1-69) with Tom20(37-183).** Comparison of the  $R_2$  relaxation rates between the GB1-pSu9(1-69) alone (on the left hand side) and GB1-pSu9(1-69) upon addition of Tom20(37-183) (on the right hand side). The residues corresponding to the pSu9(1-69) are inside the gray box. The difference in the scale of the rate axis is due to big error bars in the plot for pSu9-Tom20. Black bars indicate two ScTom20 interacting regions of pSu9(1-69), observed from the CSP and intensity ratios upon adding ScTom20 in the sample (from Figure 4.11).

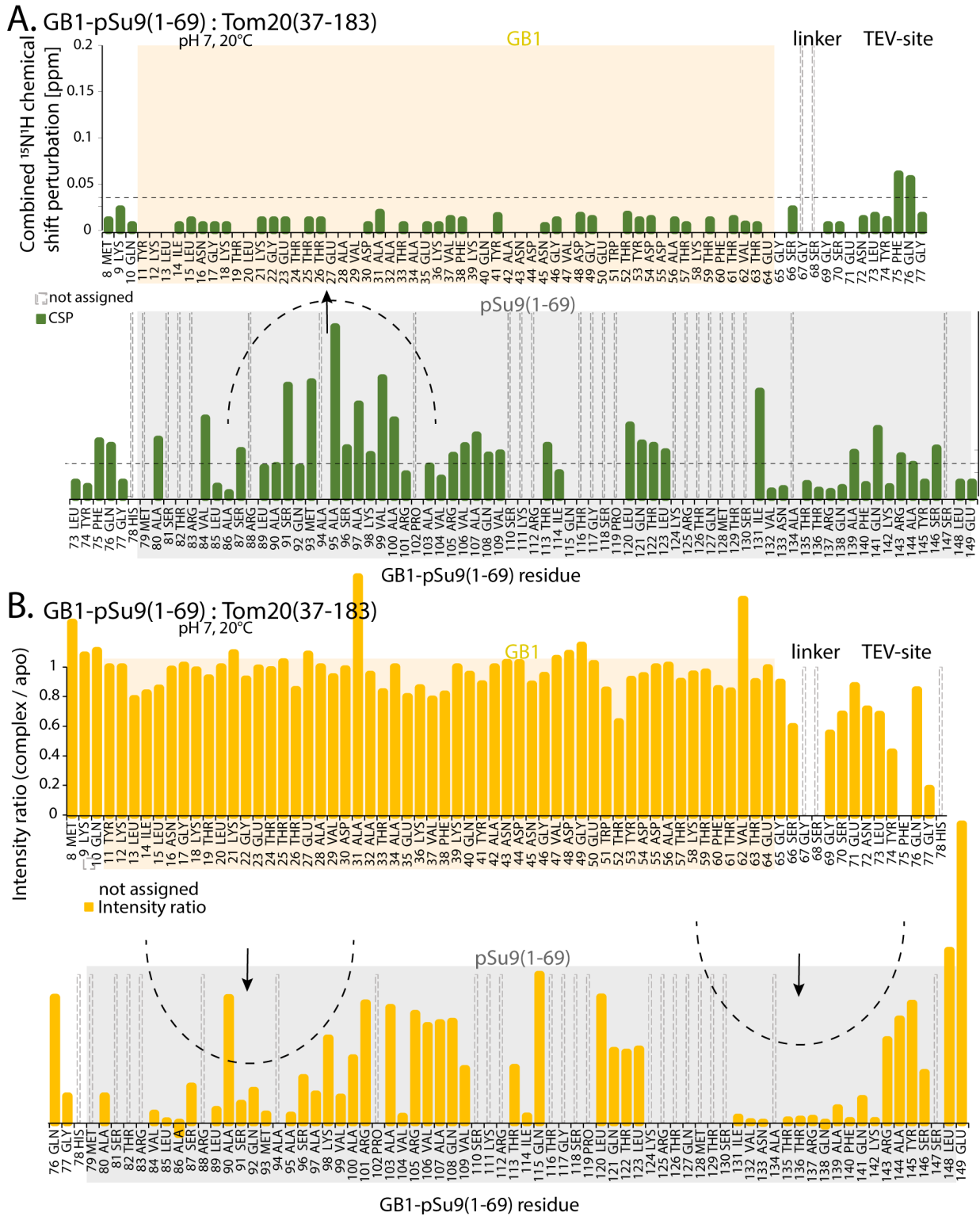


FIGURE 4.11: **Interaction of GB1-pSu9(1-69) with Tom20(37-183).** **A.**  $^{15}\text{N}-^1\text{H}$  chemical shift perturbation of GB1-pSu9(1-69) residues upon addition of Tom20(37-183). The residues corresponding to the pSu9(1-69) are inside the gray box. The standard deviation over all CSPs is indicated with horizontal dashed line - below this value the CSP is considered non-significant. **B.** Intensity ratios of GB1-pSu9(1-69) residues upon adding ScTom20(37-183). Same experiment as in panel (A.). Intensity ratio calculated as a ratio between peak intensities in the pSu9 sample with Tom20 added over the intensity ratios of the peaks in the apo pSu9 sample. Plots with CSPs and intensity ratios are split in two parts for clarity due to long protein sequence.

Having the chemical shift assignment and characterized secondary structure of our pSu9 construct we could now perform binding experiments and identify behaviour of the presequence once it binds to the Tom20 receptor. For that we added unlabeled yeast Tom20(37-183) protein to  $^{15}\text{N}$ -labeled GB1-pSu9(1-69) to monitor the binding from the pSu9(1-69) side. We observed shifts in the  $^{15}\text{N}^1\text{H}$  frequencies and the broadening of certain peaks (Figure S5). Calculated chemical shift perturbations (CSPs) and intensity ratios plotted as a function of GB1-pSu9(1-69) residue number can be seen in Figure 4.11A. and B.

As expected, residues of the GB1 part of the fusion construct did not show any significant CSP neither changes in the peak intensities, indicating that the GB1 part does not interact with Tom20 (Figure S5). The largest CSPs of pSu9(1-69) could be observed for the region between residues 13 and 22 (residues 91-100 of the full GB1-pSu9(1-69)) (Figure 4.11A.). Additionally, significant peak broadening can be seen for the regions between residues 7 and 22 (residues 85-100 of the full GB1-pSu9(1-69)), and residues 53 and 64 (residues 131-142 of the full GB1-pSu9(1-69)) (Figure 4.11B.). This indicates presence of two Tom20 binding regions of pSu9(1-69) construct. This is in agreement with previous observations that pSu9(1-69) has at least two regions that are important for mitochondrial import, where the first 14 residues are essential for initial docking to the TOM complex (Esaki et al., 2004).

From the NMR relaxation experiments performed on the sample containing labeled GB1-pSu9(1-69) and unlabeled Tom20(37-183), we could observe a slight increase in the hetNOE and  $R_2$  values, compared to what was observed on GB1-pSu9(1-69) alone (Figure 4.9 panel on the right, and Figure 4.10). pSu9(1-69) residues that showed the largest CSPs also have increased hetNOE and  $R_2$  values. The increase in hetNOE values of these residues upon interaction with Tom20 indicates a decrease in backbone flexibility upon binding to Tom20. The observed global increase in the  $R_2$  values measured for residues of the pSu9(1-69) part can be explained by the increase in the apparent size of the pSu9 domain in the GB1-pSu9 construct upon binding to Tom20. No changes of the hetNOE or  $R_2$  values were observed for residues of the GB1 part. This demonstrates no detectable interaction of GB1 and Tom20. Due to severe peak broadening of certain pSu9(1-69) residues upon interaction with Tom20, the constants for the complex sample could not precisely be determined, which results in a significant uncertainty reflected by the large error bars in Figure 4.10.

Direct interaction between the GB1-pSu9(1-69) and Tom22(1-97) receptor domain could not be tested. Immediately upon mixing of these two soluble protein samples precipitation was observed. Due to the loss of peak intensities of  $^{15}\text{N}$  labeled GB1-pSu9(1-69) when mixed with unlabeled Tom20, it was concluded that the precipitation is coming from the GB1-pSu9(1-69). This observation, in addition to the titration experiments performed with the presequence peptides and the cytosolic domain of Tom22 (shown above), are indicating that there is no direct interaction between the presequence and the Tom22 when these two are present alone in the sample.

Together, the interaction studies of GB1-pSu9(1-69) and the cytosolic domain of yeast Tom20 receptor revealed a site specific binding of the pSu9(1-69) presequence to the Tom20 presequence binding site. This unusually long presequence can bind with two distinct regions, residues 7-22 and 53-64 (residues 85-100 and 131-142 of the full GB1-pSu9(1-69)). Interestingly, both regions show a certain tendency to populate folded structure even in absence of Tom20, as revealed by the relaxation data. In addition, residues 7-22 clearly have some tendency to form an alpha-helix in solution. It would be interesting to perform a full backbone assignment of GB1-pSu9(1-69) in presence of Tom20 to determine possible secondary structure changes upon binding.

### 4.3 Interaction of Tom20 and Tom22 cytosolic domains

The interaction between Tom22 and Tom20 has been suggested to play a role in docking the precursor-loaded Tom20 to the TOM core complex, however it is unclear whether Tom22 is directly involved in the presequence recognition or it is modulating the binding of Tom20 and the presequence (Van Wilpe et al., 1999; Yamano et al., 2008b). Our results (presented above 4.2.1) show no direct interaction of the Tom22 cytosolic domain and the presequence, suggesting that the physiological role of Tom22 within the TOM complex does not involve specific recognition of presequences on their way to the TOMcc.

On the other hand, we confirmed the presequence binding sites on yeast and rat Tom20 cytosolic domains from NMR measurements.

The binding of Tom20 and Tom22 was shown by the *in vitro* cross-linking experiments, suggesting a proximity of the precursor binding site of Tom20 and the conserved acidic region of Tom22 (just upstream of the cytosolic  $\alpha$ -helix) (Shiota et al., 2011). However, some cross-linking agents have a quite long range, 9.6 Å for BPA for example, raising the question of the precision at which we can conclude about the exact binding sites.

In order to address the functional role of Tom22 within the TOM complex, our next aim was to determine whether Tom22 binds to Tom20, and eventually to characterize the interface between Tom20 and Tom22.

#### 4.3.1 Tom22 binds the presequence-binding site of the Tom20 receptor

To identify the Tom22 binding site on Tom20 we titrated isotopically labeled ScTom20 with unlabelled Tom22 (see spectra in Appendix B, Figure S3). When a 2-fold excess of Tom22(1-97)-gb1 was added, several peaks shifted or were broadened (Figure 4.12, residues Thr92, Thr94, Ile133, Tyr134). Remarkably, the corresponding residues belong to the same region that has previously been shown to interact with the presequences (Figure 4.6B.).

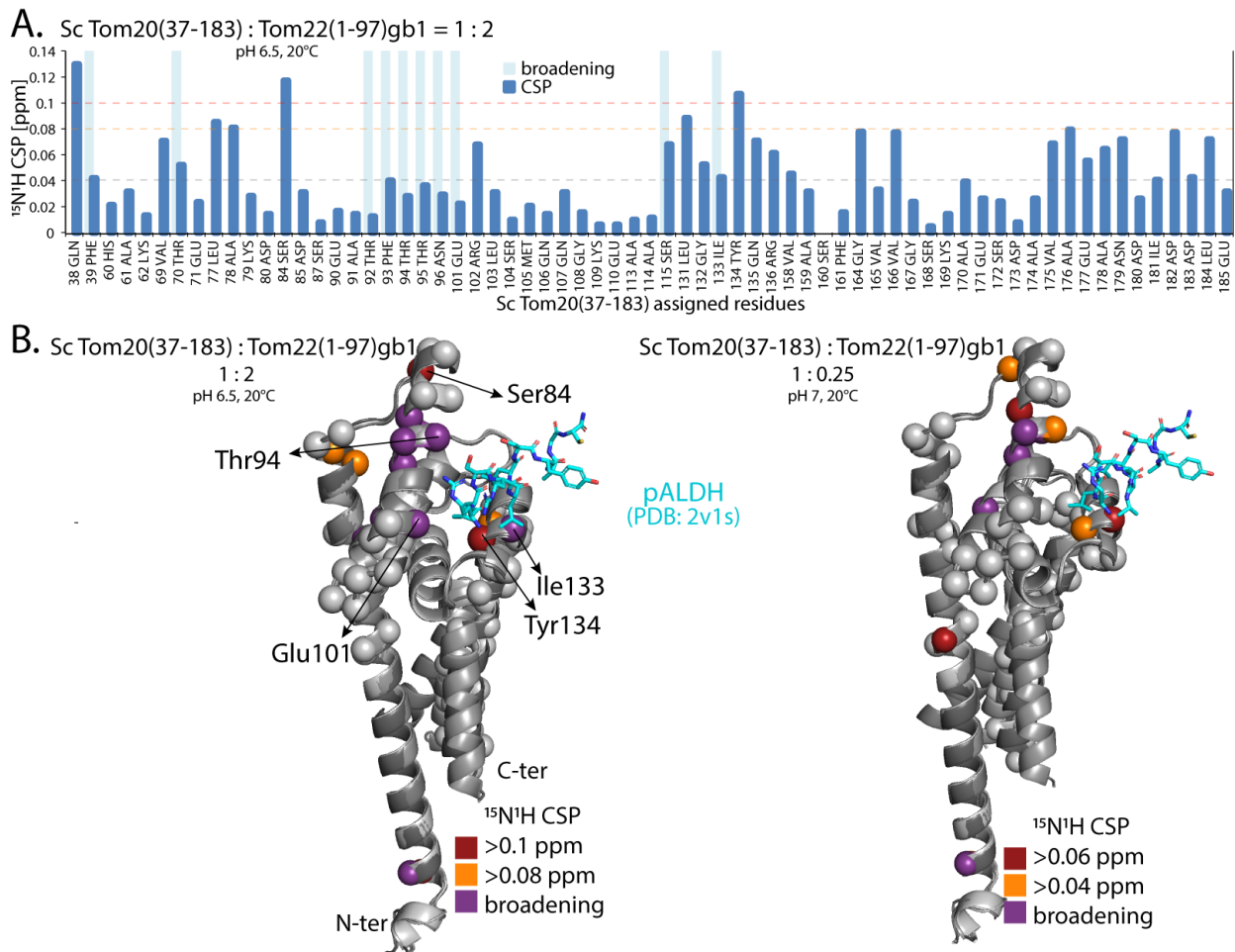


FIGURE 4.12: **Yeast Tom20 receptor binds the Tom22 in the presequence binding site.** **A.** Combined  $^{15}\text{N}/^1\text{H}$  CSP (blue) of the assigned yeast Tom20(37-183) residues upon adding two fold excess of Tom22(1-97). The residues for which a significant peak broadening caused by addition of Tom22 was observed are shown with light blue bars and the standard deviation over all CSP values (0.017 ppm) is indicated with horizontal dashed line - below this line the CSP is considered non-significant. Dashed red and orange horizontal lines indicate an arbitrary cut-off for CSPs shown on the ScTom20 model in the panel (C.) of this figure. **B.** Projection of the chemical shift deviations and line-width changes observed on five superimposed Robetta models. On the left hand side, Sc Tom20 residues with significant CSPs or peak broadening upon adding two fold excess of Tom22(1-97) are shown by colored spheres (see plot in panel (A.) of this figure). On the right hand side, same interaction in slightly different buffer conditions and protein ratio is mapped on Tom20 structural model. In all structural models of ScTom20 the peptide of pALDH presequence is shown (cyan) to indicate the presequence binding site observed for the rat Tom20 (Abe et al., 2000). In both conditions, the Sc Tom20 residues affected by the presence of Tom22 are mostly located in the confirmed presequence binding site.

As the ScTom20 assignment is incomplete, we again performed equivalent experiments with rat Tom20, assuming that the interaction patterns are conserved in the yeast and rat Tom receptors.

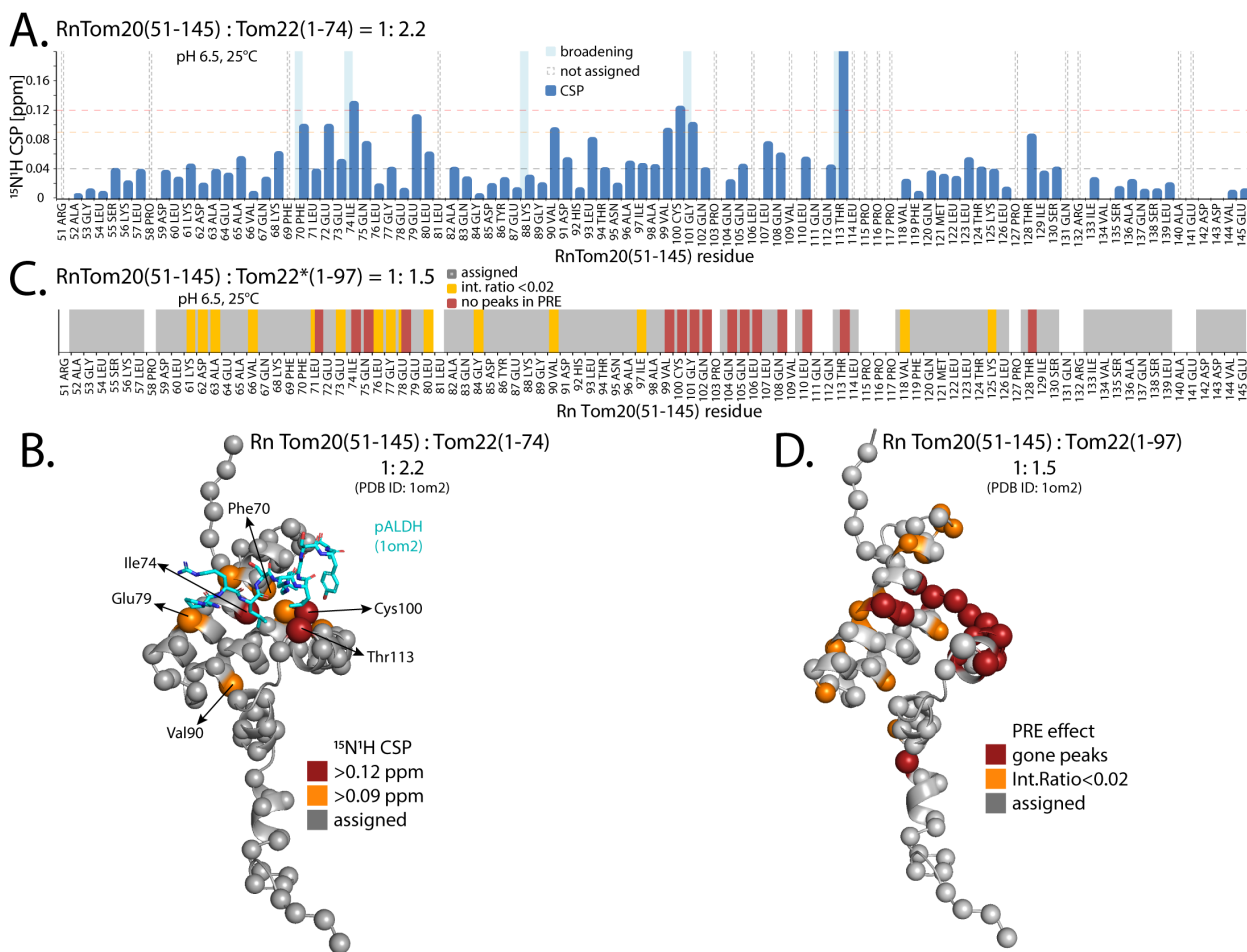


FIGURE 4.13: **Tom20 receptor from rat binds Tom22 in the presequence binding site.** **A.** Combined  $^{15}\text{N}^1\text{H}$  chemical shift perturbations (CSPs) of the RnTom20(51-145) residues upon adding 2.2 fold excess of the Tom22(1-74) (blue bars). The residues for which a significant peak broadening caused by addition of Tom22 was observed are shown with light blue bars. The standard deviation over all Tom22 induced CSPs (0.035 ppm) is indicated with gray horizontal dashed line. Dashed red and orange horizontal lines indicate an arbitrary cut-off for CSPs shown on the RnTom20 structure. **B.** RnTom20 residues with significant CSPs upon adding 2.2 fold excess of Tom22 mapped on the NMR structure of RnTom20 bound to the pALDH presequence peptide (Abe *et al.*; PDB:1om2). **C.** The relaxation enhancement effect (PRE) of the paramagnetically labeled Tom22(1-97) on the  $^{15}\text{N}^1\text{H}$  RnTom20(51-145) residues. The residues for which the NH peaks were not recovered by reducing the OXYL-1-NHS are labeled in red and the ones for which the significant peak broadening between the oxidized and reduced RnTom20-Tom22 sample was observed are shown in orange (see Methods7.14). **D.** Mapping the PRE affected residues induced by paramagnetically labeled Tom22 on the NMR structure of RnTom20 confirms the binding of the presequence peptides and the Tom22(1-74) in the same binding site of the rat Tom20 receptor.

Addition of a 2.2 fold excess of the unlabeled ScTom22(1-74) to the isotopically labeled RnTom20(51-145) caused CSPs on RnTom20 that are located in the presequence-binding site (Figure 4.13), with the largest effects observed for Ile74, Phe70, Thr113 and Cys100. Some of these residues are identical to the ones involved in binding the presequence while others are in close vicinity. To further confirm the binding site additional paramagnetic relaxation enhancement (PRE) experiments were performed. In such experiments, a paramagnetic tag is attached to a protein. The unpaired electron of the radical induces a distance-dependent increase of both transverse and longitudinal relaxation rates, that leads to peak broadening or even disappearance in spatial vicinity to the paramagnetic tag. Paramagnetically labelled ScTom22(1-97) was added to  $^{15}\text{N}^1\text{H}$ -RnTom20(51-145). Presence of the paramagnetic tag on Tom22 then leads to a decrease in the peak intensity for Tom20 residues that are in or close to the binding site. By mapping the RnTom20(51-145) residues for which significant peak broadening in the PRE experiment was observed on the structure of the

RnTom20, we confirmed the binding of the Tom22<sub>cyt</sub> in the presequence binding site (Figure 4.13C. and D.).

Altogether, our results show that the presequence binding site on the Tom20 receptor is the same for both yeast and rat protein and that it is comparable to the previously observed presequence binding site of the rat protein (Abe et al., 2000). Interestingly, Tom22 binds exactly into this presequence-binding site of both yeast and rat Tom20, suggesting that Tom22 competes with the presequences for the same binding site on Tom20. Comparing the residues of yeast and rat Tom20 involved in the presequence and Tom22 binding, we noticed that many of them correspond to conserved hydrophobic residues: Leu71, Ile74, Val99, Leu107, Leu110 and Thr113 of rat Tom20 and Thr94, Val97, Val123, Leu129, Leu131, Ile133 and Ser137 of the yeast Tom20. The multiple sequence alignment (Figure 4.5) shows that these positions are conserved in Tom20 proteins.

Having determined the Tom22 binding site on the cytosolic domain of Tom20 receptor and establishing that Tom22 occupies the presequences binding site, the next step in elucidating the possible role of Tom22 in presequence-holding precursors import is to determine the Tom20-binding region of Tom22 receptor.

#### **4.3.2 Tom22 employs the residues of the cytosolic $\alpha$ -helical region for binding to Tom20**

To study the interaction and binding of Tom20 on the cytosolic domain of Tom22 receptor, we used both the longer Tom22(1-97) and the shorter Tom22(1-74) constructs and the cytosolic domain of yeast Tom20(37-183) protein (Methods7.4.3; AppendixS1). For the Tom22-detected NMR experiments isotope-labeled  $^{15}\text{N}^1\text{H}$  Tom22 was mixed with unlabeled Tom20 sample at the indicated molar ratios.

The observed residue specific changes in peak positions and the severe peak broadening observed in the  $^{15}\text{N}^1\text{H}$  BEST-TROSY spectrum of Tom22 upon adding an equimolar quantity of yeast Tom20 indicated specific interaction between the two proteins (Figure 4.14).



$^{15}\text{N}$  Tom22(1-97)gb1 apo = 60  $\mu\text{M}$   
 $^{15}\text{N}$  Tom22(1-97)gb1 : ScTom20(wt) = 1 : 1

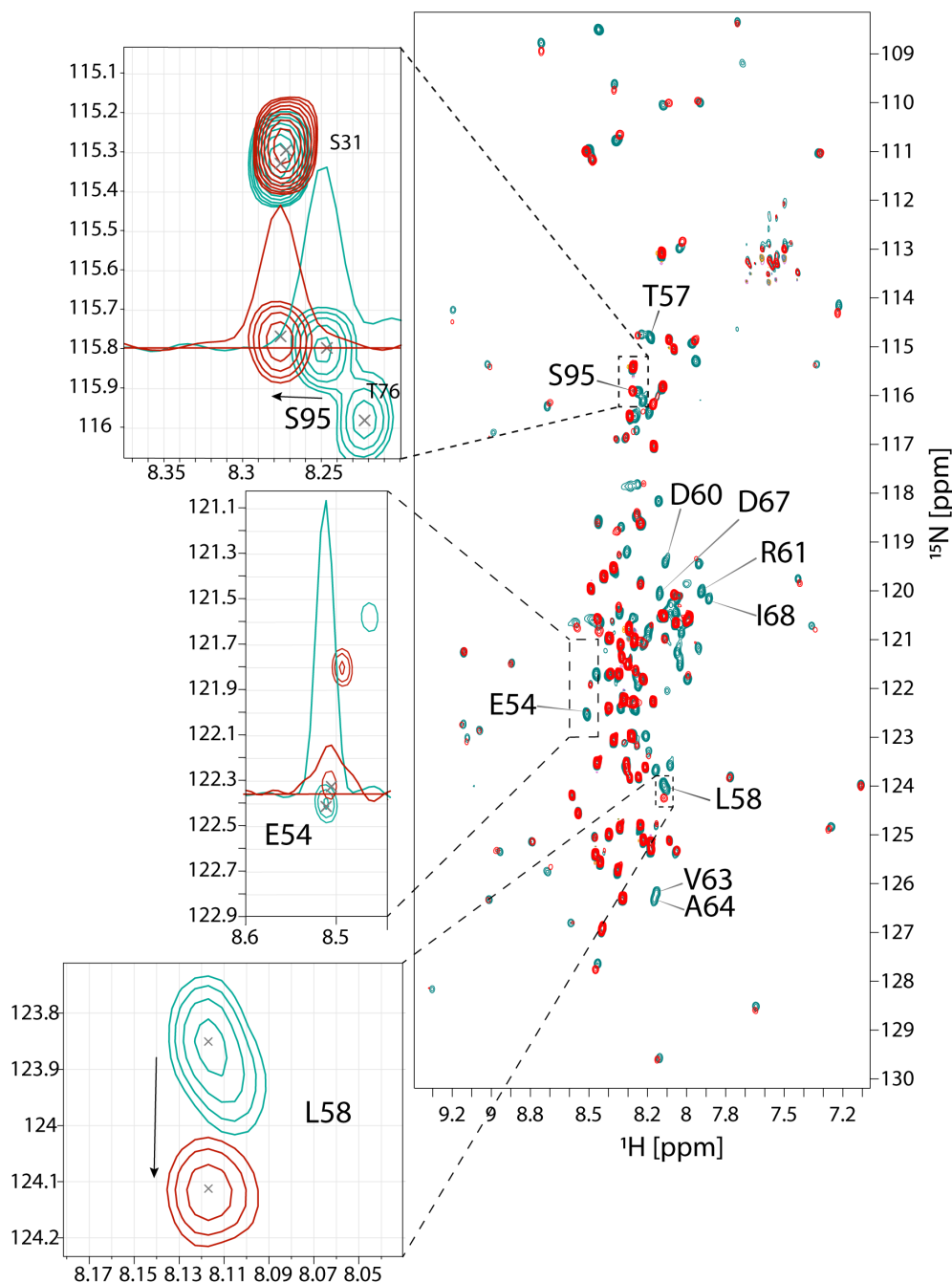


FIGURE 4.14: **Cytosolic domain of Tom22 interacts with the cytosolic domain of yeast Tom20.** The superposition of the  $^{15}\text{N}$  $^1\text{H}$  BEST-TROSY spectra of Tom22(1-97)GB1 construct alone (cyan spectrum) and in the presence of equimolar ScTom20(37-183) (red spectrum). Upon addition of the Tom20, changes in the peak positions and peak intensities could be observed, as illustrated for few selected isolated peaks that are enlarged and their 1D traces (signal intensities at exact nitrogen frequency) are shown. The more dispersed peaks, outside of the 7.8-8.7 ppm range, correspond to the GB1-solubility tag which is a folded protein domain with known chemical shifts.

The chemical shift perturbations and decrease of peak intensities observed in Tom22(1-97)GB1 spectra upon addition of the Tom20 protein (Figure 4.14), were plotted as a function of the Tom22 sequence (Figure 4.15A). We observed that residues that are affected by the presence of Tom20 are mostly located in the

region of the Tom22<sub>cyt</sub> that is experimentally shown to have an  $\alpha$ -helical secondary structure propensity (residues 56-68).

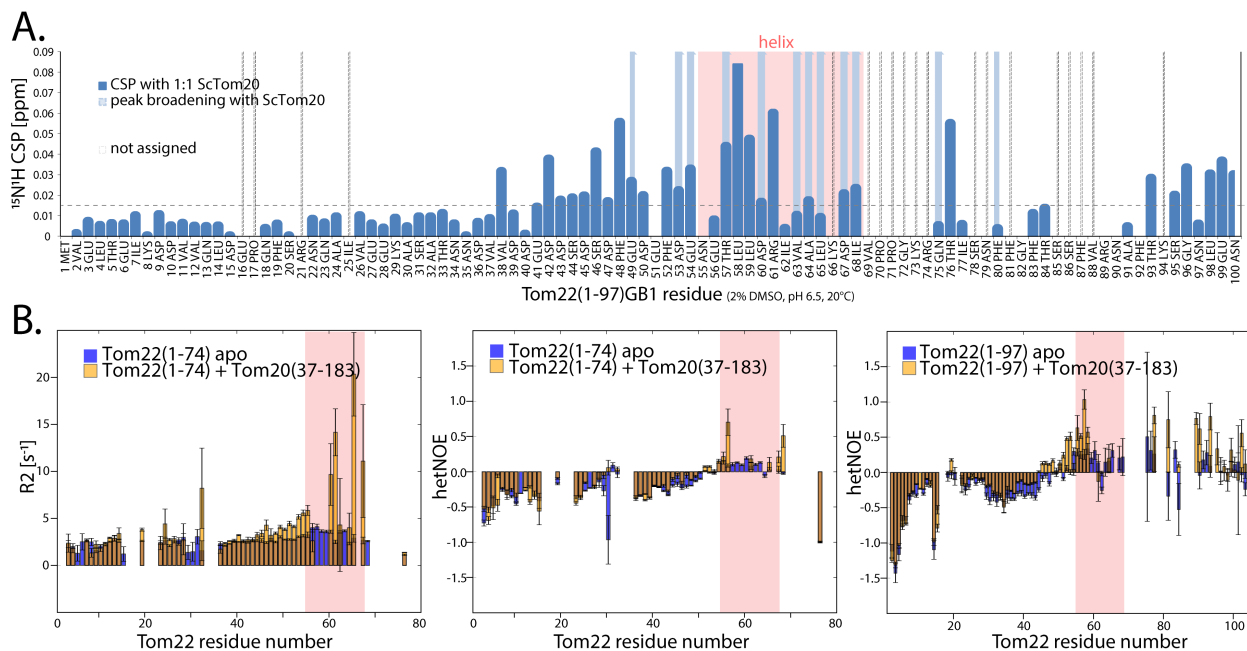


FIGURE 4.15: **Tom22 employs the residues of the  $\alpha$ -helical region for binding to Tom20.** **A.** Combined  $^{15}\text{N}/^1\text{H}$  chemical shift perturbations (CSP) of Tom22(1-97) upon addition of yeast Tom20 in equimolar ratio. The significant peak broadening caused by addition of the Tom20 is indicated in light blue color. The Tom22 residues that are not assigned, or for which no peaks were present in the apo-Tom22 sample, are marked with vertical, dashed lines in gray. The standard deviation over all CSP values (0.016 ppm) is indicated with horizontal dashed line - below this line the CSP is considered non-significant. Red box indicates Tom22 residues that are experimentally shown to have a propensity to form an  $\alpha$ -helix. **B.** Relaxation parameters of apo-Tom22 and of Tom22 in presence of Tom20 are shown with a blue and yellow bars, respectively. The panel shows from left to right: R2 relaxation rates, hetNEOs observed for the short and the long Tom22 constructs.

The relaxation properties of Tom22 in the presence of Tom20 were studied and compared to those of the apo-Tom22 with the goal of determining site-specific differences in backbone dynamics that would confirm the binding site (Figure 4.15B). Comparison of R2 relaxation rate constants in absence or presence of Tom20 (Figure 4.15B, left panel) show similar relaxation properties of the N-terminus of Tom22(1-74), while in the C-terminus, higher relaxation rates are observed in presence of Tom20. This indicates an increase in the apparent size of this region due to interaction with Tom20. The hetNOE values are plotted as a function of Tom22(1-74) (Figure 4.15B, middle panel) or Tom22(1-97) (Figure 4.15B, panel on the right) residue number. Negative hetNOE values for the residues in the N-terminus of Tom22 were observed for all samples indicating high flexibility that does not change upon addition of Tom20. The increase in the hetNOE values in the C-terminus upon addition of Tom20 indicates slower motion experienced by these residues in the presence of Tom20, compared to the apo Tom22.

Interestingly, both higher R<sub>2</sub> rates and hetNOE values in presence of Tom20 were observed for the residues comprising the small cytosolic  $\alpha$ -helix. Together, the NMR relaxation experiments indicate that the flexibility of the Tom22 N-terminus is independent of the presence of Tom20 while towards the C-terminus of Tom22<sub>cyt</sub>, in the  $\alpha$ -helical region, the flexibility decreases and the apparent size increases when Tom20 is present in the sample.

The relaxation parameters measured above suggest that, upon binding to Tom20, Tom22 adopts a more

stable structure. To assess whether the Tom20 interacting region of Tom22 (mainly residues 56-68) is subjected to secondary structure changes, we collected 3D NMR experiments that provided us with the H, N,  $C\alpha$ ,  $C\beta$ ,  $C_O$  chemical shifts of Tom22 in the Tom20-bound state. The observed chemical shifts report on the apparent state of Tom22 (i.e. population averaged state). However, due to severe peak broadening of Tom22 at saturating concentrations in the  $^{15}\text{N}^1\text{H}$  BEST-TROSY spectra we were not able to do a complete backbone assignment. Therefore, non-saturating conditions were chosen to assign as many resonances as possible from the 3D experiments. Taking into account the concentration of Tom22 in the sample and the calculated dissociation constant (Section 4.6, Table 4.1), a 1:0.75 molar ratio of Tom22:Tom20 was chosen. At this molar ratio  $\sim 55\%$  of Tom22 is in the bound state (binding curves simulated with the help of <https://protein-ligand-binding.herokuapp.com/>). We calculated the secondary  $C\alpha$  and  $C_O$  chemical shifts as the difference of the experimentally obtained chemical shift and the random coil chemical shift. The random coil chemical shifts were obtained by the POTENCI software (Prediction Of TEMperature, Neighbor and pH Corrected shifts for Intrinsically disordered proteins; Nielsen and Mulder, 2018) where the Tom22 sequence was given as an input. The secondary chemical shift is dependent on secondary structure. Both  $C\alpha$  and carbonyl carbons ( $C_O$ ) experience a downfield shift when they are located in helices and an upfield shift when they are located in  $\beta$ -strands (while the  $C\beta$  resonances experience an upfield shift when located in helices and are a bit less sensitive when located in  $\beta$ -strands). Indeed, the region of apo Tom22 for which a partially populated helix was found (see Section 4.1.1), shows significant downfield (positive)  $C\alpha$  and  $C_O$  secondary chemical shifts (residues 54-69, Figure 4.16, left panels). Comparing the  $C\alpha$  and  $C_O$  secondary chemical shifts of Tom22 alone and Tom22 bound to Tom20 a slight increase for the residues in predicted  $\alpha$ -helical region could be observed for Tom20-bound Tom22 whereas for the rest of the Tom22 sequence no such change is observed (Figure 4.16). Unfortunately, assignments could not be provided for all residues in this region in the Tom20-bound state.

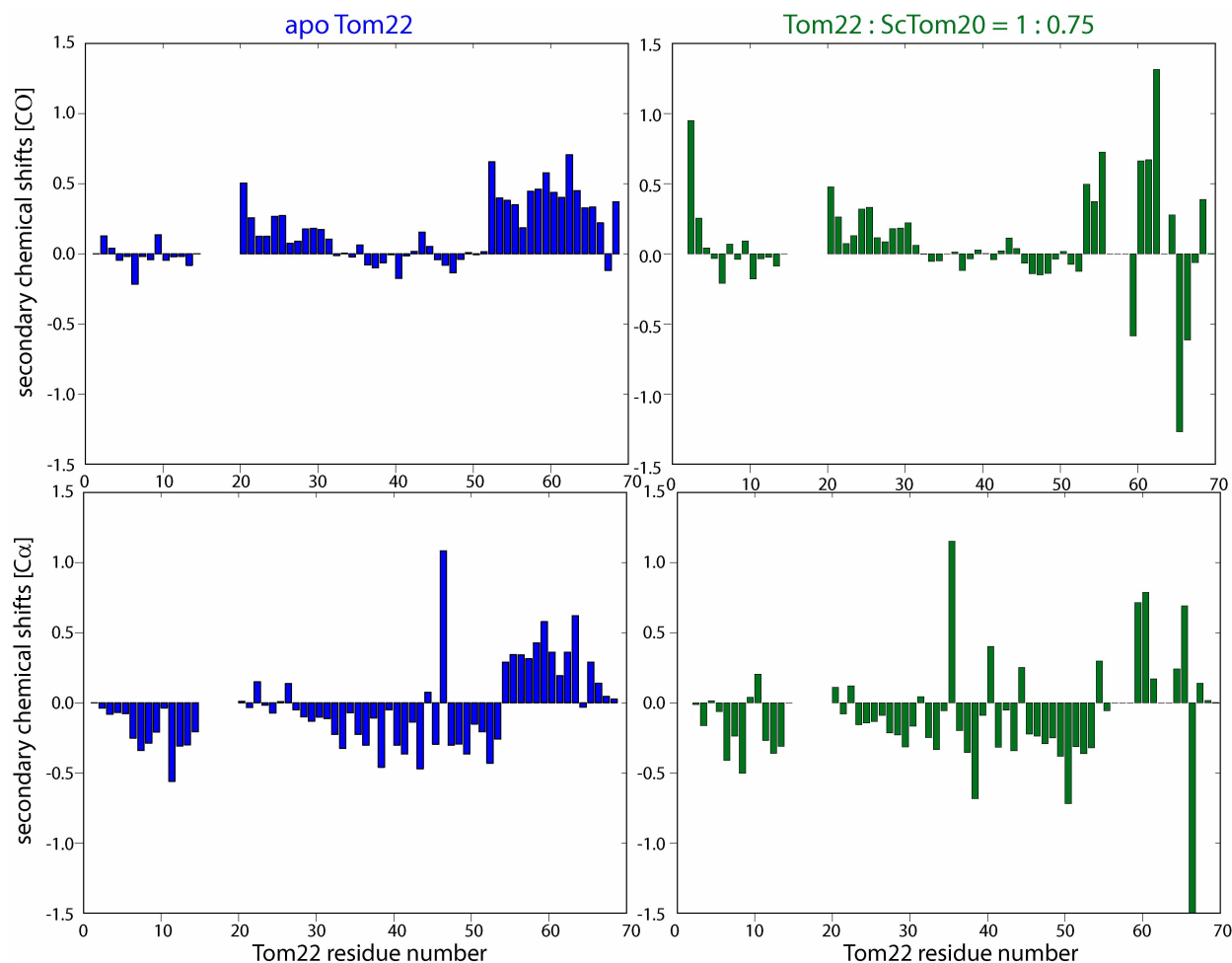


FIGURE 4.16: **Tom22 secondary chemical shifts in the presence or absence of Tom20.** Residue-wise  $C\alpha$  and  $C\beta$  secondary chemical shifts calculated as differences of the experimentally obtained chemical shifts and the random coil chemical shifts. The latter were obtained by POTENCI software (Prediction Of Temperature, Neighbor and pH Corrected shifts for Intrinsically disordered proteins; Nielsen and Mulder, 2018) with the Tom22 sequence as an input.

We initially used TALOS-N and CheSPI software to predict secondary structure propensities of the apo and Tom20-bound Tom22. Based on provided H, N,  $C\alpha$ ,  $C\beta$ ,  $C\gamma$ ,  $C\delta$  chemical shifts, TALOS-N software predicted an increase (of  $\sim 20\%$ ) in foldedness of residues 56-62 of Tom20-bound Tom22 (Figure S6 in Appendix B). For the Tom22 residues 62-69 the propensity for  $\beta$ -strand significantly increases. However, since the  $C\alpha$ ,  $C\beta$  and  $C\gamma$  resonances of residues 56, 57, 58 were not provided (due to peak broadening when interacting with Tom20) TALOS-N provides propensities based only on the sequence and its matches in the database. Therefore, from the TALOS-N propensity predictions we could not draw definite conclusion on the Tom22 secondary structure changes.

From the changes in the Tom22 peak positions and intensities, together with the changes in the relaxation properties of the C-terminus of Tom22 in presence of Tom20, we conclude that residues of the partially populated  $\alpha$ -helix (55-68) are involved in binding to the presequence binding site of the yeast Tom20 receptor. A slight increase in the secondary chemical shifts of  $C\alpha$  and carbonyl carbons observed for residues (55-65) of Tom22 upon addition of Tom20 indicate a possibly increased helical propensity of these residues when bound to Tom20.

### 4.3.3 Reduced binding of Tom22 with Tom20 variants

NMR analysis of the Tom20–Tom22 interaction revealed that conserved hydrophobic residues from Tom20 are in the presequence binding site that also bind to Tom22. The importance of these conserved hydrophobic residues in binding of the transient  $\alpha$ -helix of Tom22 was further studied by site-directed mutagenesis of selected Tom20 residues. We chose two amino acid exchanges in positions where Tom20 sequences carry a conserved hydrophobic residue (compare alignment in Figure 4.5), and replaced them by charged residues: (V97K) and (V97K, T94K). If the binding of Tom20 and Tom22 was mostly driven by hydrophobic interactions, we would expect reduced binding affinity of Tom22 to the variants. However, if Tom22-Tom20 binding was driven by electrostatic interactions, the more positively charged binding site of these Tom20 variants would increase the binding of the conserved residues of Tom22 forming the negatively charged side of the small cytosolic, amphiphilic helix (Figure 4.2 and Figure 4.3).

We probed the binding by NMR titration, using isotope-labeled Tom22 and unlabeled variants of Tom20 (V97K) and (V97K, T94K). We decided to detect Tom22 rather than yeast Tom20 due to the more complete assignment of Tom22 providing us with more probes to monitor the interaction. To compare the binding efficiency of the variants and the wild-type Tom20 on Tom22, NMR samples with the same concentration of Tom22 were prepared with equimolar amounts of either WT, V97K or V97K,T94K Tom20. In the NH spectra of Tom22 bound to Tom20 variants we observe smaller chemical shift perturbations and lower peak intensities compared to the ones caused by the wt Tom20 (Figure 4.17). The changes in the HN spectrum of Tom22 caused by Tom20 variants point to lower binding affinity. Lower binding affinity results in lower population of the bound state and since with NMR we are observing the population weighted average, under fast chemical exchange we see smaller perturbations compared to the chemical shifts expected for the 100% bound state. These observations confirmed that the chosen hydrophobic residues of Tom20 are involved in binding, and that hydrophobic interaction, presumably with the hydrophobic side of Tom22 helix, appears to be important for the Tom20-Tom22 binding.

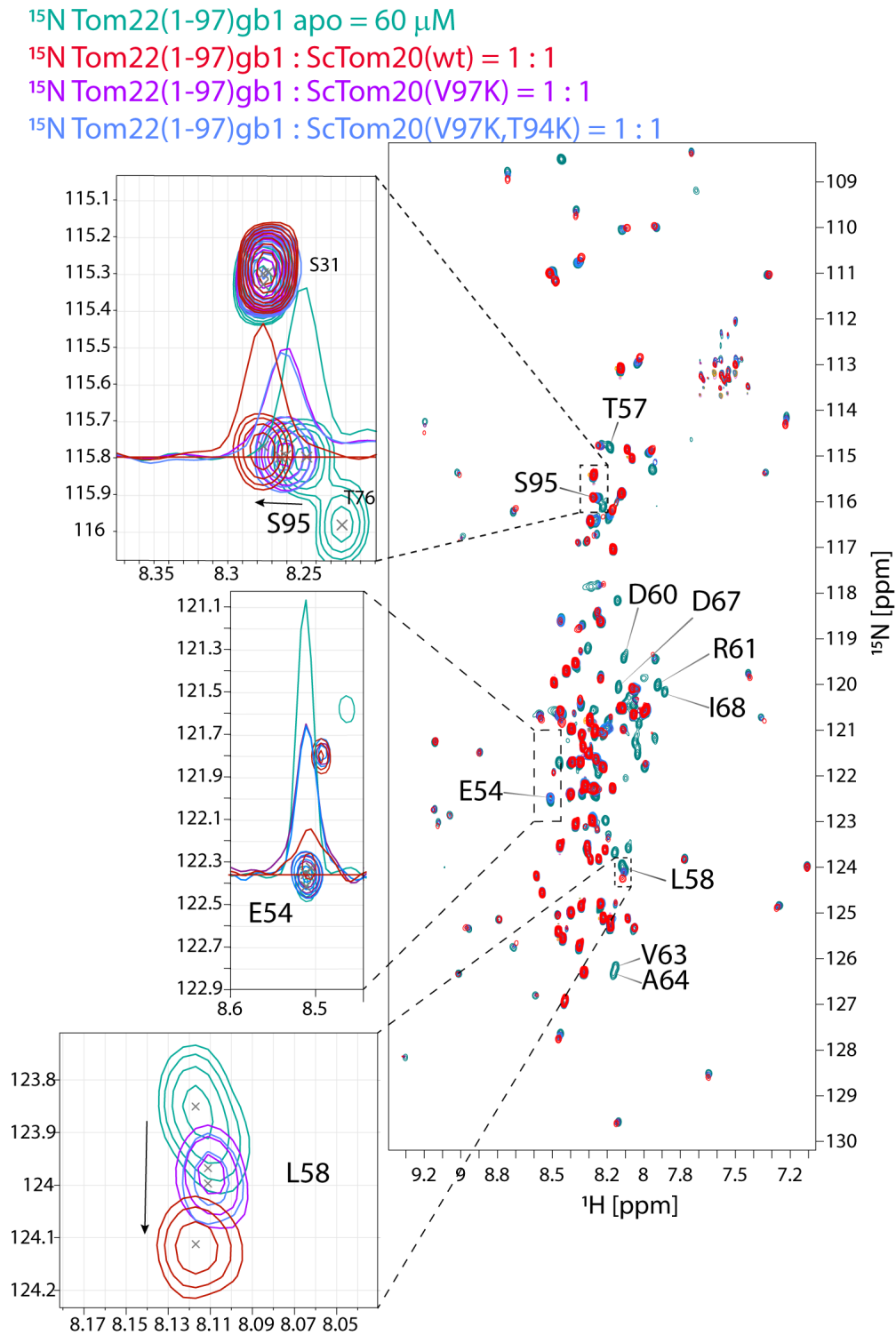


FIGURE 4.17: **Cytosolic domain of Tom22 interacts with the cytosolic domain of yeast Tom20.** The superposition of the  $^{15}\text{N}$  $^1\text{H}$  BEST-TROSY spectra of Tom22(1-97)GB1 construct alone (cyan spectrum) and in the presence of either wild-type ScTom20(37-183) (red spectrum), V97K ScTom20(37-183) (magenta) or V97K,T94K ScTom20(37-183) (cyan). Upon addition of the Tom20 variants, compared to the addition of the wild-type Tom20, lower peak intensities or smaller shifts in the peak positions could be observed. All samples were prepared with the same concentration of Tom22 and with equimolar Tom20. Few selected isolated peaks are enlarged and their 1D traces (signal intensities at exact nitrogen frequency) are shown. The more dispersed peaks, outside of the 7.8-8.7 ppm range, correspond to the GB1-solubility tag which is folded protein domain with known chemical shifts.

The NMR experiments showed weaker binding of Tom22 to the yeast Tom20 variants with less hydrophobic (more positively charged) presequence binding sites, confirming the importance of these hydrophobic residues for the Tom20-Tom22 interaction. Altogether, we observe binding of the transient  $\alpha$ -helix of Tom22 to the presequence binding site of the Tom20 receptor, where Tom20 residues in conserved hydrophobic positions are important for the binding.

#### 4.4 Testing the possibility of a ternary Tom20-Tom22-presequence complex formation

Our data indicate that the presequences, pSu9(1-69) and pSu9(1-25), do not interact directly with the cytosolic domain of Tom22 receptor (Fig.4.4). We hypothesized that the interaction between Tom22 and presequence may be stabilized by the presence of the Tom20 receptor. We tested whether a ternary-complex between presequence, Tom22 and Tom20 would be formed.

First, interaction of isotopically labeled Tom22 with unlabeled Tom20 and pSu9 were studied. Upon mixing of the  $^{15}\text{N}$ -Tom22(1-97) with ScTom20(37-183) at an equimolar ratio, we observed chemical shift perturbations and peak broadening as expected and described above. Unexpectedly, upon mixing  $^{15}\text{N}$ -Tom22(1-97) with ScTom20(37-183) and GB1-pSu9(1-69) at a 1:1:1 ratio, we observed additional changes in peak positions that already showed changes upon interacting with Tom20 only (Figure 4.18 left panel). With the concentration of Tom22 that was used in the sample (100  $\mu\text{M}$ ), with the known dissociation constant of ScTom20-Tom22(1-97) interaction of  $\sim 20 \mu\text{M}$ , and at the 1:1 molar ratio of Tom22(1-97):ScTom20, we can expect  $\sim 65\%$  of the complex to be in a bound state. Similarly, due to the dissociation constant in the same order of magnitude between the pSu9(1-69) and ScTom20,  $\sim 65\%$  of the pSu9-ScTom20 complex is in the bound state. Additional linear shifts in the peak position upon adding pSu9(1-69) could be explained by the increased population of the Tom22 bound-state. However, knowing that the presequences do not directly interact with the cytosolic domain of Tom22 receptor, these observations indicate increased affinity of the presequence-bound Tom20 towards the Tom22. The ternary complex formation, where the hydrophobic side of pSu9 would be bound in the presequence binding site of Tom20 and the positively charged side of pSu9 would bind to Tom22, does not seem plausible because then we would observe changes in the chemical shifts for different Tom22 residues instead of additive shifts of Tom20 interacting ones. However, performed experiment is not sensitive enough to detect possible transient interactions that might occur between the negatively charged side of amphiphilic helix of Tom22 and the negatively charged side of the presequence, mainly because of the peak broadenings at more saturated Tom22.

Similar behaviour, additive linear shifts of Tom22-binding residues of RnTom20, was observed when additionally pSu9(1-25) was added (see Figure 4.19). Interaction of pSu9 peptide and RnTom20 is  $\sim 60$  times weaker (NMR titration measured  $K_d \sim 1400 \mu\text{M}$ ; see Table4.1) compared to ScTom20 and the long presequence pSu9(1-69). At the 1:4 molar ratios of RnTom20-pSu9(1-25) we could expect  $\sim 20\%$  of RnTom20 in a pSu9 bound state. The changes observed in RnTom20 spectra correspond to changes in population of the bound state. This could be seen from the chemical shifts where the perturbation of the peak in Tom22-Tom20-pSu9 sample is linearly following the perturbation of Tom20-Tom22 and Tom20-pSu9 samples (peaks marked with black arrow in Fig.4.19). Perturbations in the chemical shifts of certain residues with only Tom22 added (blue arrows in Fig.4.19) or with only pSu9 added (green arrows in Fig.4.19) appears not to be linear, however it is a clear indication of the same binding site on Tom20. For these residues the peaks in Tom22-Tom20-pSu9 sample appears to be at the averaged peak position of the Tom20-pSu9 and Tom20-Tom22 peaks.

The bound residues appear to be mostly the same for both pSu9 and Tom22, with a few involved in binding of Tom22 only (shown in red squares in Figure 4.19). These few Tom22-only binding residues could be observed due to the lower affinity of Tom20 and pSu9 constructs used in this experiment resulting in only 20% of Tom20 being bound, hence the smaller effect observed by NMR, and/or due to the size of the Tom22 construct compared to the short peptide.



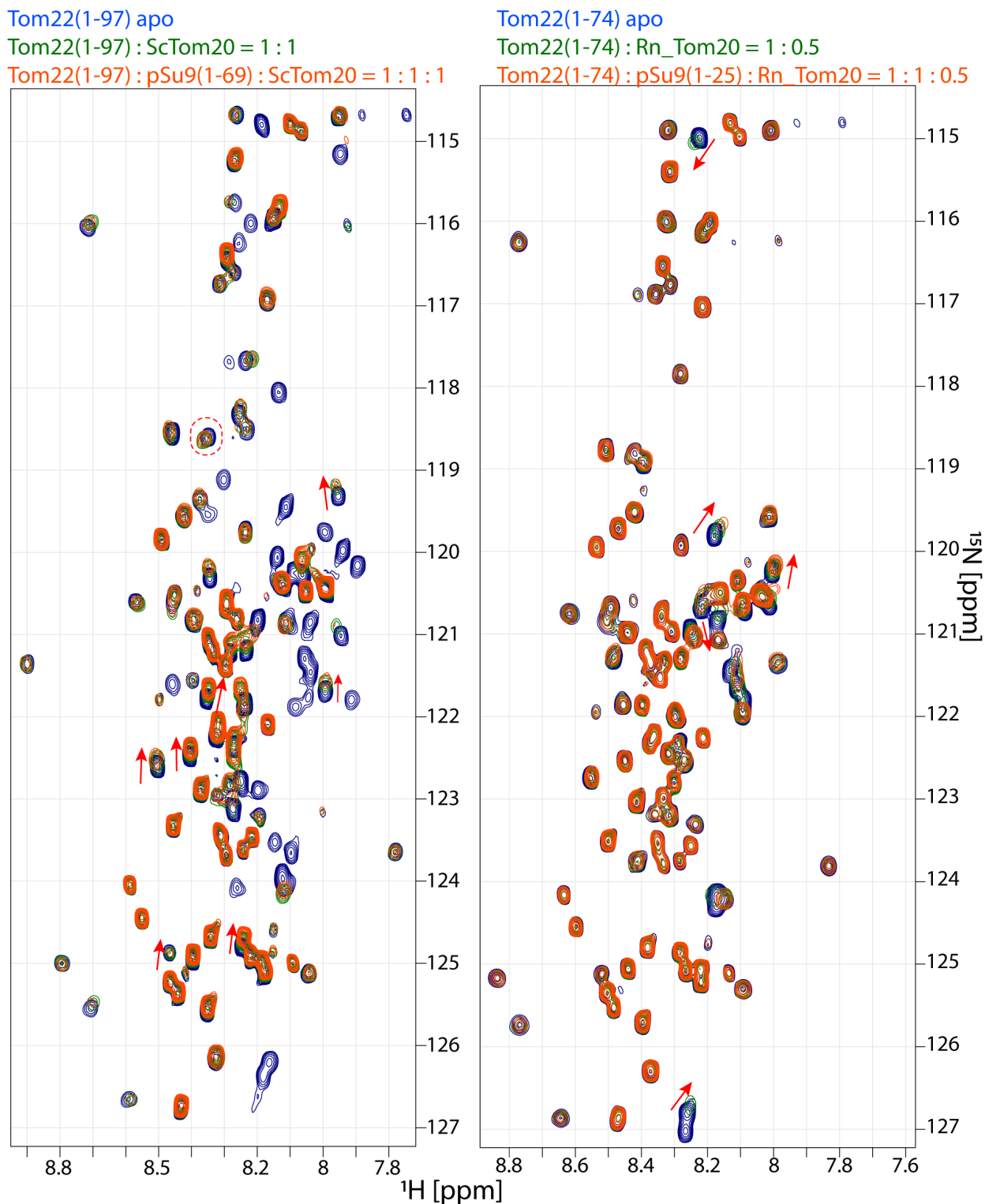


FIGURE 4.18: **Tom22 observed interaction with Tom20 and pSu9.** HN BEST-TROSY spectra of  $^{15}\text{N}$ -Tom22(1-74)GB1 (right hand side) and of  $^{15}\text{N}$ -Tom22(1-97)GB1 (left hand side). Blue spectra are showing the Tom22 protein alone. Spectra in green are showing Tom22 upon adding either ScTom20(37-183) (on the left side) or RnTom20 (on the right side), and spectra in orange are showing Tom22 upon adding both the presequence and Tom20 into the sample. Red arrows are indicating observed additional linear shift upon adding both Tom20 and pSu9 to the Tom22 sample, compared to when only Tom20 is added to the Tom22 sample.

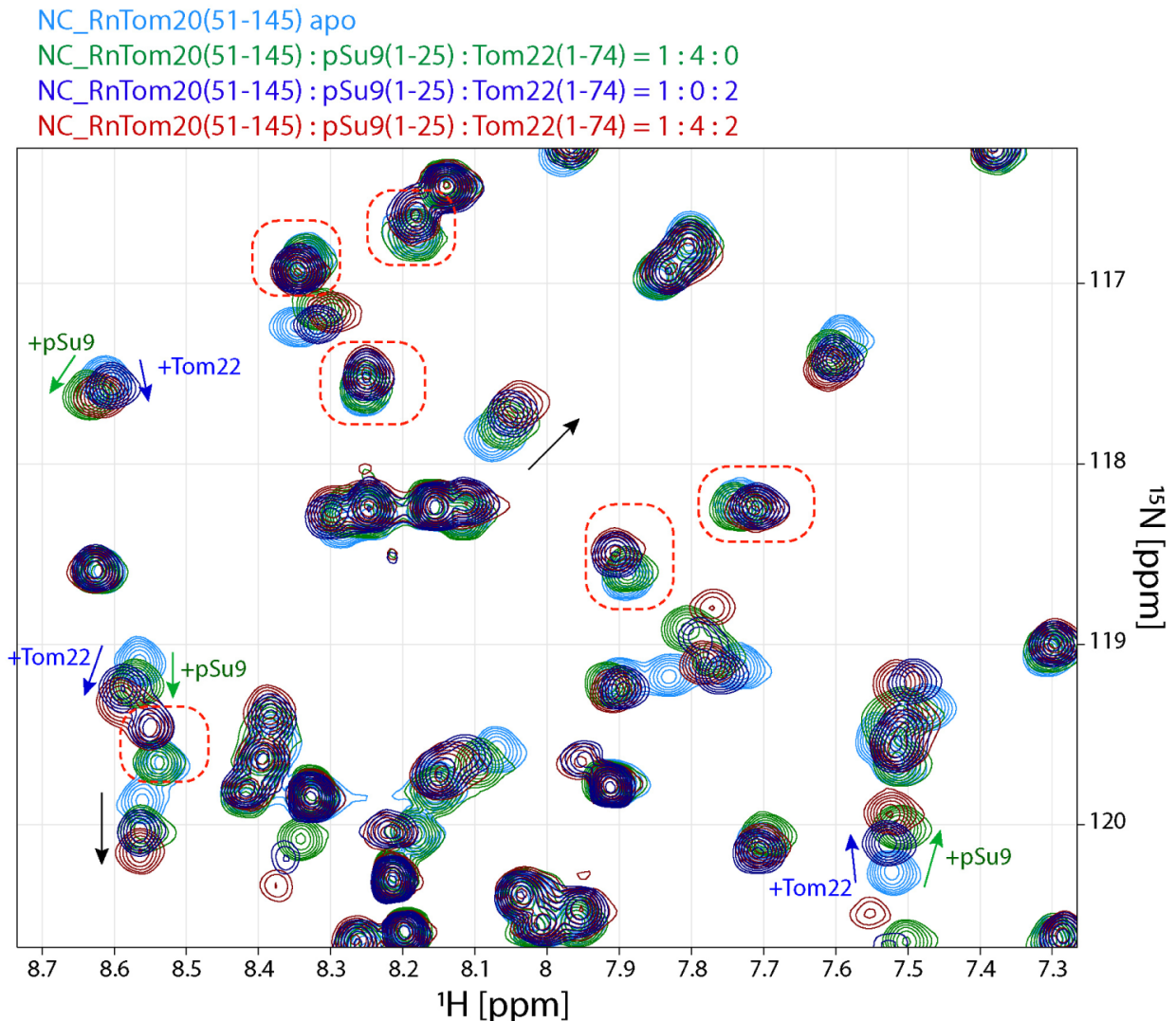


FIGURE 4.19: **RnTom20 observed interaction with Tom22 and pSu9.** HN BEST-TROSY spectra of  $^{15}\text{N}$ -RnTom20(51-145). Blue spectra are showing the RnTom20 protein alone. Spectra in green are showing RnTom20 upon adding four times excess of pSu9(1-25) peptide, spectra in purple are showing RnTom20 upon adding two times excess of Tom22(1-74) and finally spectra in red are showing RnTom20 upon adding both four times excess of pSu9(1-25) peptide and two times excess of Tom22(1-74). Blue arrows are indicating observed shift upon adding Tom22 alone and green arrows shifts of the same peaks but in a different direction upon adding pSu9 peptide alone. Black arrows indicate shifts induced by Tom22 and pSu9 that are linear and appear additive upon having three proteins in a sample. Peaks in red dashed boxes shift only upon adding Tom22 (dark blue spectrum), but not upon adding pSu9 (green spectrum) and show no additive effect with three proteins in a sample (brown spectrum).

Our experiments, observing by NMR isotopically labeled Tom22 or Tom20 upon adding one or two binding partners, did not result in changes that would indicate the presence of the ternary complex. Instead, our experiments confirm that only one set of residues, the cytosolic  $\alpha$ -helix of Tom22, is involved when either Tom20 or Tom20-pSu9 are present in the sample. Additionally, there is indication of increased affinity of the presequence-bound Tom20, compared to Tom20 alone, towards the Tom22<sub>cyt</sub>. The experiments observing Tom20 by NMR, upon adding Tom22 alone or Tom22 and the presequence, are not so straightforward to interpret due to ability of Tom20 to bind to both Tom22 and the presequence. However, (mostly) the same residues can be identified interacting with Tom22 and the presequence. The perturbations in the chemical shifts when both Tom22 and the presequence are added to Tom20 are reporting on the averaged state of Tom20, where one part is pSu9 bound and the other is Tom22 bound. Altogether, these observations indicate that pSu9 and Tom22 are competing for the same binding site on Tom20. However, further

experiments are required to confirm these observations. For example, titration of Tom22 in a sample with constant concentration of pSu9-Tom20 could provide better information on the changes in the affinity.

## 4.5 Interaction of Tom22 and Tom70 cytosolic receptor domains

### 4.5.1 Tom22<sub>cyt</sub> employs its small helix for binding to the Tom70 receptor

The third receptor of the TOM complex, Tom70, is loosely attached to the core complex, similarly to Tom20, and is not observed in the cryo-EM structures (Araiso et al., 2019; Tucker, 2019). The proposed docking role of the cytosolic domain of Tom22 for the peripheral receptors Tom20 and Tom70 (Van Wilpe et al., 1999; Yamano et al., 2008b) suggests that Tom22 can also bind the Tom70 receptor. However, no details of such an interaction are known.

First, we wanted to find out if there is an interaction between the cytosolic domains of Tom22 and Tom70 and to determine which residues of Tom22 are involved. Therefore, isotope-labeled Tom22 was mixed with unlabeled Tom70(39-617), which we recombinantly expressed, purified and biochemically characterized (see AppendixB.0.1: Fig.S9). Analysis of the 2D  $^{15}\text{N}^1\text{H}$  BTROSY-spectra of the Tom22(1-74) alone and in presence of Tom70 revealed chemical-shift perturbations and a decrease in peak intensities (Figure S7). It is worth noting that the observed peak broadening in the Tom70-Tom22 sample was less drastic compared to the Tom20-Tom22 sample at the same molar ratios and concentrations. This observation indicates a weaker binding affinity and/or different binding on/off rate constants. Indeed, the dissociation constant of Tom70-Tom22 interaction ( $\sim 180 \mu\text{M}$ ; determined from NMR titration data, see Table 4.1) is about one order of magnitude lower than the one of Tom20-Tom22 interaction (with  $K_D \sim 20 \mu\text{M}$ ).

Mapping the CSPs and intensity decrease in Tom22(1-74) caused by presence of Tom70(39-617) onto the sequence indicates that Tom22 employs mostly the residues of the helical region (residues 56-68) and the upstream negatively charged region when interacting with Tom70 (Figure 4.20A).

To assess whether the Tom70-interacting region of Tom22 (mainly residues 56-68) is subject to changes in secondary structure we collected 3D NMR experiments that provided us with the H, N,  $C\alpha$ ,  $C\beta$ , carbonyl ( $C_O$ ) chemical shifts of the Tom22 residues in the Tom70-bound state. The observed chemical shifts report on the population-weighted average over the conformations, including the ratio of bound vs free Tom22. We observed that when attempting to saturate Tom22 with Tom70, extensive line broadening of the residues involved in binding makes it impossible to retrieve the chemical shifts in particular in 3D experiments. Therefore, we chose a molar ratio of 1:1.5 (Tom22:Tom70), which, taking into account the concentrations and dissociation constant (Section 4.6, Table4.1) results in ratio  $\sim 50\%$  of Tom22 in the bound state (binding curves simulated with the help of <https://protein-ligand-binding.herokuapp.com/>).

Comparing the  $C\alpha$  and  $C_O$  secondary chemical shifts of Tom22 alone and Tom22 bound to Tom70 shows a slight increase in the  $C\alpha$  and  $C_O$  secondary chemical shifts for the residues in the small cytosolic  $\alpha$ -helical region whereas for the rest of the Tom22 sequence no such change is observed (Figure 4.20). Analysis of the chemical shifts with TALOS-N and CheSPI, shown in (Figure S8) is qualitatively similar. Thus, binding to Tom70 stabilises the  $\alpha$ -helix of Tom22.

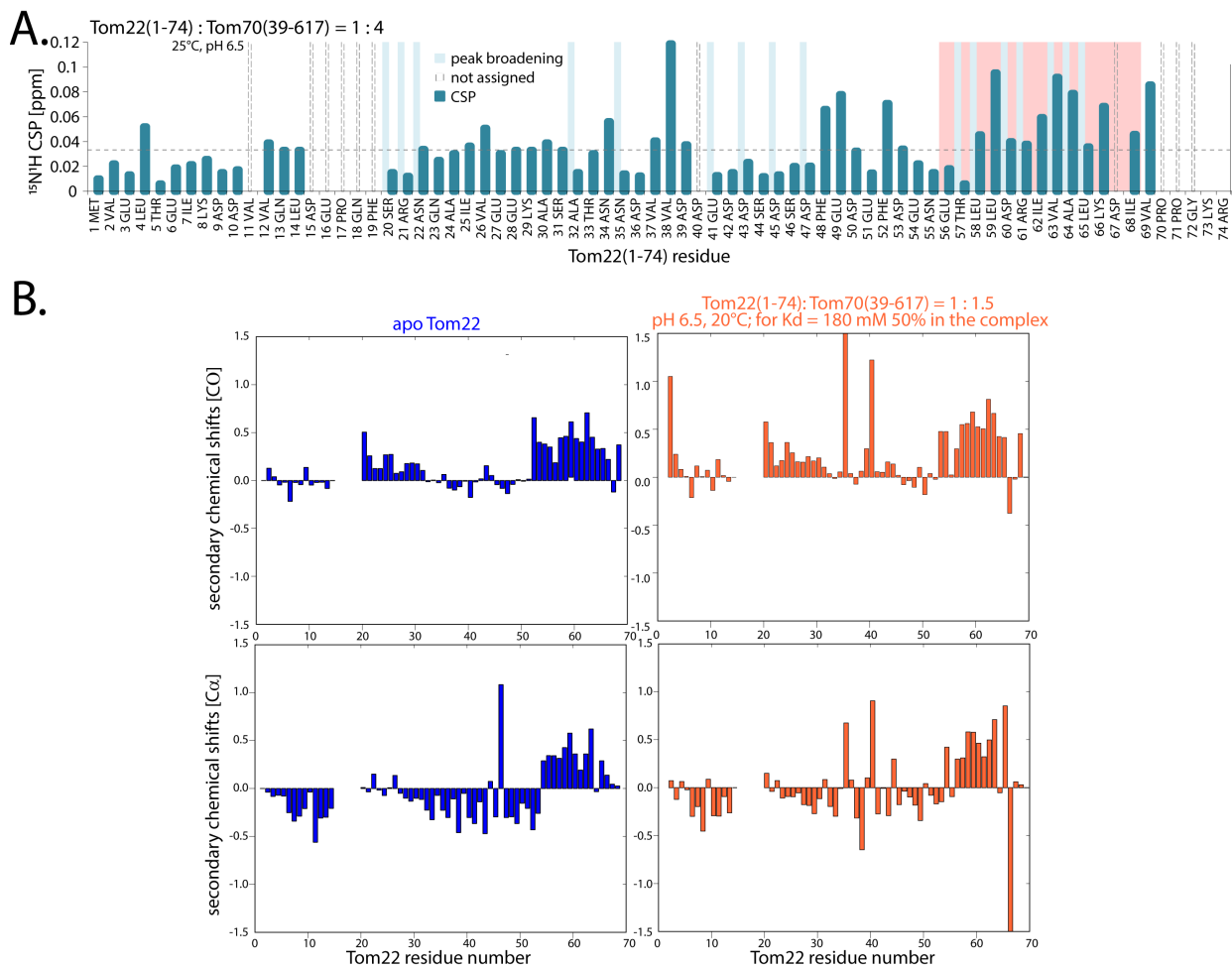


FIGURE 4.20: **Binding of the cytosolic domain of Tom22 on the Tom70 receptor.** **A.** Combined  $^{15}\text{N}^1\text{H}$  chemical shift perturbations (CSPs) of Tom22(1-74) upon adding four time excess of Tom70(39-617). The residues for which a significant peak broadening was observed upon adding the Tom70 are shown with light blue bars. The Tom22 residues that are not assigned are marked with vertical, dashed lines in gray. The standard deviation over all CSPs is indicated with horizontal dashed line - below this value the CSP is considered non-significant. **B.** Residue-wise  $\text{C}\alpha$  and  $\text{C}\text{O}$  secondary chemical shifts calculated as differences of the experimentally obtained chemical shifts and the random coil chemical shifts. The chemical shifts used as random coil chemical shifts were obtained by POTENCI software (Nielsen and Mulder, 2018) with Tom22 sequence as an input.

#### 4.5.2 Cytosolic helix of Tom22 binds to the putative client binding site of Tom70 receptor

Having determined that the cytosolic domain of Tom22 employs mostly the small cytosolic helix (residues 56-68) when interacting with the Tom70 receptor, similarly to when interacting with the Tom20 receptor, we wanted to probe where Tom22 binds on Tom70. The Tom70 client binding site is not known. Currently, the only information comes from structures of human Tom70 with the viral protein Orf9b from SARS-CoV2. The binding site of the internal matrix targeting signals was proposed based on the Tom70 structure and conservation of the residues in the C-terminal cleft (Wu and Sha, 2006) whereas the presequence binding was proposed to involve distinct residues however still in the C-terminal binding cleft (Melin et al., 2015). The structural models in Figure 1.4 illustrate these two proposed binding sites. Using the same protein construct and an identical approach, we wanted to compare binding of Tom22 and precursor proteins to Tom70.

Site-resolved NMR studies of proteins as large as Tom70 is a significant challenge due to the large number of resonances, the linewidth and, consequently, difficulties in assigning them to individual atoms. A BEST-TROSY spectrum of  $^2\text{H}$ ,  $^{15}\text{N}$ -labeled Tom70, shown in Figure S10A., shows a rather wide peak dispersion, as expected for a folded helical protein; attempts to obtain 3D experiments for assigning the backbone atoms suffered from low sensitivity. Deuteration and specific  $^{13}\text{C}$ ,  $^1\text{H}_3$  labeling of methyl groups, together with a methyl-TROSY pulse sequence (Ollerenshaw, Tugarinov, and Kay, 2003), is the most sensitive way of studying such large proteins. We have prepared samples with methyl labeling at either the Ala, Ile ( $\delta^1$ ), Val ( $\gamma^1$ ), Met or Thr positions, and finally chose to work with Met, Thr and Ile ( $\delta^1$ ) labeling. We have obtained the assignment of the Met, Thr, Ile-labeled Tom70 by preparing 59 point mutants (see Appendix B.0.1; Fig S11). Although this mutagenesis-based approach is labor-intensive, it has been successful in providing assignments of 100 % of Met, 87 % of Ile ( $\delta^1$ ) and 77% of Thr methyl groups (Figure S12).

The changes in the chemical environment of these specifically labeled methyl groups upon adding approximately two time excess of unlabeled Tom22(1-74) resulted in the changes of the peak positions or peak intensities (Figure 4.21).

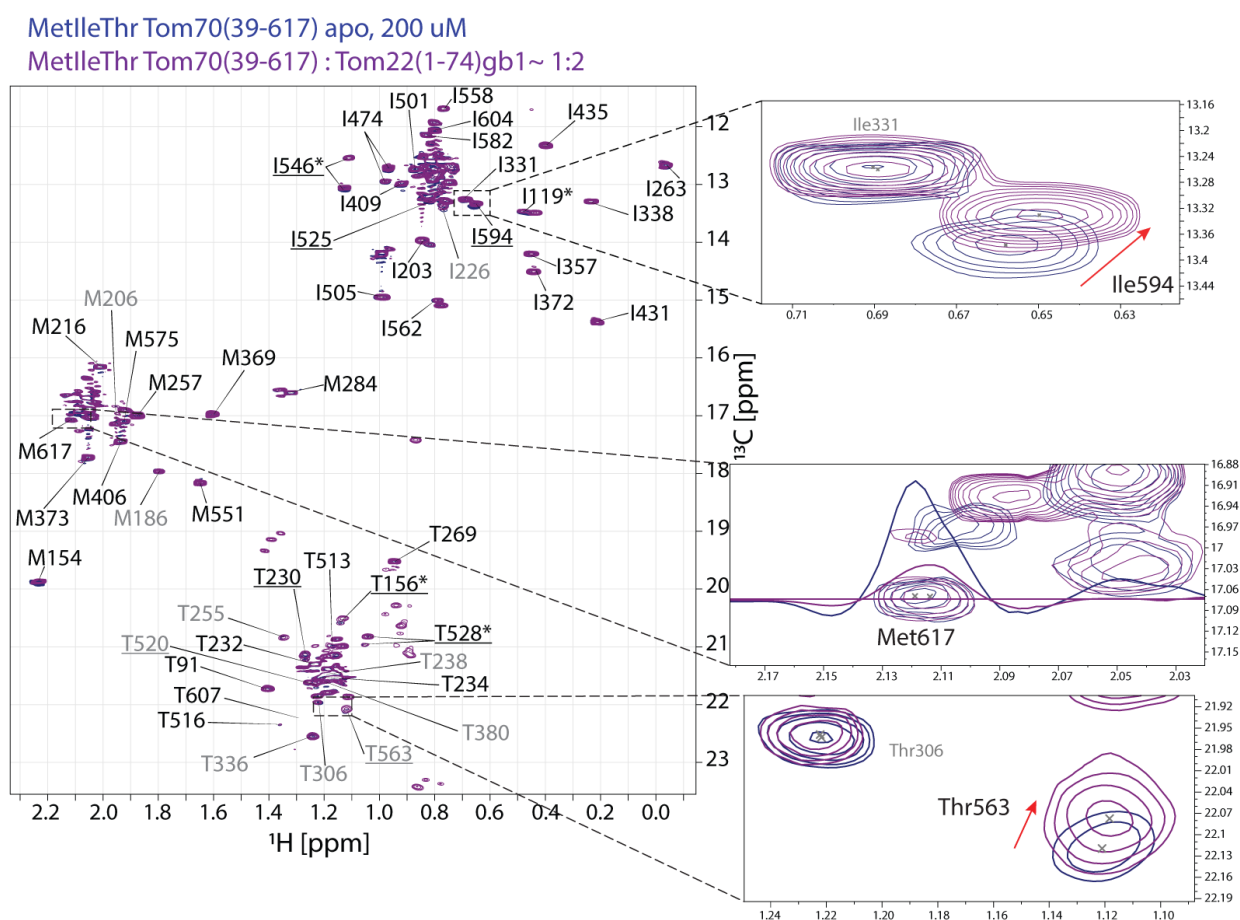


FIGURE 4.21: Changes in the  $^{13}\text{C}^1\text{H}$  HMQC spectrum of Ile $^{\delta^1}$  Met $^{\epsilon}$  Thr $^{\gamma}$  methyl groups of the Tom70(39-617) upon adding cytosolic domain of Tom22. Methyl SOFAST-HMQC spectra overlay of specifically labeled Ile $^{\delta^1}$  Met $^{\epsilon}$  Thr $^{\gamma}$  methyl groups of the Tom70(39-617) alone (in blue) and upon adding Tom22(1-74)GB1 (in purple). Some of the methyl groups for which significant changes in the peak position (Ile594 and Thr563) or intensity (Met617) was observed upon interaction with Tom22 are showed in enlarged areas.

Combined  $^{13}\text{C}^1\text{H}$  chemical shift perturbations of Met, Ile and Thr residues of Tom70 involved in the Tom22 binding were plotted residue wise (Figure 4.22A.) and mapped on the Tom70(39-617)–Tom22(1-74) protein

complex model (Figure 4.22B). The protein complex models were obtained using AlphaFold-Multimer-v2 (Evans et al., 2022). As input, protein sequences Tom70(39-617) and Tom22(1-74) were used, and the run was set without a user-provided template and using Amber (post-prediction relaxation of the structure by gradient descent in the Amber32 force field). The output was 5 structural models of a Tom70(39-617)–Tom22(1-74) complex. Interestingly, AlphaFold predicted the partially populated helix in Tom22 which is found in the interface in the model structures. Three out of five protein complexes show Tom22 bound in the lower part of the cleft forming the putative client binding site. This lower binding site comprises the residues proposed to be involved in the presequence binding based on *in vivo* import studies in yeast and cross-linking experiments (Melin et al., 2015). In the other two structural models Tom22 is bound in the upper part of the C-terminal cleft which matches the previously proposed client binding site proposed from the conservation and the crystal structure (Wu and Sha, 2006). From the results obtained by AlphaFold we cannot conclude where exactly Tom22 binds to the Tom70. However, by mapping the experimentally observed CSP and peak broadenings of specifically labeled Met, Ile and Thr of Tom70 upon adding Tom22 on the obtained structural models we can observe that most of the Tom22-interacting residues of Tom70 are located near the lower binding site on Tom70 (Figure 4.22B.).



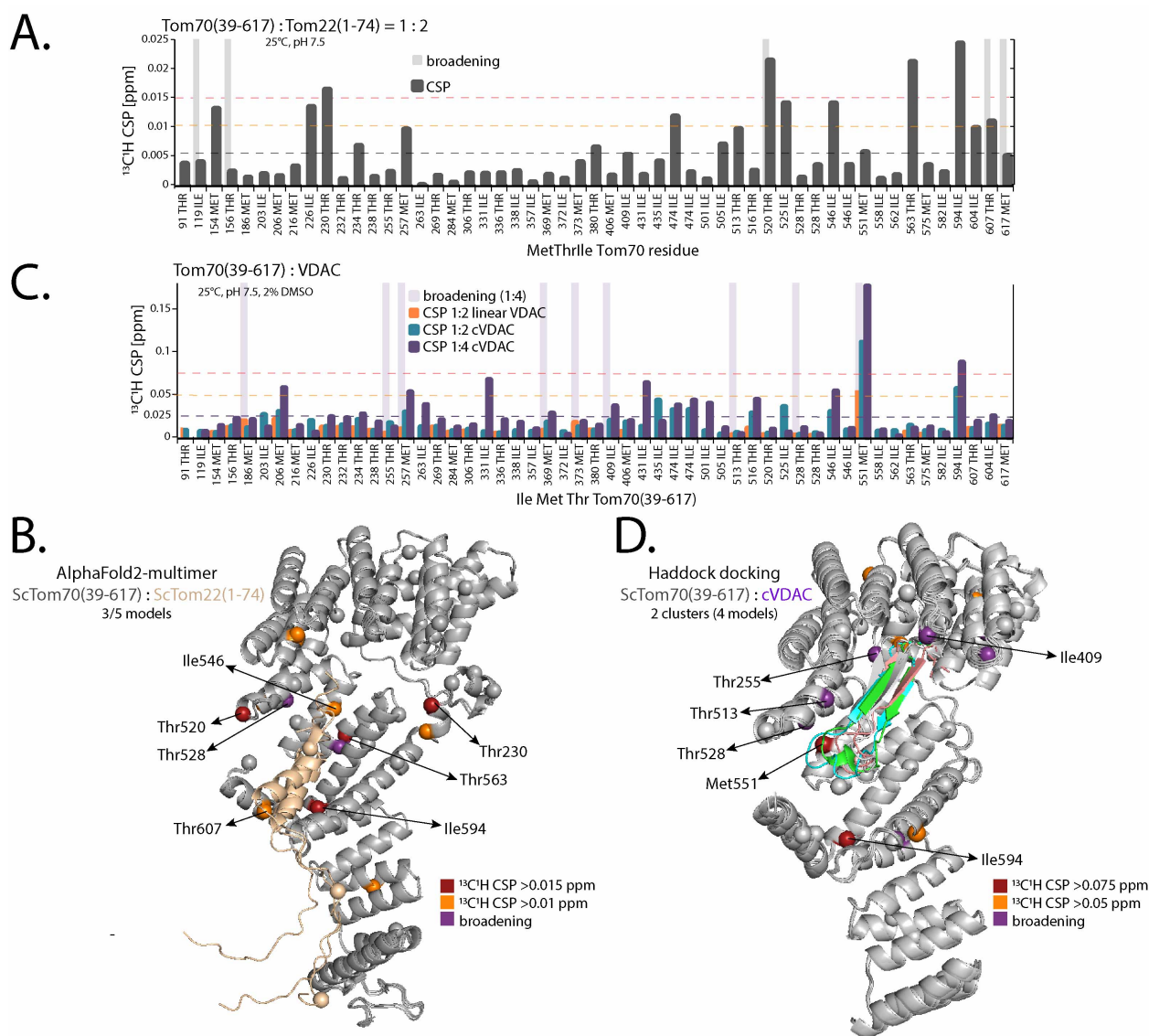


FIGURE 4.22: **Binding site of the cytosolic domain of Tom22 on the Tom70 receptor.** **A.** Combined <sup>13</sup>C<sup>1</sup>H chemical shift perturbations of Ile<sup>δ1</sup> Met<sup>ε</sup> and Thr<sup>γ</sup> methyl groups of Tom70(39-617) upon adding two time molar excess of Tom22(1-74). The residues for which a significant peak broadening was observed upon adding Tom22 are indicated with light gray bars. The standard deviation over all CSPs induced by adding Tom22 (0.058 ppm) is indicated with a gray horizontal dashed line - below this value the CSPs are considered non-significant. Dashed red and orange horizontal lines indicate an arbitrary cut-off for Tom22 induced CSPs that are mapped on the Tom70(39-617) structural model in panel (B.) of this figure. **B.** The Tom70(39-617)–Tom22(1-74) protein complex predicted using AlphaFold-Multimer. Tom70 residues showing significant CSPs (red and orange) or peak broadening (purple) upon adding Tom22 are shown as spheres. The Tom22 protein is shown in wheat color and the AlphaFold predicted helix matches the experimentally shown helical region of Tom22. Residues shown as gray spheres are assigned methyl groups of Tom70, which are not affected by the Tom22 in the sample. **C.** Combined <sup>13</sup>C<sup>1</sup>H CSP of Ile<sup>δ1</sup> Met<sup>ε</sup> and Thr<sup>γ</sup> methyl groups of Tom70(39-617) upon adding either a two time molar excess of linear VDAC<sub>257-279</sub> peptide (orange bar plot), a two time molar excess of cyclic VDAC<sub>257-279</sub> peptide (cyan/blue bar plot) or a four time molar excess of cyclic VDAC<sub>257-279</sub> peptide (purple bar plot). The residues for which a significant peak broadening was observed upon adding four time molar excess of cyclic VDAC<sub>257-279</sub> peptide are indicated with light purple bars. The standard deviation over all CSPs induced by 4x excess of cVDAC<sub>257-279</sub> is indicated with a gray horizontal dashed line (0.027 ppm) and the dashed red and orange horizontal lines indicate an arbitrary cut-off CSPs induced by 4x molar excess of cVDAC<sub>257-279</sub> mapped on the Tom70(39-617) structural model in panel (D.) of this figure. **D.** Structure of a protein–peptide complex of Tom70(39-617)–cVDAC<sub>257-279</sub> obtained by protein modeling software HADDOCK2.4. The Tom70 structural model that was uploaded to HADDOCK was obtained by SWISS-MODEL (<https://swissmodel.expasy.org/>) where a Tom71 (PDB: 3LCA) structure was used as a template. The CSPs on Tom70 caused by addition of cVDAC (panel (C.) of this figure) were used as docking constraints. The 2.4 HADDOCK version enables cyclisation and docking of a peptide using the peptide sequence (Charitou, Van Keulen, and Bonvin, 2022).

Further on, we tested the binding of a peptide fragment of the voltage-dependent anion channel (VDAC) comprising the two C-terminal  $\beta$ -strands, VDAC<sub>257-279</sub> to the cytosolic domain of Tom70 receptor. Previous studies demonstrated that this specific peptide, when in a cyclic form, forms a  $\beta$ -hairpin element with one highly hydrophobic side that is a mitochondrial targeting signal recognized by the Tom20 receptor (Jores et al., 2016). However in the same study the formation of the cVDAC peptide and Tom70 cross-linking adducts was shown. The multi-transmembrane carrier proteins are the biological substrates/clients of Tom70. The internal matrix targeting elements scattered over the primary structure of these precursors are assumed to be specific for the recognition by the Tom70 receptor (Wiedemann, Pfanner, and Ryan, 2001; Backes et al., 2018). However, due to the partially overlapping function of the Tom receptors shown by *in vivo* and *in organello* import experiments we decided to test the binding of this 'not primary' client of the Tom70 receptor.

To test the binding of cyclic VDAC<sub>257-279</sub> and Tom70(39-617) by NMR we prepared [U-<sup>2</sup>H]-[<sup>13</sup>C<sup>1</sup>H<sub>3</sub>-Ile<sup>1</sup>Met <sup>$\epsilon$</sup> Thr <sup>$\gamma$</sup> ]-labeled Tom70(39-617) and mixed it with a cVDAC peptide (see Methods 7.5). Addition of the linear VDAC<sub>257-279</sub> peptide was also tested; in previous interaction studies it was found that the cyclic peptide interacts much more strongly than does the linear one. Chemical-shift perturbations of <sup>13</sup>C<sup>1</sup>H<sub>3</sub>-Ile <sup>$\delta$ 1</sup>Met <sup>$\epsilon$</sup> Thr <sup>$\gamma$</sup>  methyl groups of Tom70(39-617) induced by adding either two-fold excess of the linear peptide, two-fold excess of the cyclic peptide or four-fold excess of the cyclic peptide were plotted residue wise (Figure 4.22C.) resulting in a different pattern compared to the plotted CSPs induced by Tom22. The chemical shift perturbations caused by addition of linear VDAC peptide (orange bar plot) could be detected, but they are approximately half the size of the CSPs caused by addition of the cyclic peptide at the same concentrations (cyan bar plot). The chemical shift perturbations caused by adding four time excess of the cyclic peptide were mapped on the structural model of cVDAC-Tom70 complex obtained with the HADDOCK software (Figure 4.22D.).

HADDOCK docking constraints used in our analysis were the chemical shift perturbations on Tom70 caused by addition of four time excess of cVDAC. The structural model of Tom70 with mapped cVDAC induced chemical shift perturbations and HADDOCK-assisted docked cVDAC, indicates that the peptide of a  $\beta$ -barrel precursor protein binds in the upper region of a previously suggested binding site in the C-terminal cleft of the cytosolic Tom70 domain (Wu and Sha, 2006).

Our results show that a cyclic peptide fragment of the voltage-dependent anion channel (VDAC) binds in the conserved hydrophobic binding pocket on the cytosolic domain of Tom70 receptor. This binding pocket was previously proposed for interaction with the internal matrix targeting signals of the carrier-precursor proteins (Wu and Sha, 2006). Surprisingly, we experimentally observed binding of Tom22 on Tom70 in the lower binding site, in the same C-terminal cleft. Residues of this lower binding site were shown to be implied in Tom70-dependant import of the presequences (Melin et al., 2015).

Both the upper (or the conserved (Wu and Sha, 2006)) and the lower (or presequence (Melin et al., 2015)) binding sites, occupy the same C-terminal pocket of Tom70 (Figure 1.4). Interestingly, the recent cryo-EM and crystal structures of the cytosolic domain of human Tom70 in complex with SARS-CoV-2 Orf9b protein (Gao et al., 2021; Gordon et al., 2020b) show that this viral protein occupies the complete pocket, both lower and the upper region (Figure 1.4). Whether the complete pocket of Tom70 is commonly involved in the client binding or this is an inhibition mechanism of the viral protein - preventing both Tom22 and client binding on Tom70, is currently being investigated in our group by characterizing the interactions of Tom70 receptor with other outer and inner membrane precursor proteins.



## 4.6 Determining kinetic parameters of the interactions

The NMR titration experiments were performed to determine the dissociation constant of interactions: (i) between the Tom receptors, (ii) Tom receptors and the presequence and (iii) Tom20 and the cytosolic chaperone Xdj1.

In the NMR titration experiments several NMR samples were prepared, all with the same analyte concentration (labeled protein). The concentration of the titrant (unlabeled protein or unlabeled peptide) varied. The  $^{15}\text{N}^1\text{H}$  spectra of each sample were recorded and the spectra were analysed with TITAN software (see Methods 7.15).

TABLE 4.1: Dissociation constants for Tom receptors, protein-protein or protein-peptide interactions.

interaction	Bootstrap (100 iterations)		Jackknife**	
	$K_D$ [ $\mu\text{M}$ ]	std error	$K_D$ [ $\mu\text{M}$ ]	error
ScTom20(37-183) / GB1-pSu9(1-69)	7.35	0.93	7.43	1.06
GB1-pSu9(1-69) / ScTom20(37-183)	17.19	6.77		
ScTom20(37-183) / pSu9(1-25) peptide	476.71	69.47		
RnTom20(51-145) / pSu9(1-25) peptide	1444.23	71.11		
RnTom20 / pALDH (Abe et al., 2000)	20-30			
Tom22(1-97) / ScTom20(37-183)	18.34	1.55		
ScTom20(37-183) / Tom22(1-97) pH 7	17.83	0.61		
ScTom20(37-183) / Tom22(1-97) pH 6.5	1.25	0.04	1.24	0.32
Tom22(1-74) / Tom70(39-617)	181.83	47.47	191.78	40.20
Tom22(1-74) / Xdj1(CTD1)	18.91	1.70	22.71	3.37
Xdj1(CTD1) / Tom22(1-74)	8.87	2.50	12.58	3.49

The dissociation constant of  $\sim 20 \mu\text{M}$  for ScTom20–pSu9(1-69) interaction was confirmed by the isothermal titration calorimetry (ITC) experiments (see TableS1 in the Appendix).

Measured dissociation constants of yeast Tom20–presequence interaction and yeast Tom20–Tom22 interaction are in the same range of  $\sim 20 \mu\text{M}$ . These measured values are in agreement with previously measured dissociation constant of rat Tom20 and pALDH presequence of 20-30 $\mu\text{M}$  (Abe et al., 2000). Interestingly, the binding of Tom70 to phosphate-carrier precursor peptide (PiC) was shown to be in the same range with a  $K_D \sim 70 \mu\text{M}$  (Mills et al., 2009). The shorter pSu9(1-25) peptide appears to have weaker affinity towards the Tom20, probably due to presence of only one recognition and binding site, compared to two interacting regions of the longer pSu9(1-69) construct. Considering similarity in the affinities between the receptors, and receptors towards the presequence, it would be interesting to measure the affinity of the presequence loaded Tom20 towards the Tom22 receptor. Additional binding affinities that should be measured to get an idea about the sequence of the import steps include the presequence peptides with Tom70 and precursor (VDAC or presequence) bound Tom70 with Tom22.

## 4.7 Conclusions on the molecular mechanism of the Tom receptor domains

The work presented in this manuscript focused on answering some of the fundamental questions concerning the first steps of the mitochondrial precursor protein import. After the initial recognition by either Tom20 or Tom70 receptor, how is the precursor protein further handled and directed towards the import pore? Do these steps include a direct interaction of the Tom22 protein and the precursor, or is the binding of Tom22 to Tom20 (or Tom70) modulating the interaction of these receptors and their clients? Previous extensive work in the field of mitochondrial protein import gave numerous mechanistic clues. However, these questions remained open.

Without a detailed atomic-view of the cytosolic domains of Tom22, Tom20 and Tom70 in the context of the whole TOM import machinery, it is not clear how the precursor proteins interact with the respective receptor domains and what is driving the precursors translocation through the OM. To get closer in answering these questions we studied the cytosolic domains of Tom20, Tom22 and Tom70, the interactions that these domains establish with different precursors and with each other. In our studies we relied mostly on NMR spectroscopy to obtain atomic level insights.

Tom22 is the central hub of the Tom40 core complex. However, it was not clear whether Tom22 is directly involved in precursor recognition as a receptor, like the better characterized Tom20 and Tom70 proteins, or whether its role is rather the recruitment of precursor-bound receptor proteins. Our initial focus was therefore the characterization of the cytosolic domain of Tom22 and of its interactions with the two Tom receptors and presequences of precursor proteins. Characterization of the cytosolic domain of yeast Tom22 showed that this domain of Tom22 is mostly intrinsically disordered. Remarkably, a short conserved sequence (residues 56-69) could be identified with the characteristics of a partially populated  $\alpha$ -helix. Predictions suggested the disordered nature, and also the possibility of a helix (Fig. 4.3). We were able to show experimentally the IDP-like structure of Tom22, as well as the partially populated amphiphilic helix. Using NMR experiments, we also demonstrated that the cytosolic domain of Tom22 does not interact with different presequence peptides under our experimental conditions. This is an interesting observation which speaks against a direct role of Tom22's cytosolic domain as a receptor for presequence-containing precursors or as a binding intermediate for precursors on their way from Tom20 to the Tom40 import channel.

To study the sequence of the events at early stages of the presequence-containing precursor proteins import, the confirmation of the precursor binding site on the yeast Tom20 receptor was needed before the interaction with Tom22 could be studied. The structure of the yeast Tom20 protein, neither in its apo state nor in the client-bound state, has been solved yet. The presequence binding site on ScTom20 was suggested based on the NMR determined binding site on the rat protein (Abe et al., 2000) and high sequence similarity of the two receptors. We characterized the presequence binding site of the ScTom20 receptor showing it is comprised of hydrophobic residues that are conserved in Tom20 proteins, suggesting a conservation of the presequence binding site. Having in mind that the presequences can form an amphiphilic helix (Schneider et al., 1998; Maduke and Roise, 1996; Pfanner, 2000), we characterized in more detail the presequence construct alone and when interacting with the Tom20 binding site. In the presequence construct, GB1-pSu9(1-69), most of the pSu9(1-69) residues are highly flexible and adopt the characteristics of an intrinsically disordered domain, compared to the well folded GB1 domain. From the chemical-shift analysis and the obtained relaxation parameters we observed that the short N-terminal region of the pSu9(1-69) presequence, comprising residues 7-22, has a tendency to form a transient  $\alpha$ helix. The interaction studies of GB1-pSu9(1-69) and the cytosolic domain of yeast Tom20 receptor revealed specific binding of the presequence to the Tom20 presequence binding site. Two Tom20-binding regions of the presequence could be

identified: one Tom20-binding region is the transient  $\alpha$ -helix, residues 7-22, and the other region comprises residues 53-64. Interestingly, relaxation data show that both Tom20-binding regions of pSu9(1-69) tend to populate a more folded state upon binding.

After showing that the mostly unfolded presequence tends to adopt a helical fold upon binding to the conserved presequence binding site on Tom20 receptor, we studied the interaction of the cytosolic domains of Tom22 and Tom20. Interestingly, our results show that both yeast and rat Tom20 bind the small cytosolic  $\alpha$ -helix of Tom22 (residues 56-69) in the conserved presequence binding site. The importance of the hydrophobic residues in the binding site on Tom20 is confirmed by weaker binding of Tom22 to Tom20 variants with a less hydrophobic (more positively charged) presequence binding site. Furthermore, the transient helix of Tom22, once bound to the presequence binding site of Tom20, shows characteristics of a slightly more folded, more  $\alpha$ -helical, protein sequence. We suggest that the cytosolic helix of Tom22 behaves like the presequence of the precursor proteins. More explicitly, the unfolded cytosolic domain of Tom22 binds to the precursor binding site on Tom20 receptor with a short protein sequence that tends to adopt an amphiphatic  $\alpha$ -helix. Additionally, we show that the binding affinity of the presequence and Tom22 towards Tom20 are in the same range,  $\sim 20 \mu\text{M}$ .

In our *in vitro* experiments we do not find any evidence for direct binding of the presequence to Tom22, and we, therefore, advocate that Tom22 does not serve as a receptor that would directly bind precursor proteins. Rather, both the presequence and Tom22 can bind to the same conserved binding site on the Tom20 receptor. However, it may be possible that Tom22 and Tom20-bound presequence interact. We show that the Tom20-bound presequence is stabilized by forming an amphiphilic helix with the hydrophobic side bound to Tom20. Therefore, it would be possible for the environment-exposed positively charged side of the presequence to interact with conserved negatively charged residues of Tom22. Our experiments, observing Tom22 or Tom20 upon adding one or two binding partners, did not indicate the presence of the ternary complex. There is an indication of a higher affinity of Tom22 towards the presequence-bound Tom20, which would support the previous observation that human Tom20 has decreased mobility in the membrane in the presence of a substrate protein (ascribed to association with the TOM core complex (Bhagawati et al., 2021)). However, further experiments measuring affinity need to be performed in order to confirm these observations. With both Tom22 and the presequence added to Tom20 we observe the averaged state of Tom20, where one part is pSu9-bound and the other is Tom22-bound, indicating competition of pSu9 and Tom22 for the same binding site on Tom20. While Tom22 is bound to Tom20 (with its helix), the precursor may interact with Tom20, thus forming a ternary complex.

Based on our results, we propose a mechanism for initial steps in presequence-containing precursor-protein import, and the role of the Tom22 interactions with the Tom20 receptor (and, similarly, Tom70). We propose that in the absence of Tom20 or Tom70, the intrinsically disordered Tom22<sub>cyt</sub> (partly) occludes the Tom40 pore entrance. Studies measuring the channel conductivity performed by Van Wilpe and colleagues indicated that the Tom40 channel is more often in an open state when the mitochondria are depleted of Tom22. They suggested that Tom22 is a negative regulator of the openness of the TOM pore (Van Wilpe et al., 1999), which supports our proposed role of Tom22<sub>cyt</sub> as occluding the TOM pore.

The peripheral Tom20 receptor recognizes and binds the presequence of the matrix-targeted precursor proteins by employing its conserved hydrophobic binding site in the cytosolic domain. The presequence-bound Tom20 is proposed to be docked to the TOM core complex by the cytosolic domain of Tom22 protein. Our data show that Tom22 interacts with the receptors, using its conserved helix, which is already transiently formed in the apo state. When Tom22<sub>cyt</sub> binds to the Tom20 receptor it is removed from the Tom40 pore entrance, thus clearing the passage (Figure 4.23). By replacing the presequence in the binding site of Tom20, the transient  $\alpha$ -helix of Tom22 would promote the dissociation of the presequence from Tom20 that

could then be imported by Tom40. This mechanism could explain previous studies where the similar effect on the import was observed upon deletion of either Tom22 or Tom20 cytosolic domains (Yamano et al., 2008b) or the import defects upon neutralizing the cytosolic domain of Tom22 with anti-bodies (Kiebler et al., 1993). Additionally, the proposed mechanism does not exclude the transient electrostatic interaction between the two amphiphilic  $\alpha$ -helices. In fact, the formation of amphiphatic helix, and hence the positively charged side of the presequence is induced by binding to Tom20 and it could well be that only then the transient electrostatic interaction with Tom22 is possible, before (or after) Tom22 replaces the presequence in the Tom20 binding site. Our results indicate that this electrostatic interaction is short lived as we could not obtain the complex of Tom22 and the presequence. This transient electrostatic interaction between the presequence and Tom22 could be paving the way towards the Tom40 pore. Thus, the binding of Tom22 to Tom20 has a dual role: extracting itself from the Tom40 pore, and helping the detachment of the precursor protein from Tom20. In fact, we propose that even in the absence of precursor protein, the Tom22-Tom20 interaction is present, and may at least contribute to keeping Tom20 - which otherwise can migrate in the membrane - close to the TOM complex. (It might even be possible that the interaction of Tom22<sub>cyt</sub> with the presequence-binding groove of Tom20 is the only interaction that keeps Tom20 close to TOM.)

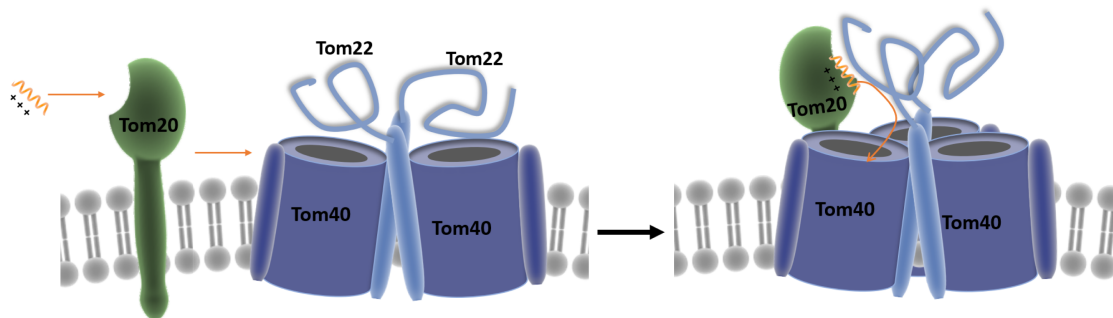
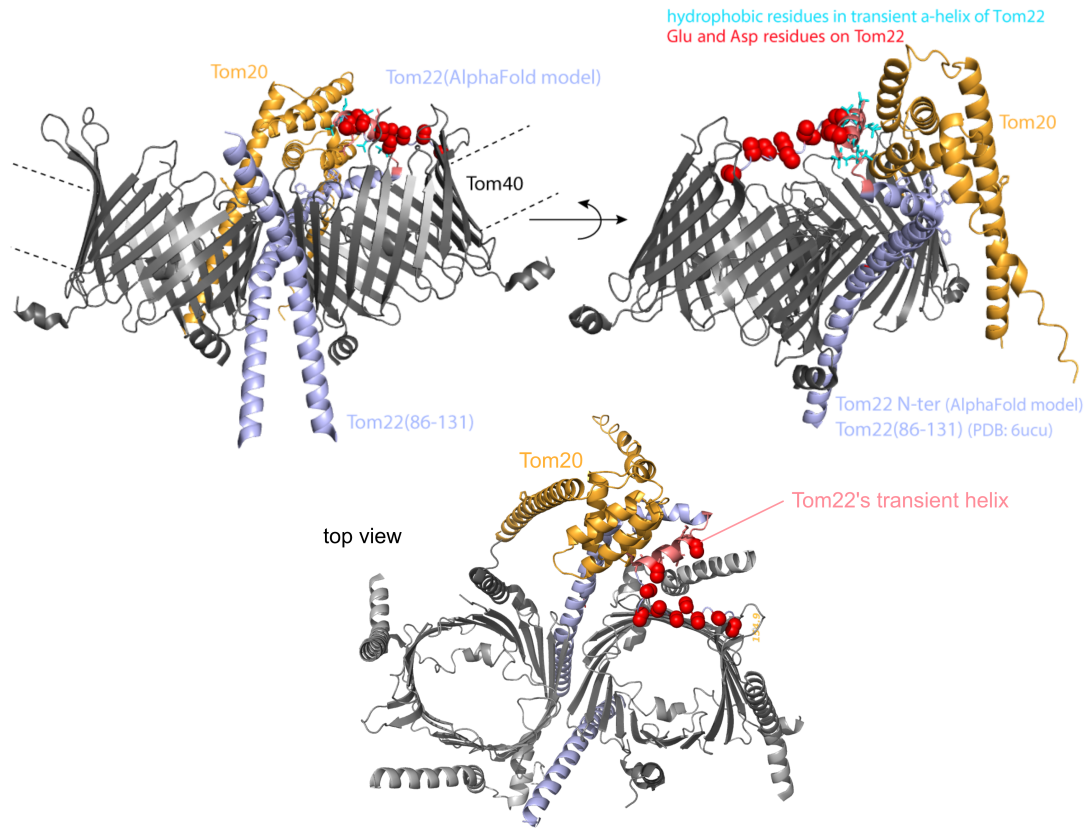


FIGURE 4.23: **Scheme of proposed recognition and interaction mechanism of the presequence by Tom20 and Tom22.** The Tom20 receptor is only loosely attached to the TOM core complex before recognizing and binding to the presequence. The presequence of the matrix targeted precursor protein is recognized and bound by Tom20 in a conserved hydrophobic binding site. The presequence-bound Tom20 is docked to the TOM core complex by the cytosolic domain of Tom22 protein. Transmembrane helix of Tom22 protein is anchored between two Tom40 pores, stabilizing the core complex, and the intrinsically disordered cytosolic domain of Tom22 is obscuring the entrance to the pore in the model shown in with this figure. In the mechanism we propose, the IDP domain of Tom22 is withdrawn from the pore entrance upon interaction of the transient cytosolic helix of Tom22 with the Tom20 receptor. By replacing the presequence in the binding site of Tom20, transient  $\alpha$ -helix of Tom22 would initiate the dissociation of the presequence from Tom20 that could then be imported by Tom40.

For the proposed mechanism to be plausible, we need to evaluate whether the interactions, in particular the one between the transient helix of Tom22<sub>cyt</sub> and Tom20, may form in the context of the entire TOM complex. In other words, can the cytosolic helix of Tom22 reach the presequence binding site on Tom20 in the outer membrane? We used the cryo-EM structure of the TOM core complex and the models we obtained using our NMR data and AlphaFold, and built a very coarse model (Figure 4.24). We obtained the AlphaFold-Multimer-v2 model of the ScTom20(34-185)-Tom22(1-97) complex and manually docked it to the TOM core complex (PDB: 6ucu; (Tucker, 2019)) using PyMOL. This was done by aligning the common region of Tom22 from the structure and AlphaFold model (residues 86-97). The AlphaFold model of Tom22-Tom20 complex is in agreement with our experimental data, showing the interface formed of hydrophobic side of the Tom22 helix and conserved presequence binding site of Tom20. In the model where Tom22-Tom20 complex is docked to the TOM core complex, the conserved negatively charged residues of Tom22 (shown in red in Figure 4.24) could provide the interface for the positively charged side of the presequence on its way from Tom20 towards the Tom40 pore. Our - admittedly coarse - model indicates that the proposed interaction

is indeed possible: the interface of Tom22-Tom20 is not hindered by the presence of other components of the TOM core complex and the cytosolic domains of Tom20 and Tom22 seem to be accessible for the interaction. More detailed modeling and molecular dynamics simulations are underway with the full-length, membrane anchored Tom20.



**FIGURE 4.24: Model of structural feasibility of the proposed mechanism.** Structure of the TOM core complex (PDB: 6ucu; (Tucker, 2019)) in this figure shows two Tom40 pores (gray) and transmembrane helices of Tom22(light blue). The AlphaFold-Multimer-v2 model of the complex ScTom20(34-185)–Tom22(1-97) is manually docked to the TOM core complex using PyMOL, by aligning the common region of Tom22 from the structure and AF model (residues 86-97). In the AlphaFold model the transmembrane helix of Tom20 receptor (residues 9-28) is not present. The interface formed of hydrophobic side of the Tom22 helix (shown in cyan) and conserved presequence binding site of Tom20 suggested by AlphaFold is in agreement with our experimental results. The conserved negatively charged residues of Tom22 are shown in red spheres. For clarity, the small Tom proteins in the TOM core complex are not shown, however they were used in modeling session.

We further investigated whether the proposed mechanism, in which Tom22 binds in the client binding site presumably to allow release of the client, is a unique mechanism of Tom22 when interacting with Tom20 or the same interaction patterns are used when Tom22 interacts with Tom70 receptor. To begin with, we established the binding site for one type of clients that Tom70 recognizes and binds. Our results show that a cyclic peptide fragment of a  $\beta$ -barrel type of client binds in the previously proposed conserved hydrophobic binding pocket on the cytosolic domain of Tom70 receptor (Wu and Sha, 2006). Surprisingly, we experimentally observed binding of Tom22 on Tom70 in the lower binding site, in the same C-terminal cleft, where the binding of the presequences was previously observed (Melin et al., 2015). Interestingly, the recent cryo-EM and crystal structures of the cytosolic domain of human Tom70 in complex with SARS-CoV-2 Orf9b protein (Gao et al., 2021; Gordon et al., 2020b) show that this viral protein occupies the complete pocket, both lower and the upper region (Figure 1.4). Whether the complete pocket of Tom70 is commonly involved in the client binding or this is an inhibition mechanism of the viral protein - preventing both Tom22

and client binding on Tom70, is currently being investigated. Further studies on Tom70 receptor ongoing in the laboratory aim to characterize the interactions of Tom70 receptor with the other outer membrane (e.g. Om14, Mim2), inner membrane (e.g. PiC) precursor proteins and the presequences of the matrix targeted precursors. Furthermore, as Tom70 receptor appears to be highly promiscuous, it will be interesting to see differences in binding affinities as well as in the binding site, if any. The interaction studies performed with Tom70 and Tom22 imply that Tom22 uses a consistent interaction pattern with both receptors: binding of a transient cytosolic helix to the precursor binding region.

The results we obtained so far in our studies of the cytosolic domains of mitochondrial Tom receptors allow us to propose a molecular model for the initial steps in precursor protein import. Tom22, the central hub of the TOM complex, along with its essential role in stabilizing the import pore, has two additional roles in precursor protein import process: (i) docking of the precursor-loaded Tom20/Tom70 receptors and (ii) facilitating release of the precursor towards the Tom40 pore by replacing the precursor in the Tom20/Tom70 binding sites.



## Chapter 5

# Results 2. Molecular basis of the client specificity of the TIM holdases.

The only two chaperones of the mitochondrial intermembrane space, TIM8-13 and TIM9-10, bind and transport hundreds of client proteins different in size, hydrophobicity and final fold.

Earlier research in the Schanda group enlightened the mode of binding of mitochondrial carrier proteins to the TIM9-10 chaperone (Weinhäupl et al., 2018) which until then was only speculated based on the chaperones crystal structures (Webb et al., 2006; Beverly et al., 2008). Obtained structure of two TIM9-10 chaperones bound to one all-transmembrane mitochondrial carrier Ggc1 revealed the interactions behind apparent contradictory roles of this holdase chaperone; (i) preserving client from detachment and aggregation in the intermembrane space and (ii) release of the client to the downstream insertase complex without significant energy barrier. The TIM9-10 bound client is highly dynamic with multiple short-lived conformations (Weinhäupl et al., 2018). Highly conserved hydrophobic residues scattered in the TIM9-10 binding site(s), between inner and outer helix of each Tim subunit, are shown to form this primary client binding site which seems to provide appropriate interface for such a dynamic but stable complex to form.

In the work presented here, we determined experimentally two distinct models of chaperone-client binding of these two structurally highly similar, 65 kDa hetero-hexamers with alternating subunits, chaperones. Combining NMR, SAXS and molecular dynamics simulations, we determine the structures of TIM8-13 and TIM9-10 bound to Tim23, integral membrane protein with an additional hydrophilic domain. TIM8-13 uses transient salt bridges to interact with the hydrophilic part of its client, but its interactions to the transmembrane part are weaker than in TIM9-10. Consequently, TIM9-10 outcompetes TIM8-13 in binding hydrophobic clients, while TIM8-13 is tuned to few clients with both hydrophilic and hydrophobic parts.

This work was done in collaboration with the group of Prof. Dr. K. Lindorff-Larsen, with Dr. Y. Wang performing and analysing the molecular dynamic simulations. In collaboration with the group of Prof. Dr. D. Rapaport, Dr. T. Jores performed and analysed VDAC cross-linking experiments. Dr. K. Weinhäupl performed and analysed initial experiments, including SAXS experiments, together with Dr. M. Brennich. This work was part of the internship projects of O. Dakhlaoui and D. Costa who, together with A. Hessel, prepared part of the protein samples and experiments.

The following results are part of published work I. Sučec, Y. Wang, O. Dakhlaoui, K. Weinhäupl, T. Jores, D. Costa, A. Hessel, M. Brennich, D. Rapaport, K. Lindorff-Larsen, B. Bersch and P. Schanda (2020). **Structural basis of client specificity in mitochondrial membrane-protein chaperones.** *Sci. Adv.* 6.51.



## 5.1 Introduction to the Client specificity of mitochondrial TIM chaperones

Cellular survival and function fundamentally rely on an intact proteome. Proteins within cells need to be correctly folded to their functional conformation, and be present at the cellular location where they function. Chaperones play a central role in maintaining this cellular protein homeostasis (Powers et al., 2009), by either helping other proteins to reach their functional three-dimensional structure after synthesis, by transporting them across the cytosol or organelles, or by sustaining their native fold along their lifetime. More than 20,000 different proteins are required to fulfill the functions of human cells, and it is believed that the majority rely on chaperones to reach and maintain their native fold (Kim et al., 2013). Given the diversity of the client proteins, many chaperones promiscuously interact with tens of different 'client' proteins that may differ widely in size, structure and physico-chemical properties. However, the need for efficient binding and refolding of their clients also calls for some degree of specificity. Chaperones operate at this delicate balance of promiscuity and specificity to their clients. The interactions that determine the chaperone–client specificity are only partly understood.

Hydrophobic interactions play a crucial role for chaperone interactions as most chaperones bind to hydrophobic patches on their clients and shield them from aggregation. Electrostatic charges also play a role in some chaperone complexes (Kim et al., 2015). The interaction motifs recognized by different chaperones differ by their physico-chemical properties (Bose and Chakrabarti, 2017). For example, for interacting with the Hsp70 chaperone family, Ile, Phe, Leu and Val residues are particularly important (Rüdiger et al., 1997; Blond-Elguindi et al., 1993); the SecB chaperone recognizes 9-residue long stretches enriched in aromatic and basic residues (Knoblauch et al., 1999); the chaperone Spy uses longer-range charge interactions for the formation of an initial encounter complex, followed by more tight binding mediated by hydrophobic interactions, (Koldewey et al., 2016) whereby structurally frustrated sites on the client protein are particularly prone to binding (He et al., 2016a).

Our understanding of the underlying principles of chaperone-client interactions is hampered by the lack of atomic-level views onto the structure and dynamics of these complexes. Their inherently dynamic and often transient nature represents a significant experimental challenge towards structural characterization. Only a very limited number of chaperone complex structures have been reported (reviewed in (Hiller and Burmann, 2018)). The modes of interactions that they revealed range from rather well-defined binding poses of client polypeptides in the chaperone's binding pockets, reminiscent of complexes formed by globular proteins, to highly flexible ensembles of at least partly disordered conformations ('fuzzy complexes'). In the latter, a multitude of local chaperone-client interactions may result in a high overall affinity despite the low affinity and short life time of each individual inter-molecular contact.

Multiple molecular chaperones are present in the cell with mutually overlapping functions and 'clientomes' (Kim et al., 2013; Genevaux et al., 2004; Bukau et al., 2000). It is poorly understood, however, whether a given client protein adopts a different conformation (or ensemble of conformations) when it is bound to different chaperones, and if different clients, when bound to a given chaperone, all show similar conformational properties.  $\alpha$ -Synuclein appears to have similar interaction patterns with six different chaperones (Burmann et al., 2020); outer-membrane proteins (OmpA, OmpX, FhuA) have similar properties – essentially fully unfolded – when bound to SurA and Skp chaperones (Burmann, Wang, and Hiller, 2013; Thoma et al., 2015), at least as judged by their NMR fingerprint spectra. Phosphatase A displays an extended dynamic conformation, but well-defined binding poses of its interacting parts, when bound to trigger factor (Saio et al., 2014), Hsp40 (Jiang, Rossi, and Kalodimos, 2019) or SecB (Huang et al., 2016). Thus, while these reports suggest that a given protein adopts similar properties on different chaperones, the scarcity of

data and absence of a direct comparison of complex structures leave open which interactions may confer specificity.

A pair of 'holdase' chaperone complexes of the mitochondrial inter-membrane space (IMS), TIM8·13 and TIM9·10, are structurally highly similar, but have different substrate binding preferences. These chaperones transport precursors of membrane proteins with internal targeting sequence (henceforth denoted as 'precursors') to the membrane-insertase machineries in the inner membrane (TIM22) and outer mitochondrial membranes (SAM) (Wiedemann and Pfanner, 2017)). The TIM chaperones form hetero-hexameric structures of ca. 70 kDa, composed of an alternating arrangement of Tim9 and Tim10 or Tim8 and Tim13. TIM9·10 is essential to cellular viability (Sirrenberg et al., 1998; Koehler et al., 1998; Lionaki et al., 2008); even single point mutations in Tim9 or Tim10 that keep the chaperone structure intact but affect precursor protein binding can impair yeast growth and cause lethality (Weinhäupl et al., 2018). Although TIM8·13 is not essential in yeast (Koehler et al., 1999), yeast cells depleted of Tim8 and Tim13 show conditional lethality (Paschen et al., 2000). Additionally, mutations in the human Tim8a protein have been identified as the cause of a neurodegenerative disorder known as Mohr-Tranebjærg (MTS) or Deafness-Dystonia-Optic Neuropathy (DDON) syndrome (Roesch et al., 2002; Kang et al., 2019).

*In vivo* experiments, predominantly in yeast, have identified mitochondrial membrane proteins whose biogenesis depends on small TIM chaperones. TIM9·10 is believed to interact with all members of the mitochondrial carrier (SLC25) family, which comprises more than 50 members in humans, such as the ADP/ATP carrier (Aac in yeast); TIM9·10 furthermore transports the central components of the TIM22 and TIM23 insertion machineries (Tim23, Tim17, Tim22) as well as outer-membrane  $\beta$ -barrel proteins (Morgenstern et al., 2017). TIM8·13 has a narrower clientome, and was shown to bind the precursors of the inner-membrane proteins Tim23 (Paschen et al., 2000; Leuenberger et al., 1999; Davis et al., 2000) and  $\text{Ca}^{2+}$ -binding aspartate-glutamate carriers (Roesch et al., 2004), as well as the outer-membrane  $\beta$ -barrel proteins VDAC/Porin, Tom40 (Hoppins and Nargang, 2004) and Tob55/Sam50 (Habib et al., 2005). There is evidence that TIM8·13 does neither bind the inner-membrane proteins ADP/ATP carrier (Aac) nor Tim17 (Paschen et al., 2000). The inner-membrane proteins that have been reported to interact with TIM8·13 have a hydrophilic domain in addition to trans-membrane domains (fig. S1) but this does not hold true for the outer-membrane  $\beta$ -barrels. Thus, the mechanisms by which TIM8·13 binds its clients remain unclear.

Recently, we obtained the first structure of a complex of a small TIM chaperone, TIM9·10, with the mitochondrial GDP/GTP carrier (Ggc1) (Weinhäupl et al., 2018). The structure, composed of two chaperone complexes holding one precursor protein, revealed a highly dynamic ensemble of Ggc1 conformers that form multiple short-lived and rapidly inter-converting ( $< 1$  ms) interactions with a hydrophobic binding cleft of the chaperone (fig. S2). The TIM9·10-Ggc1 complex can be described as a "fuzzy complex", in which the high overall affinity is driven by a multitude of individually weak interactions with the hydrophobic trans-membrane (TM) parts of its clients.

To understand what confers specificity in the mitochondrial IMS chaperone system, we studied chaperone complexes of TIM9·10 and TIM8·13 with two precursor proteins, the Ggc1 and the insertase component Tim23. In their native state, Ggc1 comprises six TM helices without soluble domains, and Tim23 four TM helices and a ca. 100-residue-long soluble inter-membrane space domain (Fig. 5.1A). By solving the complex structures of the two chaperone complexes holding Tim23, we reveal that the differential specificity of the two chaperones is based on an interplay of hydrophobic and hydrophilic interactions, which leads to different conformational properties of the precursor protein bound to these chaperones.

## 5.2 TIM8·13 and TIM9·10 interact differently with membrane precursor proteins

We have developed an experimental protocol (Weinhäupl et al., 2018) to prepare complexes of the inherently insoluble membrane-protein precursors and chaperones (Fig. 5.1B,C). Briefly, the approach involves the recombinant production of His-tagged precursor protein, binding it to a His-affinity column (NiNTA) in denaturing conditions, followed by removal of the denaturant and simultaneous addition of a chaperone. The chaperone-precursor complex is then eluted for further biochemical, biophysical and structural investigations.

The measurement of dissociation constants of chaperones and membrane precursor proteins, using methods such as isothermal titration calorimetry or surface plasmon resonance, is not possible, because the complexes cannot be formed in solution (e.g. flash-dilution methods, which work for other chaperones (Burmann, Wang, and Hiller, 2013), failed; data not shown). Thus, to characterize the relative affinities of the precursor proteins to the two chaperones, we performed different types of competition experiments. In a first experiment, precursor protein was bound to the affinity resin, and both chaperones were simultaneously added, before washing excess chaperone, and eluting the chaperone-precursor complexes (Fig. 5.1C). NMR spectroscopy shows that the two chaperones do not form mixed hetero-hexameric complexes, implying that TIM9·10 and TIM8·13 stay intact in such competition experiments (fig. S3). In a second class of experiments, we prepared one type of complex (e.g. TIM9·10-Tim23) and added the other chaperone (e.g. TIM8·13) in its apo state, allowing the precursor protein to be transferred. These experiments also demonstrate that membrane precursor proteins can be transferred between these two chaperones, on the time scale we investigated (minutes to hours). We used SDS-PAGE analyses and electrospray ionization mass spectrometry (ESI-MS) to systematically quantify the amount of obtained complexes (Fig. 5.1D and fig. S4). Consistently, we find that Ggc1 has a strong preference for TIM9·10 (ca. 5- to 10-fold), while Tim23 shows a slight preference for TIM8·13 (ca. 1.5-fold).

Taken together, we established that the two chaperones bind with different affinities to two inner-membrane precursor proteins, whereby TIM8·13 is barely able to hold Ggc1, in contrast to TIM9·10, while it can hold Tim23 slightly better than TIM9·10.

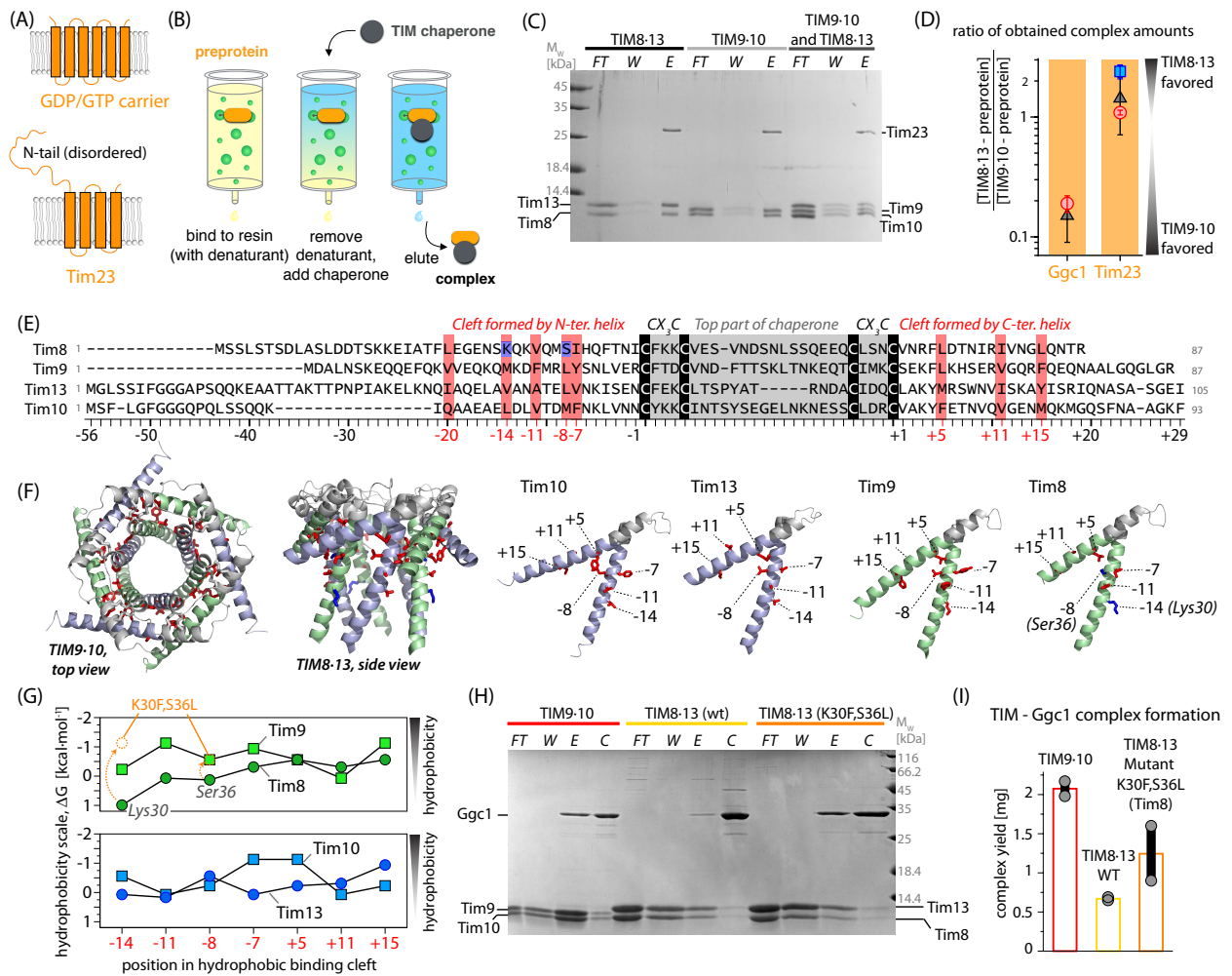


FIGURE 5.1: **Biochemical characterization of TIM chaperone - membrane protein complexes.** (A) Native topology of the two precursor proteins used in this study. (B) Schematic view of the pull-down experiment used to prepare chaperone-precursor complexes. (C) Formation of Tim23-chaperone complexes, monitored by SDS-PAGE. Either TIM8-13, TIM9-10 or a 1:1 mixture was added to NiNTA-bound Tim23. The lanes correspond to flow-through after applying chaperone (FT), additional wash (W), and imidazole elution (E). Protein bands corresponding to Tim10 and Tim8 overlap. (D) Relative amounts of complexes with Ggc1 and Tim23, obtained from three different experiments: (i) a pull-down assay where both chaperones were applied to bound precursor protein (black), (ii) preparation of a TIM9-10-precursor protein complex and addition of TIM8-13, and SDS-PAGE and mass spectrometry analysis after 1 and 3 hours (red), (iii) preparation of TIM8-13-Tim23 followed by TIM9-10 addition and SDS-PAGE (blue) as in (ii). The protein amounts were determined from LC/ESI-TOF-MS (fig.S4 in A); error estimates from  $\geq$  two experiments. (E) Sequence alignment of the small Tims, numbered from the N- and C-terminal conserved Cys residues ("0"). Red: conserved hydrophobic residues; blue: hydrophilic Tim8 residues, K30 and S36L. See fig.S5 in A for comprehensive alignment. (F) Location of the residues in the hydrophobic cleft. (G) Comparison of Kyte-Doolittle hydrophobicity of the residues in the binding cleft of wild-type (WT) native Tim proteins and Tim8<sub>K30F,S36L</sub>. (H) Pull-down experiment of Ggc1 with TIM9-10, TIM8-13 and TIM8-13(Tim8<sub>K30F,S36L</sub>). Lane descriptions are as in (C); additionally, the fraction obtained after final wash with urea and imidazole, to control the Ggc1 initially loaded onto the column, is shown (control, C). (I) Amount of complex obtained from pull-down experiments of WT and mutant chaperones; the same amount of Ggc1 was applied in all three experiments, and the total amount of eluted complex was determined spectroscopically.

### 5.3 The small TIM chaperones use a conserved hydrophobic cleft for membrane precursor protein binding

To understand the different binding properties, we performed a sequence alignment of the small TIMs, which reveals a well conserved set of hydrophobic residues that point towards the binding cleft formed between the inner (N-terminal) and outer tentacles (Weinhäupl et al., 2018) (Fig. 5.1E,F). The overall hydrophobicity of these residues is lower in Tim8 and Tim13 than in Tim9 and Tim10 (Fig. 5.1G). In particular, Tim8 has a charged residue in position -14 (Lys30). (The sequences are numbered starting with negative numbering at the twin CX<sub>3</sub>C motif towards the N-terminus, and positive numbering from the last Cys to the C terminus and the hydrophobic motif residues are at positions -20, -14, -11, -8, -7, +5, +11 and +15.) This positive charge at position -14, either Lys or Arg, is conserved among eukaryotes (fig. S5). In yeast, position -8 of the hydrophobic motif is polar (Ser36), although this position is not strictly conserved. Overall, these residues make the hydrophobic binding cleft of Tim8 less hydrophobic than in the other small Tims.

We speculated that the less hydrophobic nature of TIM8·13's binding cleft reduces its affinity to TM parts of membrane precursor proteins. To test this hypothesis, we generated a mutant TIM8·13 with increased hydrophobicity (Tim8<sub>K30E,S36L</sub>; Fig. 5.1G). This more hydrophobic TIM8·13(Tim8<sub>K30E,S36L</sub>) chaperone allowed us to obtain significantly larger amounts of complex with Ggc1 than native TIM8·13, under otherwise identical conditions (Fig. 5.1H,I). This observation establishes the importance of the hydrophobic cleft for binding hydrophobic TM parts of precursor proteins. Equivalent experiments with the full-length Tim23, shown in fig. S6, reveal that the additional hydrophobic residues in the binding cleft of Tim8<sub>K30E,S36L</sub> do not improve its capacity to bind Tim23. This observation suggests that the binding mechanisms in place for binding these two different precursor proteins differ.

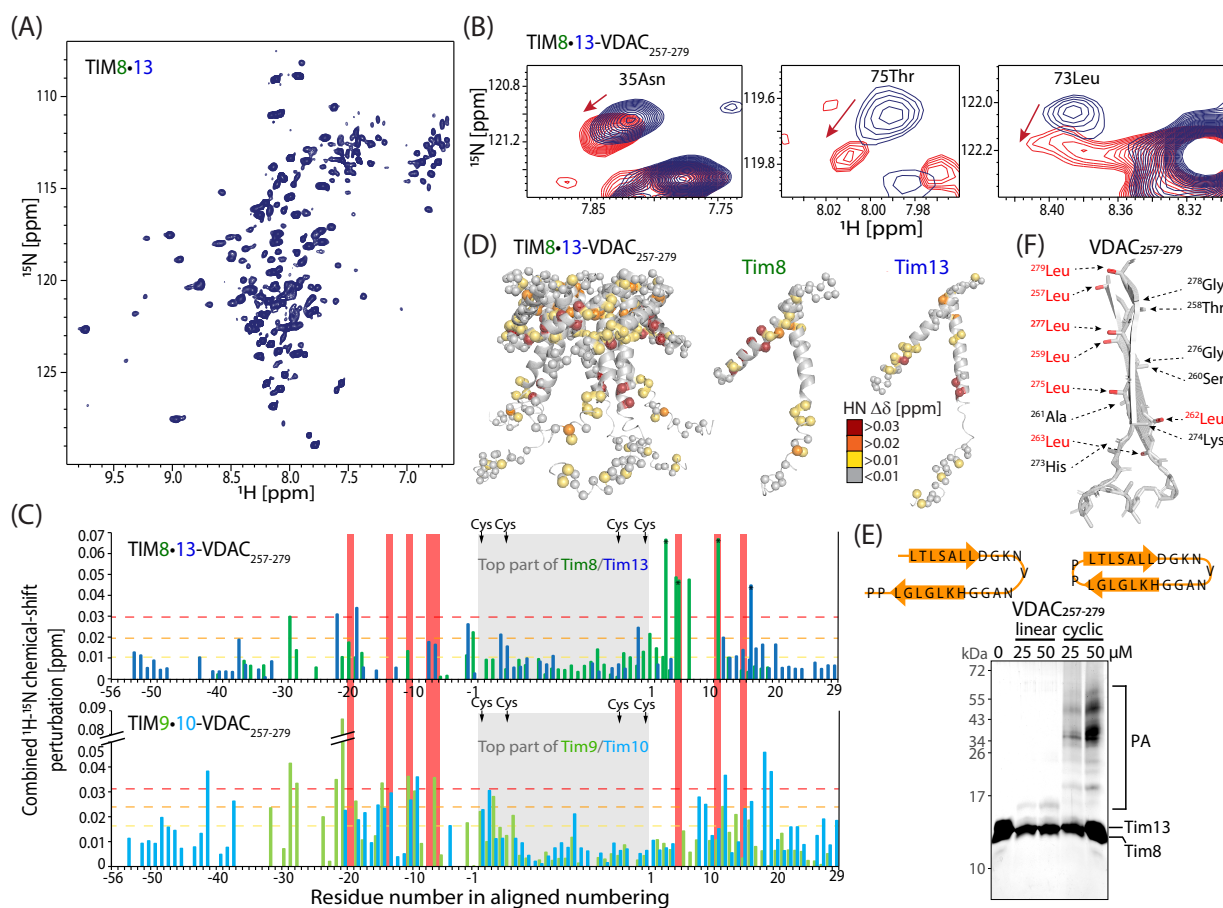


FIGURE 5.2: **Solution-NMR and binding of a VDAC fragment to TIM8-13.** (A)  $^1\text{H}$ - $^{15}\text{N}$  NMR spectrum of TIM8-13 at 35 . (B) Chemical-shift perturbation (CSP) in TIM8-13 upon addition of 5 molar equivalents of cyclic VDAC<sub>257-279</sub>. (C) CSP effects of VDAC<sub>257-279</sub> binding. The data for TIM9:10 are from ref. (Weinhäupl et al., 2018). (D) Plot of CSP data on the TIM8-13 structure. (E) Photo-induced cross-linking of the linear (left) and cyclic (right) VDAC<sub>257-279</sub> peptides to TIM8-13. While hardly any adducts are observed for the linear one, the cyclic peptide forms cross-linking photo-adducts (PA), including of higher molecular weight, resulting from multiple cross-links, as reported earlier (Jores et al., 2016; Jores et al., 2018; Weinhäupl et al., 2018). (F) Schematic structure of the two last strands of VDAC, as found in the NMR structure (Hiller et al., 2008) of the full  $\beta$ -barrel, showing that the hydrophobic and hydrophilic side chains cluster on the two opposite faces of the  $\beta$ -turn.

To better understand the client-binding properties of the two chaperones, we turned to structural studies. Solution-NMR spectra of apo TIM8-13 (Fig. 5.2A) and residue-wise resonance assignments allowed identifying the residues forming secondary structure and estimating their local flexibility. In agreement with the crystal structure, the core of rather rigid tentacles comprises the top part of the chaperone between the CX<sub>3</sub>C motifs and ca. 15-25 residues before and after these motifs. About 10-20 residues on each the N- and C-termini are flexible (fig. S7). To probe the binding of a trans-membrane segment of a membrane precursor protein, we performed NMR-detected titration experiments of TIM8-13 with a cyclic peptide corresponding to the two C-terminal strands of the  $\beta$ -barrel voltage-dependent anion channel (VDAC<sub>257-279</sub>) that has a propensity to form a  $\beta$ -turn (Jores et al., 2016). Addition of this cyclic VDAC<sub>257-279</sub> induces chemical-shift perturbations (Fig. 5.2B), that are primarily located in the hydrophobic cleft formed between the inner and the outer rings of helices (Fig. 5.2C,D). This binding site matches very closely the site on TIM9-10 to which VDAC<sub>257-279</sub> binds (Weinhäupl et al., 2018) (Fig. 5.2C). Interestingly, the VDAC<sub>257-279</sub>-induced chemical-shift perturbation (CSP) effects in TIM8-13 are overall only about half of the magnitude of CSPs found in TIM9-10, pointing to a higher population of the TIM9-10-VDAC<sub>257-279</sub> complex compared to TIM8-13-VDAC<sub>257-279</sub> at comparable conditions (Fig. 5.2C). This finding suggests a lower affinity of TIM8-13 to

VDAC<sub>257-279</sub>, as expected from its lower hydrophobicity.

Photo-induced cross-linking experiments of a Bpa-modified VDAC<sub>257-279</sub> peptide to TIM8·13 show that only the cyclic peptide forms cross-linking adducts while the linear, mostly disordered (Jores et al., 2016) form does not (Fig. 5.2E). The same behavior was also found for TIM9·10 (Weinhäupl et al., 2018) and yeast cytosolic chaperones Ssa1, Ydj1, Djp1, and Hsp104 (Jores et al., 2018). A rationale for this finding is the fact that in a  $\beta$ -turn the side chains of consecutive residues point to the two opposing faces thus creating one hydrophobic and one more hydrophilic face (Fig. 5.2F). In contrast, due to its disorder, the linear VDAC<sub>257-279</sub> peptide does not have a stable hydrophobic face, reducing its affinity to the hydrophobic binding cleft on the chaperone. In line with these findings, NMR titration data with the linear peptide shows small CSPs that are spread across the protein, thus pointing to unspecific interaction (fig. S8 in A).

Collectively, the experiments with the client fragment VDAC<sub>257-279</sub> provide a first evidence that both chaperones use the same conserved binding cleft to interact with hydrophobic membrane precursor protein sequences, and that TIM9·10 interacts more efficiently with TM parts, and thus with Ggc1 and the VDAC fragment. We propose that the more hydrophobic nature of the binding cleft in TIM9·10 allows it to interact more strongly with TM parts of its clients. In light of this observation, how does TIM8·13 achieve a binding affinity to Tim23 which is slightly higher than the one of TIM9·10 (Fig. 5.1D)?

## 5.4 Hydrophilic fragments interact differently with TIM8·13 and TIM9·10

Tim23 has a hydrophilic N-terminal segment in addition to four TM helices (Fig. 5.3A), and we investigated whether this part interacts with the chaperones. NMR spectra of the soluble Tim23<sub>IMS</sub> fragment (residues 1-98) in isolation show the hallmark features of a highly flexible intrinsically disordered protein with low spectral dispersion of <sup>1</sup>H-<sup>15</sup>N NMR signals (Fig. 5.3B,C, orange spectrum), as previously reported (Cruz et al., 2010). Upon addition of TIM9·10, the Tim23<sub>IMS</sub> <sup>1</sup>H-<sup>15</sup>N spectrum (Fig. 5.3B, left) shows only small changes: all cross-peaks are still detectable, and small chemical-shift perturbations (CSPs) are only observed for a few residues at the N-terminus, which has higher hydrophobicity (Fig. 5.3D). This finding suggests only very weak, possibly non-specific interactions between the very N-terminus of Tim23<sub>IMS</sub> and TIM9·10. In line with this finding, the interaction is not detectable by isothermal titration calorimetry (ITC) measurements (Fig. 5.3E).

The interaction of the hydrophilic Tim23<sub>IMS</sub> fragment with TIM8·13 is significantly stronger, with pronounced binding effects detected by ITC, and a dissociation constant of  $K_d = 66 \pm 8 \mu\text{M}$  (Fig. 5.3E, right panel; see Table S1 in A). The <sup>1</sup>H-<sup>15</sup>N NMR spectrum of Tim23<sub>IMS</sub> in the presence of TIM8·13 shows strongly reduced peak intensities for the majority of the residues (Fig. 5.3C, left). Such a peak broadening is expected when a highly flexible polypeptide binds to a relatively large object such as TIM8·13, thereby inducing faster nuclear spin relaxation and thus broader signals of lower intensity. Analysis of the peak-intensity reduction reveals two regions of Tim23 that are particularly involved in the binding: (i) the N-terminal hydrophobic residues, which are also involved in interacting with TIM9·10, and (ii) a long sequence stretch comprising residues from ca. 30 to 80 (Fig. 5.3F).

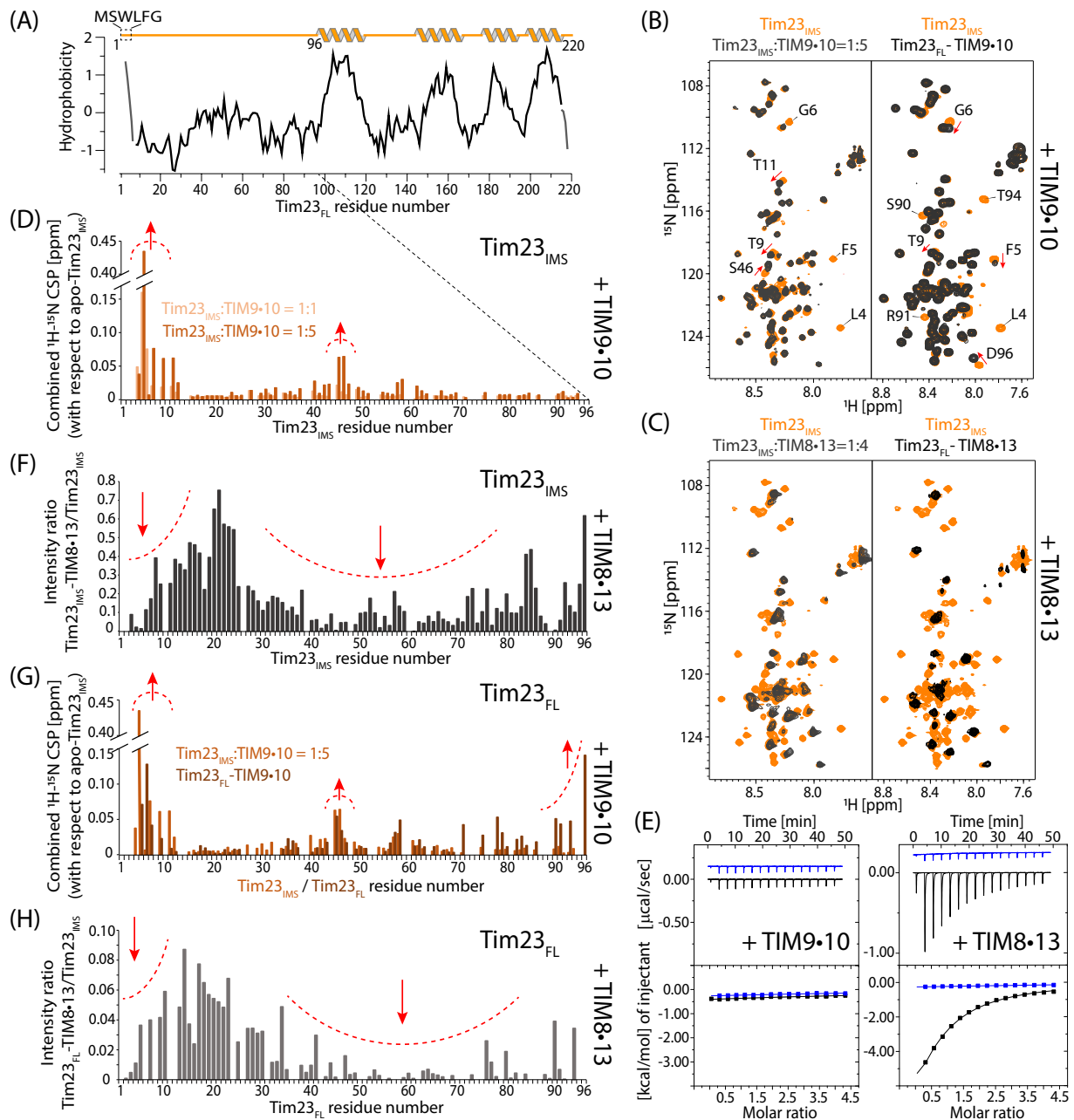


FIGURE 5.3: **Tim23 has markedly different properties when binding to TIM8-13 and to TIM9-10.** (A) Hydrophobicity of Tim23 (Kyte-Doolittle). (B) NMR spectra of the  $^{15}\text{N}$ -labeled soluble Tim23<sub>IMS</sub> fragment in the presence of TIM9-10 (left, black), and of full-length Tim23 bound to TIM9-10 (right, black) are compared to the Tim23<sub>IMS</sub> fragment in isolation (orange), under identical buffer conditions and NMR parameters. (C) As in (B) but with TIM8-13 instead of TIM9-10. (D) Chemical-shift perturbation (CSP) of residues in Tim23<sub>IMS</sub> upon addition of 1 (light orange) or 5 (dark orange) molar equivalents of TIM9-10. (E) Calorimetric titrations for the interaction of TIM9-10 or TIM8-13 (54  $\mu\text{M}$  in the calorimetric cell) with Tim23<sub>IMS</sub> (1.15 mM in the injecting syringe). Thermograms are displayed in the upper plots, and binding isotherms (ligand-normalized heat effects per injection as a function of the molar ratio,  $[\text{Tim23}_{\text{IMS}}]/[\text{chaperone}]$ ) are displayed in the lower plots. Control experiments, injecting into a buffer, are shown in blue. (F) Intensity ratio of residues in Tim23<sub>IMS</sub> in the presence of 4 molar equivalents of TIM8-13 compared to Tim23<sub>IMS</sub> alone. (G) CSP of the detectable residues in full-length Tim23 attached to TIM9-10 (brown), compared to the soluble Tim23<sub>IMS</sub> fragment. (H) Intensity ratio of detectable residues in Tim23<sub>FL</sub> attached to TIM8-13. Note that the ratio was not corrected for differences in sample concentration, and the scale cannot be compared to the one in panel (G).

To investigate whether TIM8-13 may interact with another soluble protein from the inter-membrane space,



we performed ITC experiments with the globular protein cytochrome c. No interaction could be detected (fig. S9), suggesting that the TIM8·13-Tim23<sub>IMS</sub> interaction may be related to the unfolded, flexible character of the latter.

To characterize the conformation of full-length (FL) Tim23 bound to TIM8·13 and TIM9·10, we prepared Tim23<sub>FL</sub>-labeled Tim23-chaperone complexes using the method outlined in Fig. 5.1B. Very similar to the experiments with the Tim23<sub>IMS</sub> fragment, the signals corresponding to the N-terminal half of Tim23 are still intense in the Tim23<sub>FL</sub>-TIM9·10 complex (Fig. 5.3B,G) revealing that the N-terminal half of Tim23<sub>FL</sub> does not interact strongly with TIM9·10. The small observed CSPs are localized primarily at the hydrophobic N-terminus. In contrast, when Tim23<sub>FL</sub> is bound to TIM8·13, the signals corresponding to its N-terminal half are severely reduced in intensity, revealing tight contact of the flexible N-terminal half of Tim23 to TIM8·13 (Fig. 5.3C,H).

In neither of the two Tim23<sub>FL</sub> complexes any additional signals, that may correspond to Tim23's TM helices, are visible. We ascribe this lack of detectable signals of residues in the TM part to extensive line broadening. The origin of this line broadening may be ascribed to the large size of the complex and likely to additional millisecond (ms) time scale dynamics of Tim23's TM parts in the hydrophobic binding cleft of the chaperones. Such millisecond motions have been found in the TIM9·10-Ggc1 complex (Weinhäupl et al., 2018).

## 5.5 TIM8·13 uses an additional hydrophilic face for protein binding

We probed the binding sites that the chaperones use to interact with Tim23<sub>IMS</sub> or Tim23<sub>FL</sub> using NMR spectroscopy on samples in which only the chaperone was isotope-labeled. Interestingly, the CSPs in the two chaperones upon addition of Tim23<sub>IMS</sub> reveal distinct binding patterns (Fig. 5.4A): in TIM8·13, the largest effects involve residues in the hydrophilic top part of the chaperone, between the CX<sub>3</sub>C motifs, as well as a few residues toward the C-terminal outer ring of helices; in contrast, the corresponding top part of TIM9·10 does not show any significant effects, but CSPs are observed at residues in the hydrophobic binding cleft, and in particular the N-terminal helix (Fig. 5.4B,C). This data, together with the Tim23<sub>IMS</sub>-detected data in Fig. 5.3 establish that TIM8·13 uses its hydrophilic top part to bind Tim23's N-terminal half, while only a short stretch of hydrophobic residues at the very N-terminus of Tim23 interacts with the hydrophobic cleft of TIM9·10, which is also the binding site of TM parts (Figs. 5.1 and 5.2).

Chaperone-labeled complexes with Tim23<sub>FL</sub> confirm these findings, and point to the additional effects induced by the bound TM part: in TIM9·10-Tim23<sub>FL</sub>, large CSP effects are located primarily in the binding cleft, in line with the view that the top part of TIM9·10 is not involved in binding Tim23. In contrast, Tim23<sub>FL</sub>-induced CSPs are found across the whole TIM8·13, including the hydrophilic top and the hydrophobic cleft (Fig. 5.4D,E and fig. S10).

We furthermore prepared complexes of a truncated Tim23 fragment (Tim23<sub>TM</sub>, residues 92-222), which allows to detect selectively the interaction of the trans-membrane part with the chaperones. The TIM9·10-Tim23<sub>TM</sub> complex features the largest CSPs in the hydrophobic binding cleft, qualitatively similar to the binding site detected with Tim23<sub>FL</sub> (Fig. 5.4F and fig. S11). The complex of Tim23<sub>TM</sub> with TIM8·13 appears to be much less stable than TIM9·10-Tim23<sub>TM</sub>: in the pull-down experiment, only a very small amount of complex could be obtained, and the complex rapidly precipitated (not shown), excluding NMR analyses. This observation reflects that the hydrophobic cleft of TIM8·13 is less capable of holding a hydrophobic polypeptide than the one of TIM9·10.

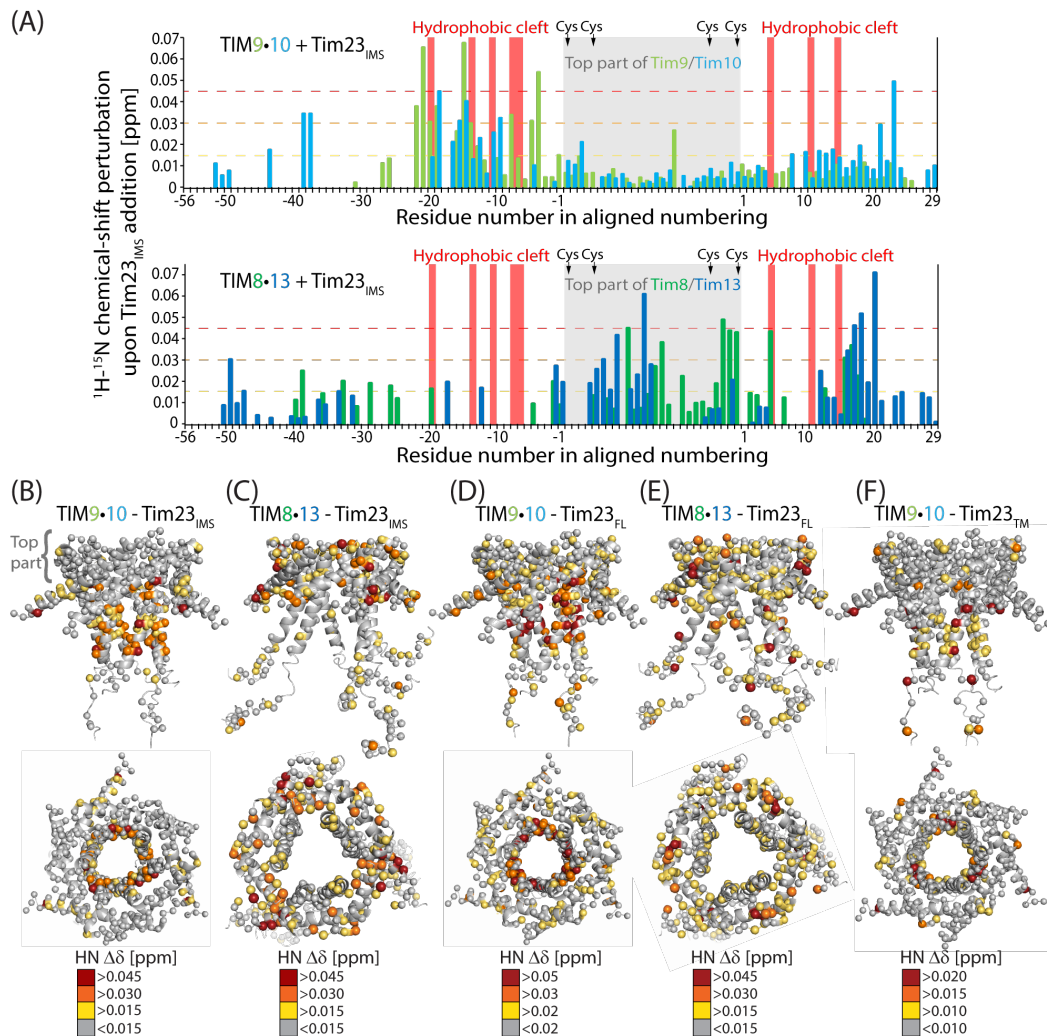


FIGURE 5.4: **Tim23<sub>IMS</sub> and full-length Tim23 differ in their interactions with TIM9-10 and TIM8-13 chaperones.** (A) Chemical-shift perturbations observed upon addition of the Tim23<sub>IMS</sub> fragment to TIM9-10 (top) and TIM8-13 (bottom). The chaperone:Tim23<sub>IMS</sub> ratios were 1:1 (TIM8-13) and 1:3 (TIM9-10). Mapping of Tim23<sub>IMS</sub>-induced CSPs on TIM9-10 (B) and TIM8-13 (C), showing that while the top part of TIM9-10 does not show any significant CSPs, the corresponding part is the main interacting region of TIM8-13. CSP in complexes of TIM9-10 (D) and TIM8-13 (E) bound to full-length Tim23. Tim23<sub>TM</sub>-induced CSP mapped on TIM9-10 (F), showing similar binding as the full-length Tim23.

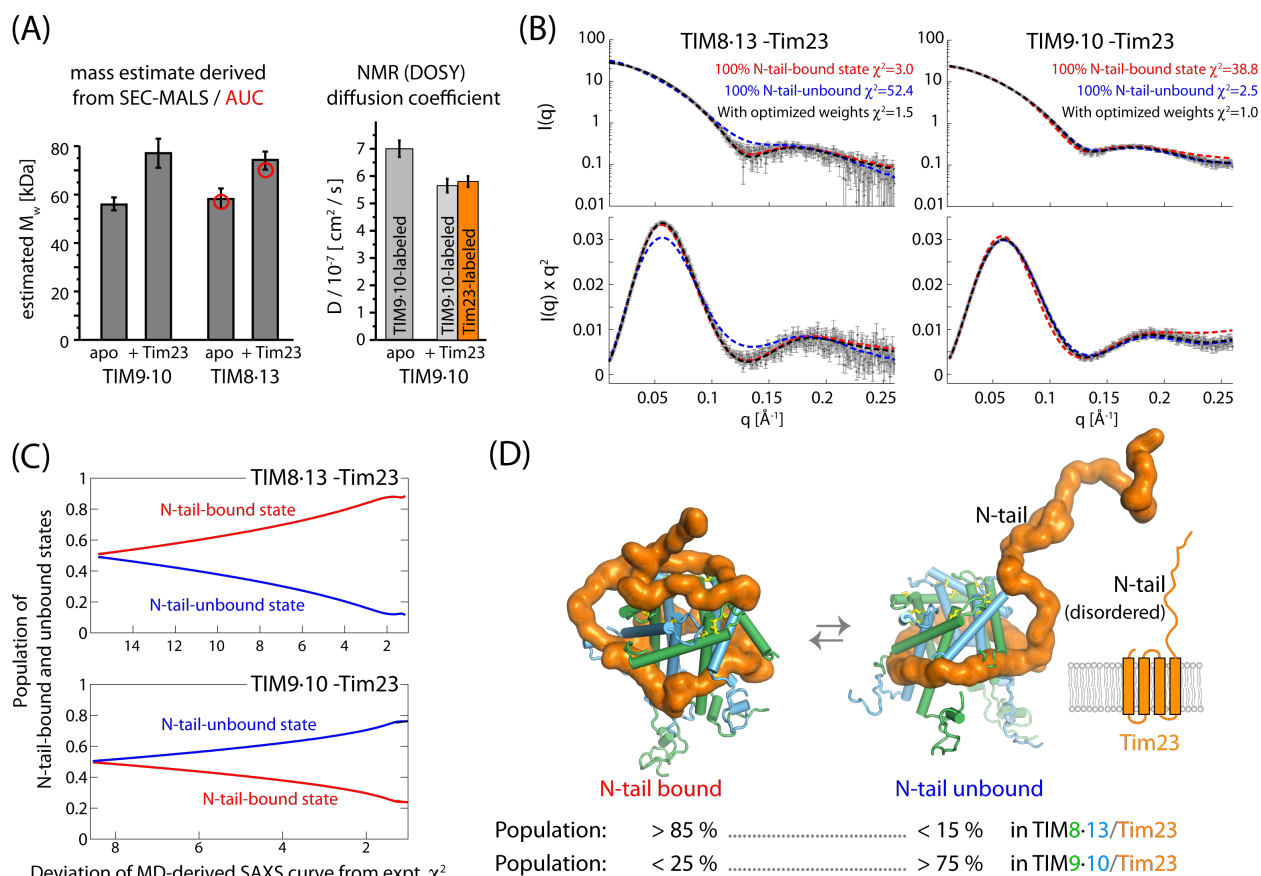
Collectively, NMR, ITC and mutagenesis have revealed that the hydrophobic cleft of both TIM8-13 and TIM9-10 are essential to hold the hydrophobic parts of the clients, and that TIM8-13, but not TIM9-10, additionally interacts with the hydrophilic part of Tim23 to increase its affinity. This interaction, which is mediated by the hydrophilic top part of TIM8-13, reduces the conformational flexibility of Tim23's N-terminal half. The observation that the interaction is driven by hydrophilic contacts supports previous findings of the protein import (Paschen et al., 2000): TIM8-13 was found to interact with hydrophobic membrane precursor only when they were fused to the hydrophilic Tim23<sub>IMS</sub> part.

We have also investigated whether a given full-length Tim23 chain may interact simultaneously with TIM9-10 and TIM8-13, using hydrophobic and hydrophilic interactions, respectively, to form a ternary complex. However, samples containing all three components do not contain detectable amounts of such complexes, and we conclude that the affinity is too low to simultaneously bind two chaperones (fig. S12 in A).

## 5.6 Structural ensembles of chaperone-Tim23 complexes

We integrated the NMR data with further biophysical, structural and numerical techniques to obtain a full structural and dynamical description of the complexes. We first investigated the complex stoichiometry using size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS), NMR-detected diffusion-coefficient measurements, and analytical ultra-centrifugation. These methods, which provide estimates of molecular mass (and shape) from orthogonal physical properties (gel filtration and light scattering; translational diffusion), reveal properties best compatible with a 1:1 (chaperone:precursor) stoichiometry (Fig. 5.5A and fig. S13). Small-angle X-ray scattering (SAXS) data of both TIM9·10-Tim23 and TIM8·13-Tim23 also point to a molecular weight corresponding to a 1:1 complex (SAXS; Fig. 5.5B). This stoichiometry contrasts the 2:1 (chaperone:precursor) stoichiometry for TIM9·10 holding the 35 kDa-large carrier Ggc1 (Weinhäupl et al., 2018) (fig. S2 in A).

Importantly, SAXS provides significantly more information, namely the overall shape of the ensemble of conformations present in solution. Given the flexibility of the complex, this SAXS data is best analyzed by considering explicitly a dynamic ensemble. We used molecular dynamics (MD) simulations to account for the breadth of possible conformations that, collectively, result in the observed scattering. To effectively sample the conformational space of the chaperone-Tim23 complex, we constructed two distinct structural models in which the N-terminal half of Tim23 is either modeled as a floppy unstructured tail or bound to the hydrophilic upper part of chaperone, denoted as 'N-tail unbound' and 'N-tail bound' conformations, respectively. In both models, the hydrophobic C-terminal transmembrane domain of Tim23 is bound to the hydrophobic cleft of the chaperone, as identified by NMR (Fig. 5.4 and fig. S10C,D in A) (Weinhäupl et al., 2018). Initiating from both conformations, explicit-solvent atomistic MD simulations ( $\sim 4.25 \mu\text{s}$  in total) were performed to collect the structures for the 'N-tail unbound' and 'N-tail bound' ensembles. In the case of TIM8·13, the 'N-tail bound' ensemble recapitulates the experimentally observed pattern better than the 'N-tail unbound' ensemble (Fig. 5.3 C and E). We then constructed a mixed ensemble consisting of a mixture of 'N-tail bound', and 'N-tail unbound' states. We used this pool of conformations for further ensemble refinement, with the relative populations of these two ensembles of states as free parameter, using the Bayesian Maximum Entropy (BME) method guided by experimental SAXS data (Bottaro2020 ; Orioli et al., 2020; Weinhäupl et al., 2018). We found that the experimental SAXS data of TIM8·13-Tim23 are very well reproduced when the mixed ensemble has  $>85\%$  of the 'N-tail bound' state (Fig. 5.5C, D). In contrast, the experimental data of TIM9·10-Tim23 are only well reproduced when the TIM9·10-Tim23 ensemble comprises predominantly the 'N-tail unbound' state. These refined ensembles guided by experimental SAXS data are in excellent agreement with the NMR data, which showed that (i) in the TIM9·10-Tim23 complex, the N-terminal part of Tim23 is predominantly free and flexible, and Tim23 makes contacts only to the hydrophobic cleft of the chaperone, while (ii) in TIM8·13-Tim23, the Tim23<sub>IMS</sub> part is largely bound to the upper part of the chaperone (Figs. 5.3 and 5.4).



**FIGURE 5.5: Architecture of the TIM8-13 and TIM9-10 holdases in complex with full-length Tim23.** (A) (Left) Apparent molecular weights of apo and holo chaperone complexes from SEC-MALS, and AUC (red circles). (Right) Translational diffusion coefficients of TIM9-10 (apo) and TIM9-10-Tim23<sub>FL</sub> from NMR DOSY measurements. Two independent samples were used for the complex, in which either the chaperone or the precursor protein was labeled, as indicated. See also fig. S13. (B) Small-angle X-ray scattering curves (top) and Kratky plot representations thereof for the two chaperone-precursor complexes. The lines are SAXS curves calculated from structural ensembles obtained over 4.25  $\mu$ s long MD trajectories, in which the N-terminal half of Tim23 was either in a conformation bound to the top part of the chaperone (red) or in a loose unbound conformation (blue), or from an ensemble in which these two classes of states were present with optimized weights. (C) Goodness of fit of the back-calculated SAXS curves to the experimental SAXS data as a function of the relative weights of the two classes of conformations (bound/unbound). (D) Snapshots of conformations in which Tim23<sub>N-tail</sub> is either bound or unbound, and the best-fit relative weights of the two classes of states as derived from SAXS/MD. More SAXS/MD data and Supplementary Movies 1 and 2 can be found in the online version of this publication DOI: 10.1126/sciadv.abd0263.

The amount of ‘N-tail bound’ relative to ‘N-tail unbound’ states is expected to depend on the affinity of the N-tail of Tim23 to the chaperone. Indeed, the ITC-derived TIM8-13-Tim23<sub>IMS</sub> affinity ( $K_d=66 \mu$ M; Fig. 5.3E) predicts that the population of N-tail-bound states is of the order of 75 - 98 % (see Methods - Chapter 7 for details), in excellent agreement with the MD/SAXS derived value (> 85 %). This good match of data from the Tim23<sub>IMS</sub> fragment and Tim23<sub>FL</sub> suggests that the binding of Tim23’s hydrophilic N-tail does not strongly depend on the presence of the TM parts. The low affinity of the N-tail to TIM9-10, reflected by the inability to detect TIM9-10-Tim23<sub>IMS</sub> binding by ITC, is mirrored by the small population of the ‘N-tail bound’ states in the full-length complex.

To identify the molecular mechanisms underlying the observed differences in N-tail binding, we studied the interactions formed between Tim23 and the chaperones along the MD simulation. An interesting pattern emerges from the analysis of the electrostatic interactions. The top part of TIM8-13 has predominantly polar and negatively charged residues, which are in transient contact with the positive charges of Tim23 N-tail, within a dynamic ensemble of conformations (Fig. 5.6). For example, three key aspartate or glutamate

residues in TIM8·13 appear to be involved in binding of lysine or arginine residues of Tim23<sub>IMS</sub> (Fig. 5.6A). In Tim9, a lysine (K51) is present in the top part, and contributes a positive charge (the equivalent position in TIM8·13 is a non-charged, polar residue) (Fig. 5.6B). We hypothesized that the less complementary electrostatic properties of TIM9·10's top part and Tim23's N-tail, as compared to TIM8·13, may diminish the affinity of the N-tail to TIM9·10.

We attempted to investigate the importance of these charged residues experimentally, and prepared single- and double-mutants that invert the pattern of charged residues. In TIM8·13, we introduced lysine or arginine instead of negatively charged residues, expecting to reduce thereby the affinity to Tim23<sub>IMS</sub>; conversely, in TIM9·10 we introduced negative charges to promote the Tim23<sub>IMS</sub> interaction. However, ITC experiments show that most of these mutants do not significantly differ in their binding affinity to Tim23<sub>IMS</sub> (fig. S9 and table S1). In one of the TIM8·13 mutants, the binding affinity even increases, despite the additional positive charge in the chaperone. These findings suggest that due to the disordered nature of Tim23's N-tail, its binding with the chaperones might not be dominated by a few strong interactions but instead be contributed by a complex interaction network with many weak and widely distributed interactions which are tolerant to introduction of the single point mutants that we explored. The MD ensemble (Fig. 5.6C) due to its limited time scale and force field imperfections (Ahmed, Papaleo, and Lindorff-Larsen, 2018) may only be able to identify a rather small number of key interaction sites. Unlike the case of this hydrophilic interaction, we were able to identify several single-point mutations in the TIM9·10 hydrophobic motif that abrogate the binding, with a strong phenotype (Weinhäupl et al., 2018).

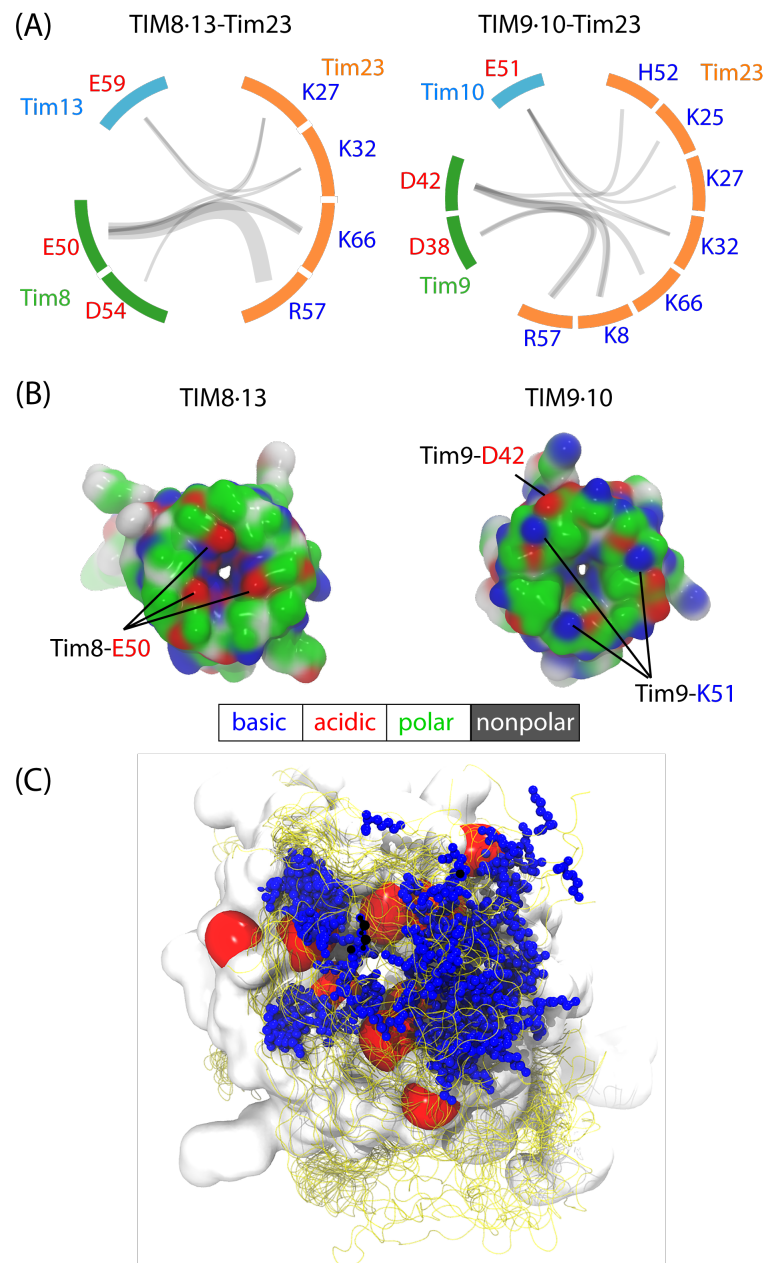


FIGURE 5.6: **Tentative identification of electrostatic interactions from the MD ensemble.** (A) The charged residue pairs forming salt bridges are connected by grey semi-transparent lines whose thickness linearly scales with the frequency of the corresponding salt bridge observed in MD simulations. Although more diverse salt bridges were observed in TIM9-10-Tim23 (10 in TIM9-10-Tim23 and 7 in TIM8-13-Tim23), these salt bridges were on average less stable than the ones in TIM8-13-Tim23, likely resulting in overall weaker interactions. (B) Snapshots of top views of the two chaperones along MD simulations of their holo forms in complex with TIM23. Residues are color-coded according to the scheme reflected below the figure. (C) Ensemble view of the N-tail-bound state of TIM8-13-Tim23. The red surface represents the negatively charged E59 of Tim13 and E50 and D54 of Tim8. Blue stick-and-ball represents the side chain of positively charged residues (K8, K25, K27, K32, R57 and K66) of Tim23, which is shown as an ensemble of 25 structures.

## 5.7 Discussion of the results obtained for the client specificity of mitochondrial TIM chaperones

Transfer chaperones (holdases) need to fulfill two contradicting requirements, holding their clients very tightly to avoid their premature release and aggregation, while at the same time allowing release at the

downstream factor. This apparent contradiction is solved by a subtle balance of multiple individually weak interactions, and a resulting dynamic complex, wherein the precursor protein samples a wide range of different conformations. This ensemble of conformations results in a high overall affinity, yet a downstream foldase/insertase can detach the precursor protein from the chaperone without significant energy barrier (Hiller, 2019). Balancing the interaction strengths is, thus, crucial to chaperone function. Herein, we have revealed a fine-tuning of chaperone-client specificity that involves hydrophobic interactions with the chaperone's binding cleft and additional hydrophilic interactions, mostly mediated by charged residues, with the chaperone's top part. Lower hydrophobicity within the binding cleft of TIM8-13 compared to TIM9-10 arises by overall less hydrophobic residues and a positively charged residue (Lys/Arg) that is highly conserved in Tim8. As a consequence, TIM8-13 is less able to hold the TM parts of its clients than TIM9-10 by ca. one order of magnitude. As we showed, replacement of two charged/polar side chains in TIM8-13's cleft brings TIM8-13 to a similar level as TIM9-10 for holding an all-transmembrane client.

For binding of its native client Tim23, TIM8-13 uses additional hydrophilic interactions to its client's IMS segment, which is ineffective in the TIM9-10-Tim23 interaction. The additional interaction effectively compensates for the lower affinity of TIM8-13 to the client's TM part. In the case of Tim23, this additional interaction involves a sequence stretch of at least 35-40 residues (Fig. 5.3E,G). Remarkably, TIM8-13 has also been shown to be involved in the transport of a  $\text{Ca}^{2+}$ -regulated mitochondrial carrier, the Asp/Glu carrier (Roesch et al., 2004), which has an additional soluble calmodulin-like domain. Whether this soluble domain is folded or disordered while the TM domain is attached to the hydrophobic chaperone cleft remains to be investigated. It is tempting to speculate that interactions between TIM8-13's top part and the calmodulin-like hydrophilic part of these carriers are important for this binding, similarly as for the case of Tim23 (fig. S1). Interestingly, membrane precursor proteins that have been shown *not* to interact with TIM8-13, such as mitochondrial carriers (Ggc, Aac) and Tim17, lack extended hydrophilic stretches, underlining the importance of those parts in binding (fig. S1). From the sequences of known clients and known 'non-clients' of TIM8-13, we propose that a minimum sequence length of about 20-25 residues is required for binding.

The nature of these additional hydrophilic interactions appears to involve primarily charged residues which form a complex and wide-connected interaction network that could be hard to suppress by mutating individual sites.

This study provides a rationale why mitochondria contain two very similar IMS chaperone complexes, the essential TIM9-10 and the non-essential TIM8-13 complex. The observation that this dual system is conserved even in humans suggests that the presence of the TIM8-13 system is not just the result of gene duplication, which appears rather often in yeast. The current results propose that for some substrates (like Tim23, or Asp-Glu carrier; see fig. S1 in A), TIM8-13 can contribute stabilizing interactions with the hydrophilic soluble parts. Our competition experiments have also revealed that mitochondrial membrane precursor proteins may be transferred from one TIM chaperone to the other, opening the possibility that these two chaperones truly cooperate in precursor protein transfer to downstream insertases.

Taken together, our study reveals how a subtle balance of hydrophobic and hydrophilic interactions is used to tune promiscuity versus specificity in molecular chaperones. We propose that a similar balance of interactions determines the clientome of the cellular chaperones.



## Chapter 6

# Results 3. New client proteins of TIM9·10 dependant carrier import pathway

## 6.1 New protein clients of TIM9·10 dependant carrier import pathway

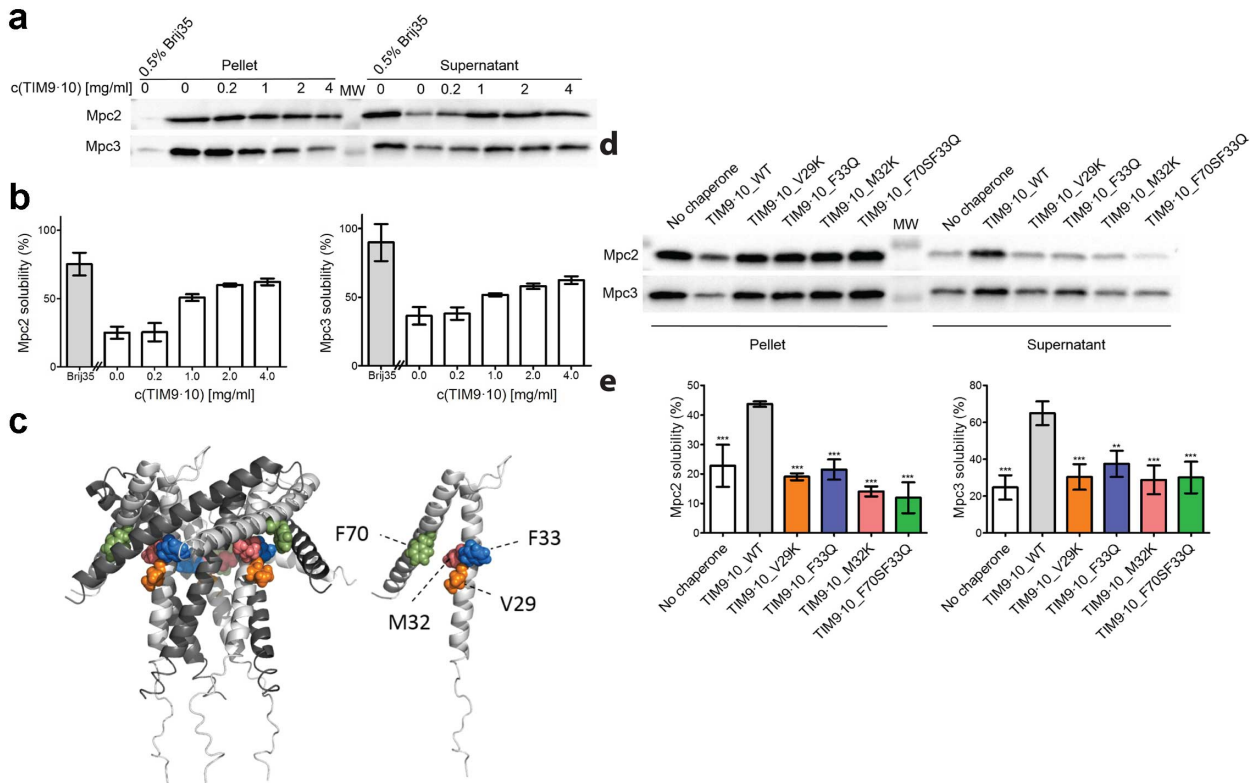
Recently discovered proteins responsible for transport of pyruvate into the matrix of mitochondria, mitochondrial pyruvate carriers (MPC) (Bricker et al., 2012; Herzig, 2012), play a central role in energy metabolism of the eukaryotic cell. The functional carrier is a heterodimer composed of Mpc1 and either Mpc2 or Mpc3 in yeast, and a larger oligomeric assembly cannot be fully excluded (review on mitochondrial pyruvate carriers (Bender and Martinou, 2016)). In contrast to the majority of mitochondrial carriers that belong to the canonical (SLC25) carrier family, Mpc's were shown to contain, in case of Mpc2 and Mpc3, three transmembrane helices where the N-terminus is facing the matrix and the number of transmembrane helices of the Mpc1 protein is either two or three (Bender, Pena, and Martinou, 2015; Tavoulari et al., 2019).

In our collaborative work with the group of Professor Nikolaus Pfanner (Rampelt et al., 2020) we reported that the import pathways of Mpc2 and Mpc3, yeast MPC proteins with an odd number of transmembrane segments and matrix-exposed N-terminus, are imported by the carrier pathway, using the receptor Tom70, small TIM chaperones, and the TIM22 complex. The TIM9·10 complex chaperones MPC proteins through the mitochondrial intermembrane space using conserved hydrophobic motifs that are also required for the interaction with canonical carrier proteins.

Here, I will briefly summarise my contribution to the published work (Rampelt et al., 2020) and the full publication is available in the appendix of this thesis.

To directly determine if the MPC precursors depend on the chaperone function of small TIMs, we synthesized cysteine-free forms of the Mpc2 and Mpc3 precursors in a cell-free translation system (Foshag et al., 2018) and performed an aggregation assay. The majority of the hydrophobic Mpc2 and Mpc3 precursors aggregated in the cell-free system in the absence of detergent (Fig. 5a, b). tesWeinhäupl et al., 2018 showed that the TIM9·10 chaperone prevented the aggregation of a canonical carrier precursor *in vitro*. We thus added recombinantly produced TIM9·10 to the *in vitro* translation reaction and observed a significant improvement of the solubility of Mpc2 and Mpc3 in a chaperone concentration-dependent manner (Fig. 5a, b). Importantly, the positive effect of TIM9·10 on the solubility of MPC precursors was abrogated with Tim10 point mutants in which hydrophobic residues in the binding cleft were replaced by hydrophilic ones (Fig. 5c–e). These mutant forms also disrupt the interaction with carrier precursors (Weinhäupl et al., 2018), suggesting that MPC precursors bind to the same hydrophobic motif of the chaperone as carriers.





**FIGURE 6.1: Interaction of Mpc2 and Mpc3 with the TIM9-10 chaperone *in vitro*.** (a) Cell-free reaction mixtures producing Mpc2 (upper panel) or Mpc3 (lower panel) were supplemented with detergent (Brij35) or different concentrations of recombinantly produced TIM9-10 complex. Immunoblot of the soluble (supernatant) and insoluble (pellet) fractions of the reaction mixtures. (b) Mpc2 and Mpc3 solubility quantification. In the presence of detergent (absence of TIM9-10), both Mpc2 and Mpc3 were largely found in the soluble fraction. In the absence of detergent and chaperone, the majority of Mpc2 and Mpc3 was found in the insoluble fraction. Increasing the concentration of TIM9-10 complex in the cell-free reaction mixture resulted in increased solubility of Mpc2 and Mpc3;  $n = 4-5$  for Mpc2;  $n = 3$  for Mpc3; error bars indicate standard deviation. (c) Structural view of the TIM9-10 complex [26, 68]. In the chaperone complex (left), Tim9 monomers are shown in dark gray and Tim10 in light gray. Altered amino acids of the mutant variants in the TIM9-10 complex (Weinhäupl et al., 2018) are shown as colored spheres. Tim10 monomer (right) and altered amino acids in the hydrophobic cleft of TIM9-10. (d) Immunoblot of the soluble and insoluble fractions of the cell-free reaction mixtures producing Mpc2 or Mpc3 in the absence of TIM chaperones or in the presence of wild-type TIM9-10 (TIM9-10-WT) or mutant variants of Tim10 in the TIM9-10 complex (TIM9-10-V29K, TIM9-10-F33Q, TIM9-10-M32K, TIM9-10-F70SF33Q). (e) Solubility quantification shows solubility of Mpc2 and Mpc3 in the presence of TIM9-10 mutant variants comparable to the reaction condition without added chaperone complex.  $n = 3$ ; error bars indicate standard deviation; \*\*\* and \*\* indicate the significant difference with  $P < 0.001$  and  $P < 0.005$ , respectively, in comparison with the reaction with the WT chaperone. This figure corresponds to Figure 5. of (Rampelt et al., 2020).

## Chapter 7

# Material and methods

### 7.1 Production and purification of the TIM chaperones

Protein subunits assembling mitochondrial intermembrane space chaperone TIM8-13 of *Saccharomyces cerevisiae*, Tim8 and Tim13, were coexpressed in *Escherichia coli* cells from a pET-Duet1 vector. This vector encodes for a tobacco etch virus (TEV) protease cleaving site between the protein sequence and an affinity tag (His<sub>6</sub>-tag) for one of the subunits (Tim13). The same approach was used for TIM9-10 chaperone, including co-expression of the two protein subunits, with a cleavable His<sub>6</sub>-tag on Tim10, as described elsewhere (Weinhäupl et al., 2018).

The TIM9-10 and TIM8-13 chaperone complexes can be obtained by over-expression in either SHuffle T7, which results in soluble protein with correctly formed disulfide bonds, or in BL21(DE3) *E. coli* cells which requires refolding from the inclusion bodies. The proteins obtained with either method have indistinguishable properties (SEC, NMR). For TIM9-10, expression in sHuffle cells results in better yield, while we obtain higher TIM8-13 yields with refolding from BL21(DE3). Accordingly, TIM9-10 and unlabeled TIM8-13 were overexpressed in the SHuffle T7 *E. coli* cells and purified as described previously (Weinhäupl et al., 2018).

Briefly, bacterial cell cultures for protein overexpression were inoculated with the over-night precultures and after reaching optical density at 600 nm (OD<sub>600</sub>) between 0.6 and 0.9 the overexpression was induced with 0.5 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Overexpressing cultures were incubated over-night at 20°C. Collected cell pellets were lysed by sonication, while on ice, in buffer A (50 mM tris(hydroxymethyl)aminomethane (Tris), 150 mM NaCl, pH 7.4) and the soluble protein fraction was recovered by centrifugation at 46 000 g for 30 minutes at 4°C. Soluble protein fraction was applied on the gravity-flow Ni-NTA Agarose (QIAGEN) affinity resin and the chaperone complex was eluted with buffer A supplemented with 300 mM imidazole. Eluted protein sample was concentrated and desalted by dialysis in buffer A after which TEV protease was added in 1:20 (m/m) ratio and the cleaving reaction was incubated over-night, at 4°C, under agitation. Reverse affinity purification step was applied and the (tag-less) chaperone sample was collected in the flow-through and wash fractions. Further size-exclusion purification step was performed on a Superdex S200 HiLoad 16/600 column for the chaperone samples analysed by NMR.

Chaperone proteins used for detection by NMR experiments were expressed in D<sub>2</sub>O M9 minimal medium and either labeled with <sup>15</sup>NH<sub>4</sub>Cl (1 g/liter) and D-[<sup>2</sup>H,<sup>13</sup>C]glucose (2 g/liter) or specifically labeled on isoleucine, alanine, leucine, and valine side chains using a QLAM- A  $\beta$  I  $\delta$ 1 L<sup>proR</sup> V<sup>proR</sup> kit from NMR-Bio (www.nmr-bio.com) according to the manufacturer's instructions. The proteins not detected by NMR in complex samples (i.e., the precursor proteins in complexes directed toward chaperone detection or the chaperone in preprotein-detected experiments) were unlabeled and produced in LB medium.

Over-expression of the isotope-labeled TIM8-13 chaperone complex from the BL21(DE3) *E. coli* cells was induced with 1 mM IPTG, and the cells were incubated for 4 hours at 37°C. Cell pellets were sonicated, and the inclusion body fraction was resuspended sequentially, first in buffer A supplemented with 1% lauryldimethylamine oxide (LDAO) and 1% Triton X-100, then in buffer A supplemented with 1 M NaCl and 1 M urea, and, lastly, in buffer B (50 mM Tris, 250 mM NaCl, pH 8.5). The last pellet fraction was solubilized in buffer B supplemented with 50 mM dithiothreitol and 3 M guanidine-HCl at 4°C overnight. The TIM8-13 complex was refolded by rapid dilution in buffer B containing 5 mM glutathione and 0.5 mM glutathione disulfide. The complex was purified on a Ni-NTA affinity column, and the affinity tag was removed with TEV protease and an additional Ni-NTA purification step.

## 7.2 Production and purification of TIM client proteins

The gene coding for full-length *S. cerevisiae* Tim23 (C98S, C209S, C213A) with a C-terminal His<sub>6</sub>-tag was cloned in the bacterial expression plasmid pET21b(+). The plasmid for expression of the intrinsically disordered N-terminal domain of *S. cerevisiae* Tim23<sub>IMS</sub> (residues 1 to 98) with an N-terminal glutathione S-transferase (GST) tag is described in (Cruz et al., 2010), and the pET10N plasmid encoding Tim23<sub>TM</sub> is described in (Truscott et al., 2001).

The *S. cerevisiae* Ggc1(C222S) construct was designed with a C-terminal His<sub>6</sub>-tag in pET21a expression plasmid, reported earlier (Weinhäupl et al., 2018).

Full-length precursor proteins, Tim23 and Ggc1, were expressed as inclusion bodies from BL21(DE3) cells, at 37°C during 1.5 and 3 hours, respectively, after adding 1 mM IPTG. Precursor proteins were solubilized in buffer A supplemented with 4 M guanidine-HCl for Tim23 and 6 M guanidine-HCl for Ggc1 at 4°C overnight. Precursor proteins were purified by gravity-flow affinity chromatography on Ni-NTA agarose in the same denaturing conditions used for solubilization. Imidazole was removed from the precursor protein sample with dialysis in buffer A supplemented with 4 M guanidine-HCl.

GST-tagged Tim23<sub>IMS</sub> was expressed in the soluble protein fraction from BL21(DE3)Ril+ cells during 4 hours at 25°C, after adding 1 mM IPTG. After sonication of the cell pellets, the soluble protein fraction was incubated with glutathione-agarose resin for 2 hours at 4°C. After washing the unspecifically bound proteins with 10 column volumes (CVs) of buffer A, the GST-tag was cleaved from the Tim23<sub>IMS</sub> by incubating the resin with 1 mg of TEV protease per 50 mg of the precursor protein, at 4°C overnight. Cleaved Tim23<sub>IMS</sub> and the protease were collected in the flow-through, and an additional NiNTA purification step was applied to remove the TEV protease from the protein sample. Soluble Tim23<sub>IMS</sub> was subjected to gel filtration on a Superdex 75 10/300 column and stored in buffer A.

Tim23<sub>TM</sub>, comprising residues 92 to 222, was produced in *E. coli* BL21(DE3)Ril+ during 3 hours at 37°C and purified in denaturing conditions as described for the full-length Tim23.

The proteins not detected by NMR were unlabeled and produced in LB medium. Proteins to be detected by NMR were labeled by growing the bacterial culture in minimal media M9 composed of <sup>15</sup>NH<sub>4</sub>Cl (1 g/liter), D-[<sup>2</sup>H,<sup>13</sup>C]-glucose (2 g/liter), 70 mM Na<sub>2</sub>HPO<sub>4</sub>, 5 mM KH<sub>2</sub>PO<sub>4</sub>, 10 mM KCl, 50 mM HEPES, supplemented with vitamin mix and 0.05 mM FeCl<sub>3</sub>, 0.05 mM ZnSO<sub>4</sub>, 1 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, 0.1 mM MnCl<sub>2</sub>. Chaperone-bound Tim23<sub>FL</sub> was additionally deuterated (produced in D<sub>2</sub>O M9 medium), while the labeled Tim23<sub>IMS</sub> fragment was prepared in H<sub>2</sub>O M9 medium.

## 7.3 Preparation of chaperone-precursor protein complexes

Purified precursor protein, i.e., either full-length Tim23, the trans-membrane Tim23<sub>TM</sub> fragment or Ggc1, was bound to NiNTA resin in 4 M guanidine-HCl. The column was washed with five column volumes (CV) of buffer A supplemented with 4 M guanidine-HCl, and with 5 CV of buffer A. A twofold excess of the chaperone complex was passed through the column twice. The column was washed with 10 CV of buffer A and the precursor-chaperone complex was eluted in 5 CV of buffer A supplemented with 300 mM imidazole. The precursor-chaperone complex was immediately subjected to dialysis against buffer A prior to concentrating on Amicon 30 kDa MWCO centrifugal filters (1000 g). Immediate removal of imidazole was particularly important for the preparation of the less stable Tim23<sub>FL</sub>-TIM9-10 and Tim23<sub>TM</sub>-TIM8-13 complex. Complexes of Tim23<sub>IMS</sub> with TIM8-13 or TIM9-10 were prepared by mixing two purified protein samples, and dialysis against buffer A. Formation of the precursor-chaperone complex was verified by size-exclusion chromatography (SEC) on a Superdex 200 column. The resulting complex was further characterized by SEC coupled to multi-angle light scattering (MALS). TIM8-13 and TIM8-13-Tim23 were furthermore analyzed by analytical ultra-centrifugation (AUC). Both experiments were performed at 10°C in Buffer A. The amount of eluted complex was estimated from the protein concentration, measured absorbance of the sample at 280 nm and the sum of the molecular weights and extinction coefficients of the chaperone and the precursor protein.

### 7.3.1 Competition assays

The first competition assay was performed by adding an equimolar mixture of TIM8-13 and TIM9-10 chaperones to the NiNTA bound precursor protein, Tim23<sub>FL</sub> or Ggc1. After washing the column, precursor-chaperone complex was eluted in Buffer A supplemented with 300 mM imidazole. In the time dependent competition assay, the complex of a precursor protein and one of the chaperones (TIM8-13 or TIM9-10) was prepared, and then an equimolar amount of the other chaperone was added (time point 0). The reaction mixture was incubated at 30°C. After 0.5, 1 and 3 hours, an aliquot of the reaction mixture was taken and (newly formed) precursor-chaperone complex was isolated on a NiNTA affinity column. The difference in the amount of specific chaperone, TIM8-13 or TIM9-10, bound to precursor protein was analysed by SDS-PAGE and liquid chromatography coupled with mass-spectrometry (LC ESI-TOF MS, 6210, Agilent Technologies, at the MS platform, IBS Grenoble). Samples for analysis by MS were heat shocked for 15' at 90°C, resulting in the dissociation and precipitation of the precursor protein, while the apo-chaperones were recovered in the supernatant after cooling the sample and centrifugation for 10' at 39k g. As a reference, samples of precursor proteins, Tim23<sub>FL</sub> and Ggc1, bound to individual chaperone, TIM8-13 or TIM9-10, were prepared and analysed in parallel. To be noted, preparation of the TIM8-13-Ggc1 complex, in quantity sufficient for the analysis, was unsuccessful. To calculate the difference in the amount of specific chaperone bound to precursor protein, normalized areas under the chromatography peaks corresponding to each Tim monomer were used.

## 7.4 Purification of the Tom receptors

### 7.4.1 Production and purification of the Tom70 receptor

The gene coding for cytosolic domain of *S. cerevisiae* Tom70 receptor protein (resn 39-617; UniProtKB - P07213) was inserted into modified/in-house created expression vector named pET-TEV-stop carrying Kanamycine resistance. Cloning of gene of interest between NdeI and XhoI restriction sites yields a protein with N-terminal His<sub>6</sub>-tag followed by the TEV protease cleavage site. Protein was produced in *E. coli*

BL21(DE3) cells by incubating the shaking cultures over-night at 20°C after inducing the overexpression with 0.75 mM IPTG. Cell pellets were resuspended in the lysis buffer (50 mM Tris, 300 mM NaCl, pH 8, 10% glycerole, 5 mM 2-mercaptoethanol ( $\beta$ -ME) with addition of cOmplete - EDTA free protease inhibitor cocktail) and lysed by sonication. Soluble Tom70 protein fraction was eluted from gravity-flow Ni-NTA agarose resin with lysis buffer supplemented with 300 mM imidazole. The elution fractions were concentrated and dialyzed for 2 h at 4°C against 50 mM Tris buffer pH 7.6 with 150 mM NaCl. In the second round of dialysis, using fresh buffer supplemented with 2 mM of dithiothreitol (DTT), 1 mg of TEV protease per 20 mg of the Tom70 protein was added in the dialysis bag and incubated over night at 4°C. Cleaved Tom70 protein was recovered in the flow-through fraction during the gravity-flow Ni-NTA affinity chromatography, while the TEV protease and cleaved His-tag were retained on the resin. Additional purification step by size-exclusion chromatography (SEC) on HiLoad 26/600 Superdex 200 pg column was applied using 20 mM Tris at pH 7.6 and 150 mM NaCl. From the SEC chromatogram of Tom70 purification one main peak could be observed, with a shoulder towards the higher molecular weights.

Protein used for detection by NMR experiments was expressed in  $D_2O$  M9 minimal medium and specifically labeled on isoleucine, alanine, and valine or on threonine, methionine and isoleucine side chains using a TLAM- Ala $^{\beta}$  Ile $^{\delta 1}$  Val $^{proR}$  or TLAM-Ile $^{\delta 1}$  Met $^{\epsilon}$  Thr $^{\gamma}$  kit from NMR-Bio ([www.nmr-bio.com](http://www.nmr-bio.com)) according to the manufacturer's instructions.

#### 7.4.2 Production and purification of the Tom22 receptor

The genes coding for two different sizes of the cytosolic domain of *S. cerevisiae* Tom22 receptor protein (resn 1-97 or resn 1-74; UniProtKB - P49334) were inserted into modified/in-house created expression vector named pET47b-TEV-GB1 carrying Kanamycine resistance. Cloning of the gene of interest between NdeI and XhoI restriction sites yields a protein with a C-terminal GB1 solubility tag followed by the His $_6$ -affinity tag and the TEV protease cleavage site. Protein was produced in *E. coli* BL21(DE3) cells by incubating the shaking cultures for 3-4 hours at 37°C after inducing the overexpression with 0.75 mM IPTG. Cell pellets were resuspended in the lysis buffer (50 mM Tris, 300 mM NaCl, pH 8 with addition of cOmplete - EDTA free protease inhibitor cocktail) and lysed by sonication. Soluble Tom22 protein fractions were eluted from the in-house packed Ni-agarose affinity resin, for the FPLC, applying gradient elution steps with the lysis buffer supplemented with 300 mM imidazole. The elution fractions were concentrated and purified by size-exclusion chromatography on a Superdex75 column in 20 mM Tris pH 7, 100 mM NaCl buffer. Removal of the His $_6$ -GB1 tag was achieved by adding 2 mM of dithiothreitol (DTT), 1 mg of TEV protease per 20 mg of the protein sample and incubating the reaction over night at 4°C. Cleaved Tom22 protein was recovered in the flow-through fraction during the gravity-flow Ni-NTA affinity chromatography, while the TEV protease and cleaved His $_6$ -GB1-tag were retained on the resin. Removing the affinity/solubility tag disables calculating the Tom22's concentration by measuring the absorbance at 280 nm due to absence of aromatic amino acids. The protein sample concentration after removal of the tag was measured by performing the BCA Protein Assay Kit (ThermoFischer Scientific).

#### 7.4.3 Production and purification of the Tom20 receptor

Cytosolic domain construct of *R. norvegicus* Tom20 (UniProtKB - Q62760) comprising residues 51-145 was chosen based on previous structural biology studies (Abe et al., 2000). The *S. cerevisiae* Tom20 (UniProtKB - P35180) cytosolic domain constructs were designed based on sequence alignment of Tom20 proteins with a goal of keeping most of the conserved residues of the cytosolic domain (Figure 4.5). The first construct that was used only for initial biochemical studies and AUC was ScTom20(77-183) fused to N-terminal His $_6$ -affinity tag. This construct was designed to start just before the conserved cytosolic domain of Tom20.

The second construct ScTom20(37-183) begins after the helix breaker (Pro residue) following the N-terminal conserved transmembrane helix (9-28).

The genes coding for the cytosolic domain of yeast Tom20(37-183) and Tom20(51-145) were inserted into modified/in-house created expression vector named pET-TEV-GB1-stop, carrying Kanamycine resistance. Cloning of gene of interest between NdeI and XhoI restriction sites yields a protein with the N-terminal His<sub>6</sub>-tag followed by the GB1 solubility tag and the TEV protease cleavage site. Cloning, production and quality control of the plasmids was performed by GeneCust.

The yeast protein was produced in *E. coli* BL21(DE3) cells while the rat protein was produced in *E. coli* BL21(DE3)Ril+ cells. The shaking cultures were incubated for 3-4 hours at 30°C after inducing the overexpression with 0.75 mM IPTG for the yeast (Sc) Tom20 and the rat (Rn) Tom20 was overexpressed over-night at 20°C after induction with 0.5 mM IPTG. Cell pellets were resuspended in the lysis buffer (50 mM Tris, 300 mM NaCl, pH 8 with addition of cOmplete - EDTA free protease inhibitor cocktail) and lysed by sonication. Soluble Tom20 protein fractions were eluted from the in-house packed Ni-agarose affinity column, used on the FPLC system, with the lysis buffer supplemented with 300 mM imidazole. The elution fractions were dialysed against 20 mM Tris pH 7, 100 mM NaCl buffer and the removal of the His<sub>6</sub>-GB1-tag was achieved by adding 2 mM of dithiothreitol (DTT), 1 mg of TEV protease per 20 mg of the protein sample and incubating the reaction while dialyzing the sample over night at 4°C. Cleaved Tom20 protein was recovered in the flow-through fraction during the gravity-flow Ni-NTA affinity chromatography and the sample was additionally purified by the size-exclusion chromatography on a Superdex75 column in 20 mM Tris pH 7, 100 mM NaCl buffer. Removing the affinity/solubility tag disables calculating the rat Tom20's concentration by measuring the absorbance at 280 nm due to absence of aromatic amino acids. The protein sample concentration after removal of the tag was measured by performing the BCA Protein Assay Kit (ThermoFischer Scientific).

#### 7.4.4 Purification of presequence-containing client protein

Client representative of precursor proteins carrying a presequence used in our studies is a mitochondrial subunit 9 of the F<sub>0</sub> ATP synthase of *Neurospora crassa*. The first 69 residues of this precursor protein carries a mitochondrial matrix targeting sequence termed a presequence, pSu9, and it has been extensively used in *in vitro* mitochondrial import assays (Yamano et al., 2008a; Vitali et al., 2018; Rapaport, Neupert, and Lill, 1997 and reviewed in Maduke and Roise, 1996; Pfanner, 2000). Presequences can form an amphiphilic helix with one hydrophobic side and one positively charged side, without significant sequence conservation between different presequences (Maduke and Roise, 1996). The gene coding for residues 1-69 of pSu9 was cloned by GeneCust (<https://www.genecust.com/>) in customized pET-GB1-TEV stop vector for bacterial protein expression, resulting in an N-terminal His<sub>6</sub>-GB1-tag, cleavable by the TEV protease. This tag ensured stable bacterial expression, ease of purification and an increased solubility. The construct carrying a C-terminal tag was also tested, however the yield of production was significantly lower compared to the construct with the N-terminal tag. The work with the labeled pSu9 construct, pSu9 assignment and the interaction with Tom20, was performed with help from A. Feignier during his five months internship in the laboratory. The GB1-pSu9(1-69) construct was produced and purified from inclusion bodies from BL21(DE3) *E. coli* cells. After refolding from denaturing conditions, the protein showed instability at low temperatures, it could not be frozen. Refolding was performed just before using the sample.

## 7.5 Peptide sequences used in the chaperones and receptors interaction studies

All the presequence (matrix targeting sequences of the mitochondrial precursor proteins) peptides were purchased from GeneCust (<https://www.genecust.com/>) with purity of > 95%.

The presequence of the mitochondrial  $\beta$ -subunit of F1-ATPase from *S. cerevisiae* (UniProtKB - P00830) pF1 $\beta$ (1-31): MVLPRLYTATSRAAFKAAKQSAPLLSTSWKR.

The presequence of the mitochondrial Cytochrome c oxidase subunit 4 from *S. cerevisiae* (UniProtKB - P04037) pCox4(1-25): MLSLRQSIRFFKPATRTLCSRYLL.

The presequence of the rat aldehyde dehydrogenase ALDH2 split in two parts (UniProtID- P11884) pALDH(1-22):MLRAALSTARRGPRLSRLLSAA and pALDH(12-22):PRLSRLLSAA.

The presequence of the subunit 9 of the mitochondrial ATP synthase from *N. crassa* (UniProtKB - P00842) pSu9(1-25):MASTRVLASRLASQMAASAKVARPA.

The presequence of the  $\alpha$ -subunit of the mitochondrial ATP synthase (UniProtKB - P07251) pAtp1(1-35):MLARTAAIRSLRSTLINSTKAARPAALASTRRL, and the internal matrix targeting signal of the same precursor protein Atp1(306-324): DDLSKQAVAYRQLSLLLRR (iMTS-L).

The fragments of human VDAC1 peptide (cyclic or linear VDAC<sub>257-279</sub>) were prepared by solid-phase synthesis as described elsewhere (Jores et al., 2016). All peptides were received lyophilized. Consequently, they were resolubilized in 100% DMSO and step-wise diluted into interactions-compatible buffer resulting in less than 2.5% DMSO in the final protein sample.

The peptide used for photo-induced cross-linking performed by the group of Prof. D. Rapaport, differed from the one used for NMR by the substitution of L263 by a Bpa side chain, as used earlier (Jores et al., 2016; Weinhäupl et al., 2018).

## 7.6 Cell-free protein production

The protocol for *in vitro* production of mitochondrial carrier proteins in presence or absence of TIM chaperones is extracted from (Rampelt et al., 2020) (in the AppendixB.1). Genes coding for *S. cerevisiae* Mpc1(C87A), Mpc2(C86A, C111S), and Mpc3(C87A) were cloned by GeneCust in customized pIVEX2.3d cell-free expression plasmids between NdeI and XhoI cloning sites. Cysteine-free variants were used since previous studies with the chaperoning assay (Weinhäupl et al., 2018) indicated that the presence of Cys residues can lead to enhanced aggregation, likely due to disulfide bond formation. The plasmid codes for the TEV-protease-cleavable N-terminal His<sub>6</sub>-tag. MPC proteins were produced in 50  $\mu$ l cell-free reaction mixtures (Schneider et al., 2010) for 2.5 h at 28°C. The final composition of the cell-free reaction buffer was 0.08 mM rUTP, 0.08 mM rGTP, 0.08 mM rCTP, 0.55 mM HEPES, 0.12 mM ATP, 6.8  $\mu$ M folinic acid, 0.064 mM cyclic AMP, 0.34 mM DTT, 2.75 mM NH<sub>4</sub>OAc, 80mM phosphocreatine, 0.208M potassium glutamate, 10.48 mM magnesium acetate, 1mM of amino acid mix, 1.25  $\mu$ g creatine kinase, 0.25  $\mu$ g T7 polymerase, 20  $\mu$ l S30 *E. coli* extract, 0.5  $\mu$ g plasmid DNA, and 0.175 mg/ml tRNAs. The reaction condition with the detergent contained additionally 0.5% of Brij35. To test the specificity of the binding of MPC proteins by TIM chaperones, the solubility of MPC proteins was monitored at increasing concentration of either TIM8-13 or TIM9-10 complexes. The concentrations of the chaperones in the reaction mixtures were 0, 0.2, 1, 2, and 4 mg/ml. To test the effect

of selected Tim10 mutant variants in the TIM9·10 chaperone complex on the binding and subsequently the solubility of MPC proteins, 4 mg/ml of the TIM9·10-WT, TIM9·10-V29K, TIM9·10-F33Q, TIM9·10-M32K, and TIM9·10-F70SF33Q were used. Chaperone complexes of TIM8·13, TIM9·10, and mutant variants of TIM9·10 for cell-free experiments were expressed and purified as described above. The cell-free reaction was stopped after 2.5 h, and the soluble fraction was separated from the insoluble pellet by centrifugation at 16.800×g. The amount of His-tagged MPC proteins in the soluble fraction and the pellet were quantified from the membranes, after the immunodecoration with anti-His antibody (Sigma-Aldrich monoclonal  $\alpha$ -polyHistidine-peroxidase antibody), as relative band intensities using BioRad ImageLab software. The solubility of the proteins was calculated from at least three experiments for each condition, as a percentage of protein in the supernatant in relation to the total amount of protein in the pellet and supernatant. Significance of the difference in solubility between wild-type TIM9·10 and the mutant variants was analyzed with GraphPad Prism 5 using one-way ANOVA and Tukey's multiple comparison test.

## 7.7 Isothermal Titration Calorimetry experiments

### 7.7.1 ITC experiments with the TIM chaperones

Calorimetric binding experiments of Tim23<sub>IMS</sub> and TIM chaperones were performed using a MicroCal ITC200 instrument (GE Healthcare). Sixteen successive 2.5  $\mu$ l aliquots of 1.15 mM Tim23<sub>IMS</sub> were injected into a sample cell containing 55  $\mu$ M TIM9·10 or TIM8·13. All ITC data were acquired in Buffer A at 20°C. Control experiments included titrating Tim23<sub>IMS</sub> into the Buffer A. The enthalpy accompanying each injection was calculated by integrating the resultant exotherm, which corresponds to the released heat as a function of ligand concentration added at each titration point. ITC data were analysed via the MicroCal Origin software using a single site binding model and nonlinear least squares fit of thermodynamic binding parameters ( $\Delta H$ ,  $K$ , and  $n$ ). An identical procedure was performed for TIM8·13-cytochrome c and TIM9·10-cytochrome c ITC experiments. Cytochrome c was from horse heart (Merck/Sigma-Aldrich). We also performed ITC experiments with the VDAC peptides; no effects could be detected, in line with a millimolar affinity, as already reported for the TIM9·10-cyclic-VDAC<sub>257-279</sub> peptide (Weinhäupl et al., 2018).

### 7.7.2 ITC experiments with the Tom receptors

Calorimetric binding experiments with Tom receptors were performed using a MicroCal ITC200 instrument (GE Healthcare). Ten successive 2.5  $\mu$ l aliquots of titrant were injected into a sample cell containing the analyte. Spacing between the injections was set to 180 sec and the steering speed was 750 rpm. All ITC data were acquired in 20 mM Tris pH 7 buffer with 100 mM NaCl at 20°C. Control experiments included titrating the titrant into the buffer. The enthalpy accompanying each injection was calculated by integrating the resultant exotherm, which corresponds to the released heat as a function of ligand concentration added at each titration point. ITC data were analysed via the MicroCal Origin software using a single site binding model and nonlinear least squares fit of thermodynamic binding parameters ( $\Delta H$ ,  $K$ , and  $n$ ). Exact concentrations of the analyte and titrant used is indicated in the Appendix Table 4.1 with the results of the experiments.

## 7.8 SEC-MALS experiments

SEC-MALS experiments were performed at the Biophysical platform (AUC-PAOL) in Grenoble. The experimental setup comprised an HPLC (Schimadzu, Kyoto, Japan) consisting of a degasser DGU-20AD, an LC-20AD pump, an autosampler SIL20-ACHT, a column oven XL-Therm (WynSep, Sainte Foy d'Aigrefeuille,



France), a communication interface CBM-20A, a UV-Vis detector SPD-M20A, a static light scattering detector miniDawn Treos (Wyatt, Santa-Barbara, USA), a dynamic light scattering detector DynaPro NANOS-TAR, a refractive index detector Optilab rEX. The samples were stored at 4 °C, and a volume of 20, 40, 50 or 90 µl was injected on a Superdex S200, equilibrated at 4 °C; the buffer was 50 mM Tris, 150 mM NaCl filtered at 0.1 µm, at a flow rate of 0.5 ml/min. Bovine serum albumine was used for calibration. Two independent sets of experiments conducted with two different batches of protein samples were highly similar.

## 7.9 Analytical ultra-centrifugation

AUC experiments were performed at 50000 rpm and 10 °C, on an analytical ultracentrifuge XLI, with a rotor Anti-60 and anti-50 (Beckman Coulter, Palo Alto, USA) and double-sector cells of optical path length 12 and 3 mm equipped of Sapphire windows (Nanolytics, Potsdam, DE). Acquisitions were made using absorbance at 250 and 280 nm wave length and interference optics. The reference is the buffer 50 mM Tris, 150 mM NaCl for samples with TIM chaperones and for samples with Tom receptors 20 mM Tris, 150 mM NaCl. The data were processed by Redate software v 1.0.1. The c(s) and Non Interacting Species (NIS) analysis was done with the SEDFIT software, version 15.01b and Gussi 1.2.0, and the Multiwavelength analysis (MWA) with SEDPHAT software version 12.1b.

Both AUC and SEC-MALLS experiments were performed at the platforms of the Grenoble Instruct-ERIC center (ISBG ; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB).

TABLE 7.1: **Protein constructs used during the studies for this thesis.** For each construct indicated are following parameters: molecular weight (MW), extinction coefficients ( $\epsilon$ ) used for protein concentration calculation from the absorption at 280 nm, yield of protein obtained from 1L of LB (or M9) bacterial culture, expression plasmid and *E. coli* cells used for protein overexpression, along with methods of purification used and whether the purification was done in denaturing conditions.

protein	MW [Da]	$\epsilon$ [ $M^{-1}cm^{-1}$ ]	pI	yield/L LB	yield/L M9	plasmid	<i>E. coli</i> cells	1 <sup>st</sup> purification	2 <sup>nd</sup> purification	denaturing conditions	refolding
Tim10	10230.47	4720	5.32								
Tim9	10201.63	1740	8.4	44 mg	25 mg (D <sub>2</sub> O)	pET_duet	Shuffle T7	gravity-flow Ni-NTA	Superdex S200	No	No
Tim8	9763.91	250	5.24								
Tim13	11211.68	10220	8.42	13 mg	4 mg (D <sub>2</sub> O)	pET_duet	BL21(DE3)	gravity-flow Ni-NTA	Superdex S200	No	Yes
Tim23 <sub>TMS</sub>	10424.48	12490	4.35			pGEX	BL21(DE3)Ril+	gravity-flow glutathione-agarose	Superdex 75	No	No
Tim23 <sub>FL</sub> -His <sub>6</sub>	24002.23	23950	7.17	6 mg	4.5 mg (D <sub>2</sub> O)	pET21b(+)	BL21(DE3)	gravity-flow Ni-NTA	-	Yes	No
His <sub>10</sub> -Tim23 <sub>TM</sub>	14994	16960	9.93	2.5 mg	-	pET10N	BL21(DE3)Ril+	gravity-flow Ni-NTA	-	Yes	No
Ggc1-His <sub>6</sub>	34264.55	28420	10.02	48 mg	-	pET21a	BL21(DE3)	gravity-flow Ni-NTA	-	Yes	No
Tom70(39-617)	66395.94	39895	5.08	60 mg	60 mg (D <sub>2</sub> O)	pET-TEV-stop	BL21(DE3)	gravity-flow Ni-NTA	Superdex S200	No	No
Tom70(39-617)PiC1 <sub>TMS-6</sub> His <sub>6</sub>	77270.44	42875	5.5	4 mg/100 mL		pET30c(+)	BL21(DE3)	gravity-flow Ni-NTA	Superdex S200	No	No
Tom70(39-617)Aac2 <sub>TMS-4</sub> His <sub>6</sub>	79226.66	58805	5.64	20 mg/100 mL		pET30c(+)	BL21(DE3)	gravity-flow Ni-NTA	Superdex S200	No	No
Tom22(1-97)_GB1	19 615.28	11 460	4.47	6 mg		pET47b-TEV-GB1	BL21(DE3)	FPLC HisTrap	Superdex 75	No	No
Tom22(1-97)	11 858.85	1 490	3.98	4 mg		-	-	-	-		No
Tom22(1-74)_GB1	17 075.49	11 460	4.36	25 mg		pET47b-TEV-GB1	BL21(DE3)	FPLC HisTrap	Superdex 75	No	No
Tom22(1-74)	9 319.06	1 490	3.83	12 mg		-	-	-	-		No
ScTom20(37-183)	16967.37	7450	5.16	24 mg	14 mg (D <sub>2</sub> O)	pET-GB1-TEV-stop	BL21(DE3)	FPLC HisTrap	Superdex 75	No	No
RnTom20(51-145)	11036.67	1490	4.48	15 mg	12 mg	pET-GB1-TEV-stop	BL21(DE3)Ril+	FPLC HisTrap	Superdex 75	No	No
GB1-pSu9(1-69)	16570.61	12950	9.77		60 mg	pET-GB1-TEV-stop	BL21(DE3)	gravity-flow Ni-NTA	Superdex 75	Yes	Yes
His <sub>6</sub> -Om14 <sub>FL</sub>	17145.11	11920	9.05	15 mg		pET-TEV-stop	BL21(DE3)	gravity-flow Ni-NTA	-	Yes	No
His <sub>6</sub> -Om14 <sub>Cter</sub>	7068.02	5960	9.63	25 mg		pET-TEV-stop	BL21(DE3)	gravity-flow Ni-NTA	-	Yes	No

## 7.10 NMR spectroscopy

All NMR experiments were performed on Bruker Avance-III spectrometers operating at 600, 700, 850 or 950 MHz  $^1\text{H}$  Larmor frequency. The samples were in the NMR buffer with 10% (v/v)  $\text{D}_2\text{O}$ , unless stated differently. All multidimensional NMR data were analyzed with CCPN (version 2 for TIM data and version 3 for Tom data) (Vranken et al., 2005). DOSY data were analyzed with in-house written python scripts. For calculating chemical-shift perturbation data, the contribution of each different nuclei was weighted by the gyromagnetic ratios of the respective nucleus: e.g. the combined  $^1\text{H}$ - $^{15}\text{N}$  CSP was calculated as  $\sqrt{[\text{CSP}_{^1\text{H}}^2 + \text{CSP}_{^{15}\text{N}}^2 \cdot (\gamma_{^{15}\text{N}}/\gamma_{^1\text{H}})]}$ , where the  $\gamma$  are the gyromagnetic ratios.

### 7.10.1 TIM8-13, Tim23<sub>IMS</sub>, Tom22, Tom20, Tom70 and GB1-pSu9(1-69) resonance assignments

For the resonance assignment of TIM8-13, the following experiments were performed : 2D  $^{15}\text{N}$ - $^1\text{H}$ - BEST-TROSY, 3D BEST-TROSY HNCO, 3D BEST-TROSY HNcaCO, 3D BEST-TROSY HNCA, 3D BEST-TROSY HNcoCA, 3D BEST-TROSY HNcocaCB and 3D BEST-TROSY HNcaCB (Schanda2006b; Favier and Brutscher, 2011) and a 3D  $^{15}\text{N}$ -NOESY HSQC. The experiments were performed with a 0.236 mM [ $^2\text{H}$ , $^{15}\text{N}$ , $^{13}\text{C}$ ]-labeled TIM8-13, at 308 K and 333K. The NMR resonance assignment of TIM9-10 was reported earlier (Weinhäupl et al., 2018).

We collected BEST-TROSY HNCA, HNCO and HNcoCA experiments to assign Tim23<sub>IMS</sub>, aided by the previously reported assignment (Cruz et al., 2010).

For the resonance assignment of Tom22 and Tom20 proteins from yeast, the following experiments were performed : 2D  $^{15}\text{N}$ - $^1\text{H}$ - BEST-TROSY, 3D BEST-TROSY HNCO, 3D BEST-TROSY HNcaCO, 3D BEST-TROSY HNCA, 3D BEST-TROSY HNcoCA, 3D BEST-TROSY HNcoCACB, 3D BEST-TROSY HNCACB, 3D hNcocaNH and a 3D hNcacoNH. The experiments with Tom22 were performed with a 0.6-0.9 mM [ $^{15}\text{N}$ , $^{13}\text{C}$ ]-labeled Tom22, at 298K and pH 6. The experiments for Tom20 were performed with 0.5-0.6 mM [ $^2\text{H}$ , $^{15}\text{N}$ , $^{13}\text{C}$ ]-labeled Tom20, at 293K and pH 7.

We collected BEST-TROSY HN, HNCA, HNcoCA, HNCO and HNcoCA experiments to assign [ $^{15}\text{N}$ , $^{13}\text{C}$ ]-labeled Tom20 protein from rat, aided by the previously reported assignment (Abe et al., 2000). The side chain assignment of the rat Tom20 was also transferred on the recorded 3D HCCH-TOCSY.

For the resonance assignment of methyl groups of Tom70, the 2D  $^{13}\text{C}$ - $^1\text{H}$ - SOFAST HSQC experiments were recorded for the wild-type and all the variants with single amino-acid exchanged. The experiments were performed with a 0.2-0.8 mM [ $\text{U}$ - $^2\text{H}$ , $\text{U}$ - $^{12}\text{C}$ ],[ $^{13}\text{C}$ , $^1\text{H}$ ]-Ile $^{\delta 1}$  Met $^{\epsilon}$  and Thr $^{\gamma}$  labeled Tom70, at 298K and pH 7.4.

### 7.10.2 NMR titration experiments

#### VDAC titration experiments

Cyclic hVDAC<sub>1257-279</sub> peptide was synthesized and lyophilized as described elsewhere (Jores et al., 2016). The peptide was dissolved in DMSO, and the DMSO concentration was reduced to 10% by step-wise addition of NMR buffer (1:1 in each step). Chaperone, TIM9-10 or TIM8-13, in buffer A was added to yield a final DMSO concentration of 6% and a chaperone concentration of 0.15 mM (TIM9-10) or 0.1 mM (TIM8-13). Combined  $^{15}\text{N}$ - $^1\text{H}$  chemical shift-perturbation (CSP) was calculated from the chemical shifts obtained from the  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra of the complex samples with molar ratio of 1:4 for TIM9-10:VDAC, and ratio of

1:5 for TIM8-13:VDAC, in comparison to the chemical shifts from the apo-chaperone spectrum. The NMR experiments were performed at 308K.

### Tim23<sub>IMS</sub> titration experiments

For each titration point individual samples were prepared by mixing two soluble protein samples, and monitored using <sup>15</sup>N-<sup>1</sup>H- BEST-TROSY HSQC experiments at 283K (for Tim23 observed experiment) or at 308K (for chaperone observed experiments). Titration samples with 100 μM [<sup>15</sup>N]-labeled Tim23<sub>IMS</sub> with the molar ratios for Tim23<sub>IMS</sub>:TIM8-13 from 1:0 to 1:4, and the molar ratios for Tim23<sub>IMS</sub>:TIM9-10 from 1:0 to 1:5, were used. For the chaperone observed experiments, used samples contained 200 μM [<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N]-labeled TIM8-13 with molar ratios of Tim23<sub>IMS</sub> 1:0 and 1:1, and 350 μM [<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N]-labeled TIM9-10 with 1:0 and 1:3 molar ratios of Tim23<sub>IMS</sub>.

### Titration experiments with Tom20, Tom22 and the presequences

The presequence peptides were dissolved in DMSO, and the DMSO concentration was reduced to 6% by step-wise addition of NMR buffer (1:1 in each step). The Tom22 or Tom20, in 20 mM K-Pi, 20 mM KCl, pH 6.5 buffer was added to yield a final DMSO concentration of less than 2.5% and a receptor concentration of 0.1-0.15 mM. Combined <sup>15</sup>N-<sup>1</sup>H chemical shift-perturbation (CSP) was calculated from the chemical shifts obtained from the <sup>15</sup>N-<sup>1</sup>H BEST-TROSY spectra of the complex samples, in comparison to the chemical shifts from the apo-receptor spectrum. The NMR experiments were performed at 293K.

### 7.10.3 NMR experiments with the Tim23<sub>FL</sub>

Complexes of the chaperones with the full-length Tim23 were prepared as indicated above (Preparation of chaperone-precursor protein complexes). Peak positions (chemical shifts) of the amide backbone sites of TIM8-13, apo- and in the complex with Tim23<sub>FL</sub>, were obtained from the <sup>1</sup>H-<sup>15</sup>N HSQC experiments at 308K, with 120 μM [<sup>13</sup>CH<sub>3</sub>-ILV]-TIM8-13-Tim23<sub>FL</sub> sample. Similarly, to calculate combined <sup>15</sup>N-<sup>1</sup>H and <sup>13</sup>C-<sup>1</sup>H CSPs, chemical shifts of the amide backbone and ILVA-<sup>13</sup>CH<sub>3</sub> groups of TIM9-10, apo- and in the complex with Tim23<sub>FL</sub>, were obtained from the <sup>1</sup>H-<sup>15</sup>N HSQC and <sup>1</sup>H-<sup>13</sup>C HMQC experiments at 288K. Sample of the [<sup>13</sup>CH<sub>3</sub>-ILVA]-TIM9-10 with the Tim23<sub>FL</sub> was at 140 μM concentration. For the CSP calculations with the complexes of [<sup>2</sup>H-<sup>15</sup>N]-labelled Tim23<sub>FL</sub> and the chaperones (190 μM complex with TIM8-13, and 61 μM complex with TIM9-10), chemical shifts from <sup>1</sup>H-<sup>15</sup>N HSQC experiments at 288K were used in comparison to the chemical shifts of the apo-Tim23<sub>IMS</sub>.

### 7.10.4 Diffusion ordered spectroscopy

Diffusion-ordered NMR spectroscopy (DOSY) experiments were performed at 288K and 600 MHz <sup>1</sup>H Larmor frequency. Diffusion constants were derived from a series of one-dimensional <sup>1</sup>H spectra either over the methyls region (methyl-selective DOSY experiments, for <sup>13</sup>CH<sub>3</sub>-ILVA-labeled apo- and Tim23<sub>FL</sub> bound TIM9-10) or over the amides region (for [<sup>15</sup>N]Tim23<sub>FL</sub>-TIM9-10). Diffusion coefficients were obtained from fitting integrated 1D intensities as a function of the gradient strength at constant diffusion delay.

## 7.11 Small-angle X-ray scattering data collection and analysis

SAXS data were collected at ESRF BM29 beam line (Pernot et al., 2013) with a Pilatus 1M detector (Dectris) at a distance of 2.872 m from the 1.8 mm diameter flow-through capillary. Data on TIM8-13 were collected in a batch mode. The X-ray energy was 12.5 keV and the accessible q-range 0.032 nm<sup>-1</sup> to 4.9 nm<sup>-1</sup>. The

incoming flux at the sample position was in the order of 1012 photons/s in 700x700 mm<sup>2</sup>. All images were automatically azimuthally averaged with pyFAI (Ashiotis et al., 2015). SAXS data of pure TIM8·13 was collected at 1, 2.5 and 5 mg/mL using the BioSAXS sample changer (Round et al., 2015). Ten frames of one second were collected for each concentration. Exposures with radiation damage were discarded, the remaining frames averaged and the background was subtracted by an online processing pipeline (Brennich et al., 2016). Data from the three concentrations were merged following standard procedures to create an idealized scattering curve, using Primus from the ATSAS package (Petoukhov et al., 2012). The pair distribution function  $p(r)$  was calculated using GNOM (Svergun, 1992).

Online purification of the TIM8·13–Tim23<sub>FL</sub> and TIM9·10–Tim23<sub>FL</sub> complexes using gel-filtration column (HiLoad 16/600 Superdex S200 PG) was performed with a high pressure liquid chromatography (HPLC) system (Shimadzu, France), as described in reference (Brennich, Round, and Hutin, 2017). The HPLC system was directly coupled to the flow-through capillary of SAXS exposure unit. The flow rate for all online experiments was 0.2 mL/min. Data collection was performed continuously throughout the chromatography run at a frame rate of 1 Hz. All SAXS data have been deposited on SASBDB.

## 7.12 Calculations of affinities and populations

### 7.12.1 Estimation of the population of Tim23<sub>IMS</sub>-bound states from ITC-derived $K_d$

We attempted to link the ITC-derived dissociation constant of the Tim23<sub>IMS</sub> fragment to the populations of bound and unbound states in the Tim23<sub>FL</sub>-chaperone complexes, using a rationale akin to the one outlined earlier for binding of disordered proteins to two sub-sites (Zhou, Pang, and Lu, 2012). Briefly, we treat the N-terminal tail of Tim23 as a ligand, and the remaining bound complex as the target protein, then the relationship between the population of the bound state ( $P_{\text{bound}}$ ) and the binding affinity ( $K_d$ ) can be written as  $P_{\text{bound}}/(1-P_{\text{bound}}) = C_{\text{eff}}/K_d$  where  $C_{\text{eff}}$  is the effective concentration of the disordered N-tail which was estimated to be between 0.2-3 mM from the MD simulations, resulting in the estimation of  $P_{\text{bound}}$  to be between 75% and 98%.

#### Estimation of the $K_d$ ratio from competition assays

Determining dissociation constants of TIM chaperones to its insoluble client proteins is hampered by the impossibility to form the complexes by solution methods such as titration, as it requires the pull-down method outlined in Fig. 5.1A. Nonetheless, the amount of TIM8·13-Tim23 and TIM9·10-Tim23 complexes obtained in the competition assays (Fig. 5.1) can provide an estimate of the relative affinities. The dissociation constants can be written from the concentrations as follows:

$$K_d^{\text{TIM8}\cdot\text{13-Tim23}} = \frac{[\text{TIM8}\cdot\text{13}] \times [\text{Tim23}]}{[\text{TIM8}\cdot\text{13-Tim23}]}$$

$$K_d^{\text{TIM9}\cdot\text{10-Tim23}} = \frac{[\text{TIM9}\cdot\text{10}] \times [\text{Tim23}]}{[\text{TIM9}\cdot\text{10-Tim23}]}$$

where  $[\text{TIM8}\cdot\text{13-Tim23}]$  denotes the concentration of the formed chaperone-precursor complex, and  $[\text{TIM8}\cdot\text{13}]$  and  $[\text{Tim23}]$  are the concentrations of free chaperone and precursor protein in solution. The latter is negligible, as no free precursor protein is eluted from the column (some aggregated precursor protein was removed from the equilibrium). Both chaperones have been applied at the same concentration  $c_0 = [\text{TIM8}\cdot\text{13}] + [\text{TIM8}\cdot\text{13-Tim23}] = [\text{TIM9}\cdot\text{10}] + [\text{TIM9}\cdot\text{10-Tim23}]$  to the resin-bound precursor protein that was present at a concentration

$b_0 = [\text{Tim23}] + [\text{TIM9}\cdot\text{10-Tim23}] + [\text{TIM8}\cdot\text{13-Tim23}]$ . Using the ratio of formed complex obtained in the competition assay,

$r = [\text{TIM9}\cdot\text{10-Tim23}] / [\text{TIM8}\cdot\text{13-Tim23}]$  leads to

$$\frac{K_d^{\text{TIM9}\cdot\text{10-Tim23}}}{K_d^{\text{TIM8}\cdot\text{13-Tim23}}} = \frac{c_0 \times (1 + r) - r \times b_0}{[c_0 \times (1 + r) - b_0] \times r}$$

The experimental protocol does not allow to determine with precision the concentrations of precursor protein ( $b_0$ ) and each chaperone ( $c_0$ ), as the former is bound to a resin. As the chaperone was added in excess, and some of the precursor protein precipitated on the column, we can safely assume  $c_0 \leq b_0$ . With an experimentally found ratio of formed complexes of  $r = 5$  and assuming that  $c_0/b_0$  assumes the values of 1-5, the  $K_d$  ratio falls in the range of 1:25 to 1:6, i.e. ca. one order of magnitude.

## 7.13 NMR relaxation experiments

T2 relaxation experiments for characterization of backbone protein dynamics consist of multiple 2D  $^{15}\text{N}^1\text{H}$ -TROSY HSQC experiments with varying relaxation delays implemented in the pulse sequence.  $^{15}\text{N}$  R2 relaxation rate constants were derived from these 2D  $^{15}\text{N}^1\text{H}$ -TROSY HSQC experiments. Relaxation delays for T2 experiments were ranging from 0 to 85 ms.

The hetNOE values were derived from the difference in two  $^{15}\text{N}^1\text{H}$  - NOE experiments, the reference and the saturated experiment. The NOE experiment incorporates through-space magnetization transfer via the dipolar coupling from the  $^1\text{H}$  to the  $^{15}\text{N}$  nucleus close in space, providing information regarding the motions of the individual N-H bond vectors (motions on the ps-ns). In saturated experiment, protons are saturated, hence the ratio of peak intensities between the reference and saturated experiment corresponds to the values of the steady-state heteronuclear NOEs.

## 7.14 Paramagnetic relaxation enhancement experiments (PRE)

The solvent accessible  $\epsilon$ -amino groups of the Tom22(1-97) lysine residues were labeled with OXYL-1-NHS (1-oxyl-2,2,5,5-tetramethylpyrrolidine-3-carboxylate-N-hydroxysuccinimide ester, Toronto Research Chemicals). Using the NAP-5 desalting column, the buffer of Tom22(1-97) was exchanged for the labeling buffer (10 mM  $\text{Na}_2\text{CO}_3$ , pH 9.2). A stock solution of the OXYL-1-NHS was prepared by first dissolving the chemical into 100% dimethyl sulfoxide (DMSO). A six-fold molar excess of OXYL-1-NHS was added to the protein solution (in the final reaction 4.1 mM of the OXYL-1-NHS and 1.1% DMSO) and the reaction was incubated for 1 hour at room temperature, followed by 4 hours at 4°C. The excess of spin label was removed by dialysis in pH 6.5, 20 mM K-Pi buffer with 20 mM KCl. The lysine labeling was verified by mass spectrometry where 5 different labeled species were observed. Species with addition of 1x 166Da, 2x 166Da, 3x 166Da 4x 166Da and 5x 166Da correspond to the label added on one, two, three, four or all lysines of Tom22(1-97).

We measured paramagnetic relaxation (PRE) enhancement values, to provide qualitative information on the interactions between the  $^{15}\text{N}^1\text{H}$ -labeled rat Tom20 and the paramagnetically labeled yeast Tom22. Peak intensities in 2D HN-TROSY NMR experiment recorded at 20°C with a RnTom20(51-145) mixed with two fold excess of OXYL-1-NHS- labeled Tom22(1-97) were compared to the reference sample obtained by reducing the OXYL-1-NHS- labeled sample by incubation with sodium ascorbate (2.2 mM) for 24 hours at 4°C. This comparison provided information about the significant peak broadening between the oxidized

and reduced RnTom20-Tom22 sample; the residues marked as "int.ratio < 0.02" in the Figure4.13C. and D. Additional comparison with the apo RnTom20(51-145) sample was performed which provided the information about the RnTom20(51-145) peaks that were not recovered by reducing the OXYL-1-NHS; the peaks labeled as "gone peaks"/"no peaks in PRE" in the Figure4.13C. and D.

## 7.15 Calculating affinities from NMR titration experiments

For calculating the affinities from the NMR titration experiments, the TITAN software package was used (TITration ANalysis; [www.nmr-titan.com](http://www.nmr-titan.com); (Waudby et al., 2016)), using resource for biomolecular NMR - NMRbox (<https://nmrbox.org/>) (Maciejewski et al., 2017).

The 2D HN spectra of the titration series were processed with NMRPipe using exactly the same processing parameters for each spectrum. The experimental and processing parameters are read by the TITAN to simulate spectra in the presence of chemical exchange. TITAN software analyses the 2D lineshapes of selected peaks by fitting the experimental lineshapes into the simulated ones. The bootstrap method, implemented within the software, was used to calculate error estimates. The fitting was run assuming the two-state ligand binding model ( $P + L \rightleftharpoons PL$ ) and in each analysis 10-20 residues were selected for the fitting (regions of interest, ROI). Detailed protocol of NMR titration analysis by TITAN can be found in Chapter 24 of (Kragelund and Skriver, 2020). Example of fitting results can be seen in Figure7.1.

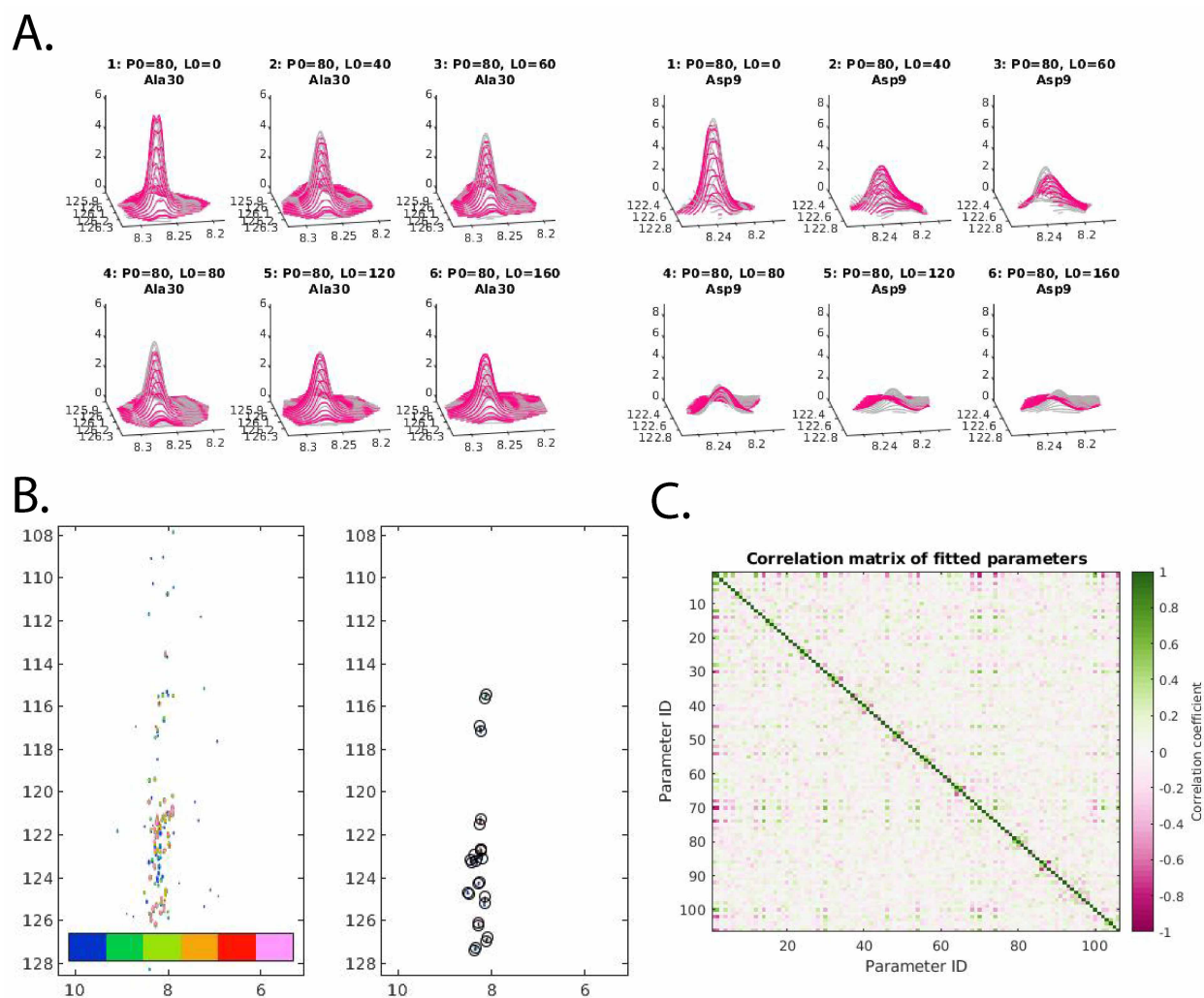


FIGURE 7.1: Analysis of NMR titration experiments of  $^{15}\text{N}$ -Tom22(1-74)-Xdj1(CTD1) interaction by TITAN software. **A.** Three dimensional view of peaks (Ala30 and Asp9) simulated by "virtual spectrometer" (in pink) and of the observed, experimental data (in gray). **B.** Overlay of all spectra of each titration point on the left hand side and selected regions of interest (ROI) around peaks affected by titration. **C.** Correlation matrix of the parameters fitted by bootstrap error analysis. The parameter are: (i) the dissociation constant  $K_d$ , (ii) the dissociation rate  $k_{off}$  (for a two state model) and (iii) the linewidths in the direct and indirect dimension for each state (free and bound) of each ROI.





## Appendix A

# Appendix A

### Data availability

The chemical shift assignments of TIM8·13 have been deposited in the BioMagResBank ([www.bmrb.wisc.edu](http://www.bmrb.wisc.edu)) under accession number 50213. All MD models and SAXS data have been deposited in the SASBDB ([www.sasbdb.org](http://www.sasbdb.org)) under accession numbers SASDH89 (TIM8·13-Tim23), SASDJP4 (TIM9·10-Tim23), SAS-DJQ4 (TIM8·13) and SASDEF2 (TIM9·10 (Weinhäupl et al., 2018)). All chemical-shift perturbation data have been deposited on Mendeley data (<http://dx.doi.org/10.17632/8cr8rvtdm.1>).

### Supplementary materials for Structural basis of client specificity in mitochondrial membrane-protein chaperones

Fig. S1: Kyte-Doolittle hydrophobicity of membrane precursor protein.

Fig. S2: Ensemble representation of the structure of TIM9·10 (blue/green) holding full-length Ggc.

Fig. S3: NMR spectra showing that TIM8·13 and TIM9·10 do not form mixed complexes.

Fig. S4: LC/ESI-TOF MS analysis used to quantify amounts of chaperone complexes.

Fig. S5: Sequence alignment of the small Tims across eukaryotes.

Fig. S6: Pull-down experiment of Tim23<sub>FL</sub> with wild-type TIM9·10, wild-type TIM8·13 and TIM8·13<sub>K30F,S36L</sub>.

Fig. S7: Local order parameters in TIM8·13 and TIM9·10 from TALOS.

Fig. S8: Chemical-shift perturbations on TIM8·13 upon addition of the linear VDAC<sub>257-279</sub> peptide.

Fig. S9: ITC curves of interactions of wild-type and mutant chaperones.

Fig. S10: NMR data showing the Tim23<sub>FL</sub> interactions with TIM9·10 and TIM8·13.

Fig. S11: NMR data showing the interaction of Tim23<sub>TM</sub> with TIM9·10.

Fig. S12: SEC-MALS data of samples containing TIM9·10, Tim23<sub>FL</sub> and TIM8·13.

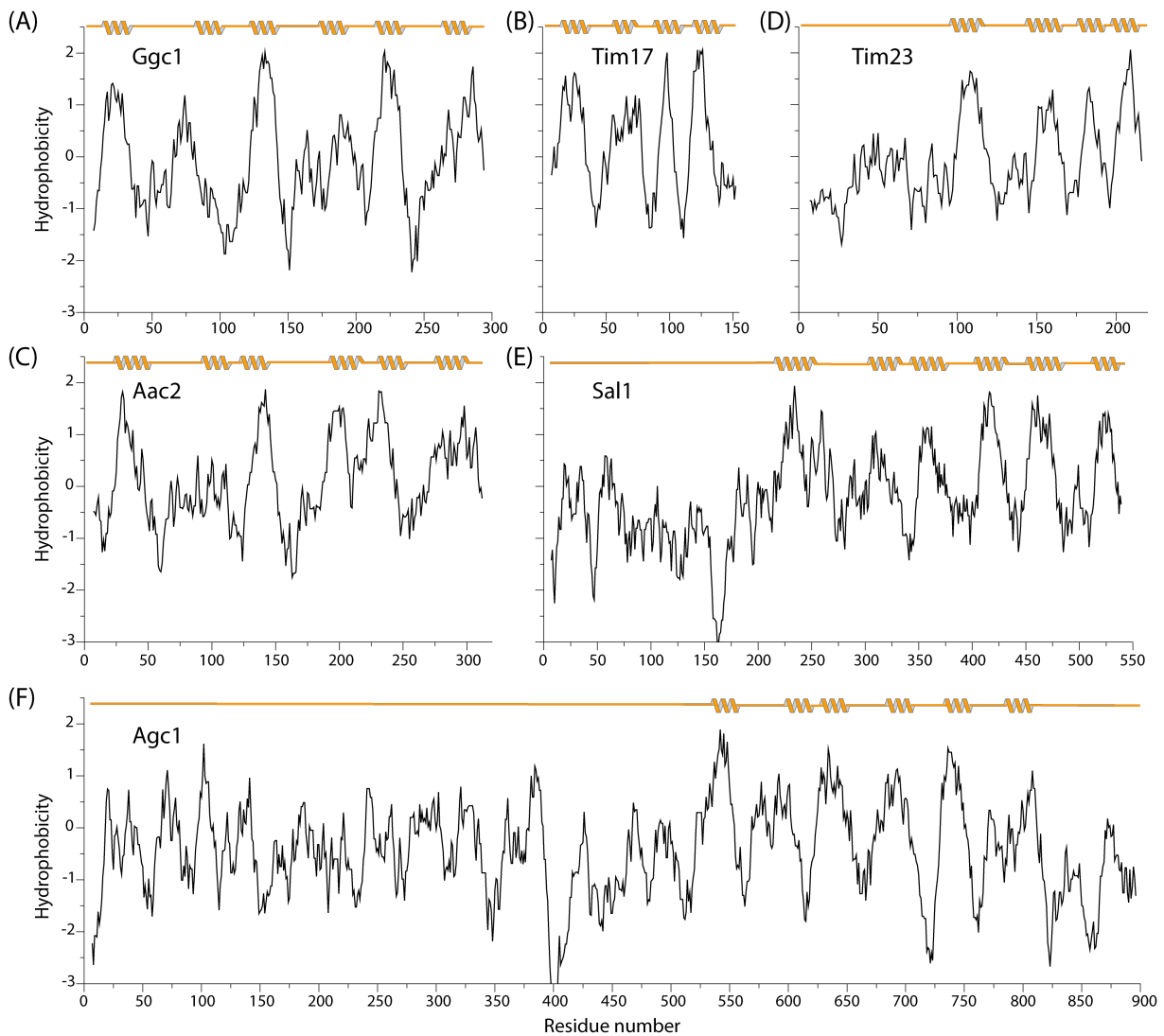
Fig. S13: Experimental characterization of the size of TIM chaperones and their precursor protein complexes.

Table S1: Thermodynamic parameters of TIM8·13 - Tim23<sub>IMS</sub> interaction from ITC.

### A.1 Supporting information for Structural basis of client specificity in mitochondrial membrane-protein chaperones

Protein	$\chi^2 / \text{DoF}$	N	K [M <sup>-1</sup> ]	- $\Delta H$ [cal/mol]	$\Delta S$ [cal/mol/deg]	K <sub>d</sub> [M]
TIM8·13 <sub>WT</sub>	592.9	2.02 ± 0.09	(1.38 ± 0.14) × 10 <sup>4</sup>	3606 ± 241	6.64	66 *
TIM8·13 <sub>E50K</sub>	789.5	2.34 ± 0.06	(2.27 ± 0.27) × 10 <sup>4</sup>	2285 ± 104	12.1	44
TIM8·13 <sub>K47D</sub>	3965	2.19 ± 0.17	(1.49 ± 0.37) × 10 <sup>4</sup>	3435 ± 455	7.37	67
TIM8·13 <sub>E59K,E50K</sub>	1195	2.50 ± 0.05	(3.44 ± 0.44) × 10 <sup>4</sup>	2120 ± 74	13.5	42 *
TIM8·13 <sub>E59K,D54T</sub>	4112	2.05 ± 0.06	(1.93 ± 0.53) × 10 <sup>5</sup>	1327 ± 53	19.7	10 *
TIM8·13 <sub>D54R</sub>	1484	1.83 ± 0.12	(1.27 ± 0.16) × 10 <sup>4</sup>	4964 ± 469	1.83	79

**TABLE S1: Thermodynamic parameters of TIM8·13 - Tim23<sub>IMS</sub> interaction from isothermal titration calorimetry measurements for wild-type and mutant TIM8·13 variants.** The mutants have selected charge inversions in the top region of the chaperone. Using the 'One Set of Sites' model, the fitted parameters were: N (number of sites), K (binding constant in M<sup>-1</sup>) and  $\Delta H$  (heat change in cal/mol). Entropy change  $\Delta S$  (in cal/mole/deg) was calculated from  $\Delta H$  and K after the fitting, and dissociation constant K<sub>d</sub> (in M) was calculated as the reciprocal of the binding constant. The goodness of the fit is reported as the sum of the squares of the deviations of the theoretical curves from the experimental points ( $\chi^2$ ) divided by the degrees of freedom (DoF, the total number of experimental points minus the number of adjustable parameters).



**FIGURE S1: Kyte-Doolittle hydrophobicity of different membrane precursor proteins.** Shown are hydrophobicity predictions of proteins known not to bind to TIM8-13 (Ggc1 (A), Tim17 (B), Aac2 (C)) and proteins known to bind (Tim23 (D)). Sal1 (E) is putatively a client of TIM8-13, as the human form of Sal1, Citrin, is known to depend on the human homolog of TIM8-13). The Glu-Asp carrier Agc1 (F) is a further known TIM8-13 client. The Kyte-Doolittle hydrophobicity (Kyte and Doolittle, 1982) has been determined with the web server of ExPASy, using a window size of 13 and the standard linear weight variation model. Trans-membrane helices (but not other secondary structure elements) are indicated above each plot, as determined from either UniProt or a modelling with SwissModeller (for Sal1, using the structure of Aac2 as a template).

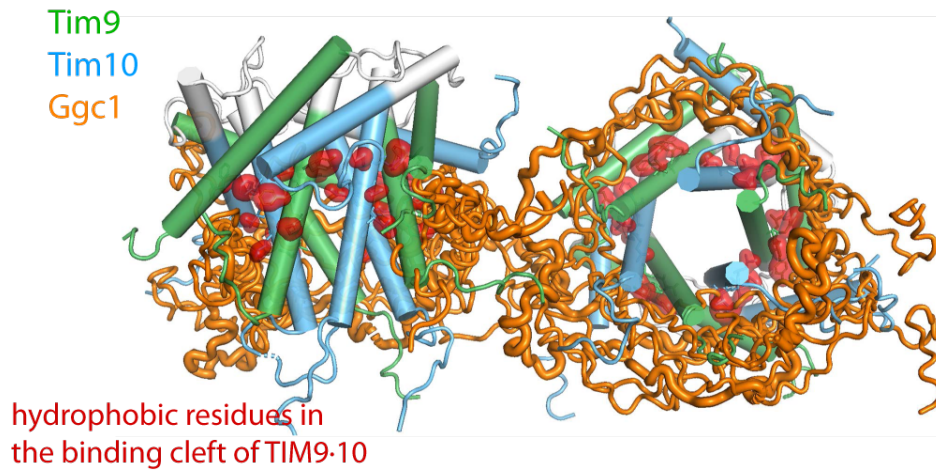


FIGURE S2: **The TIM9-10 chaperones binds Ggc1 in a highly dynamic complex.** Ensemble representation of the structure of TIM9-10 (blue/green) holding full-length Ggc1 (orange), as reported in ref. (Weinhäupl et al., 2018). In contrast to the complexes presented in this study, the complex has a 2:1 stoichiometry, and essentially the entire precursor protein is located to the hydrophobic binding cleft of the chaperone.

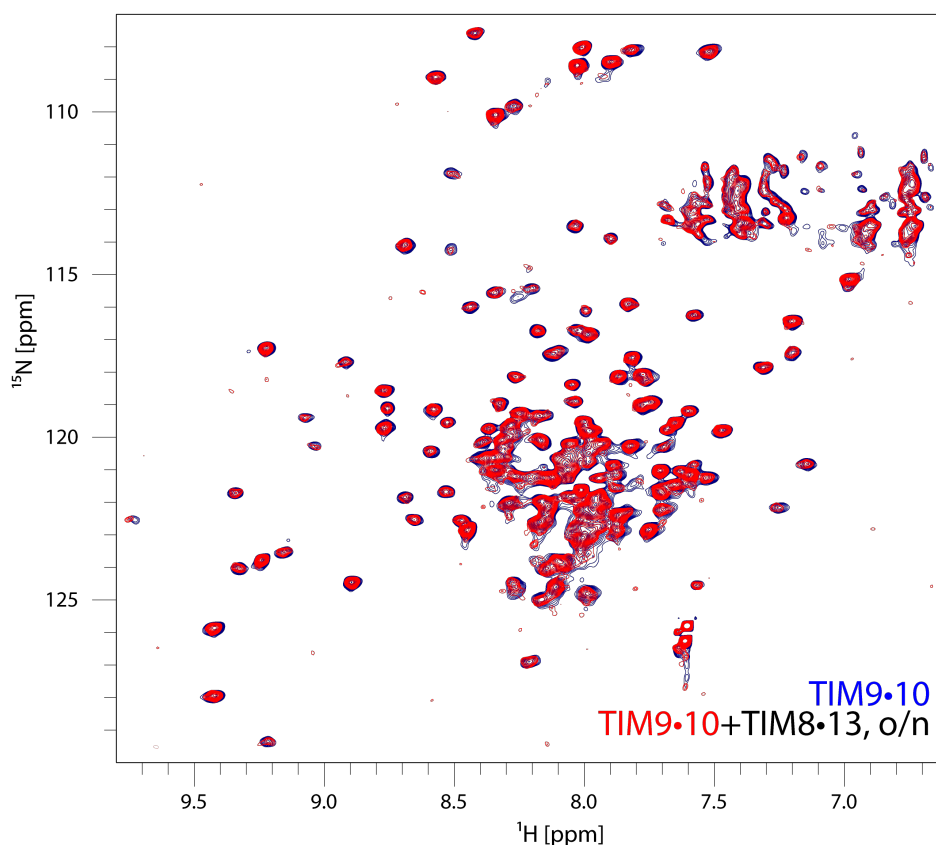
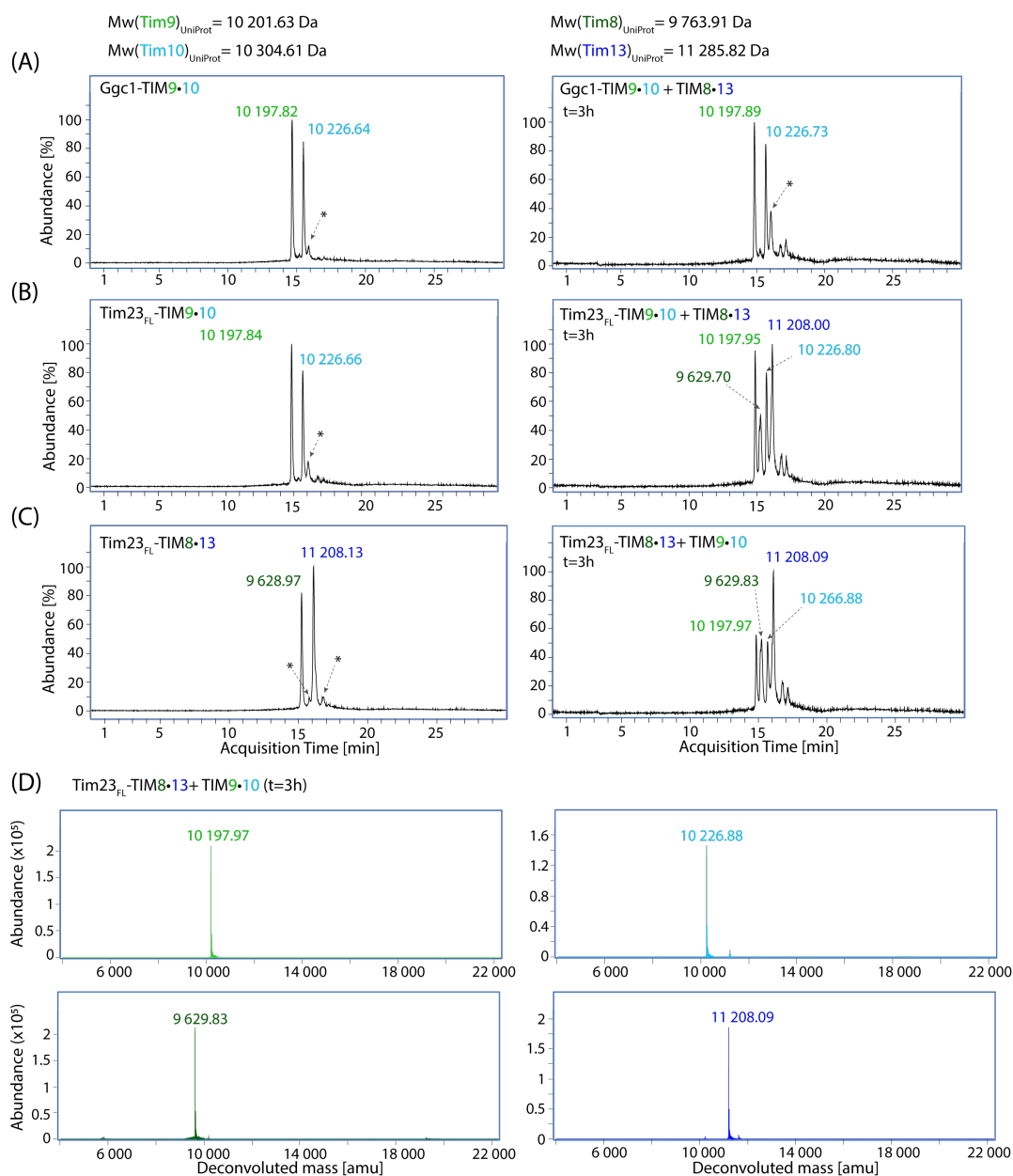


FIGURE S3: **TIM8-13 and TIM9-10 do not form mixed complexes.** NMR spectra of a <sup>2</sup>H,<sup>15</sup>N,<sup>13</sup>C labelled TIM9-10 sample (blue) and a mixture of this sample with unlabelled (NMR-invisible) TIM8-13 after overnight incubation. If TIM9-10 formed mixed oligomers with TIM8-13, the environment around each of the Tim9 and Tim10 subunits, as they would be surrounded by Tim8 or Tim13 subunits, would be different. Thus, the spectrum of TIM9-10 would feature additional peaks corresponding to those alternate environments. The spectrum after over-night incubation does not show any additional peaks and features, only the one set of peaks corresponding to the hexameric TIM9-10. Therefore, this data demonstrates that TIM9-10 does not form mixed hetero-oligomers with TIM8-13.



**FIGURE S4: Quantification of the amount of chaperone complexes from mass spectrometry.** Liquid chromatography coupled with electrospray ionization, time-of-flight mass spectrometry (LC/ESI-TOF MS) analysis used to read out the difference in the amount of specific chaperone, TIM8-13 and TIM9-10, bound to precursor protein. **(A)** Mass chromatograms of the control sample, complex of TIM9-10 chaperone bound to Ggc1 (left), and the competition reaction three hours after adding TIM8-13 to the pre-formed Ggc1-TIM9-10 complex (right). Deconvoluted mass values are indicated next to the corresponding peaks. Mass obtained for the unspecific peak, indicated with asterisk, was 10 226.53 (left) and 10 226.53 amu (right). **(B)** Same as in panel A, with the Tim23<sub>FL</sub> as a substrate precursor protein. Mass of the unspecific peak in the left chromatogram was 10 198.69 amu. **(C)** Mass chromatogram of the TIM8-13 chaperone bound to Tim23<sub>FL</sub> (left) as a control, and of the competition reaction three hours upon adding TIM9-10 to the pre-formed Tim23<sub>FL</sub>-TIM8-13 complex (right). In the left chromatogram, obtained mass from the unspecific peaks (impurities) was 10 207.94 (left asterisk) and 11 208.27 amu (right asterisk). **(D)** Deconvoluted mass spectra for each of the chromatography peaks from the chromatogram shown in panel C, right. Reaction mixtures were heat-shocked before the analysis, resulting in the presence of the chaperone only in the analysed sample (see methods). Additional experiments were performed with a mixture of TIM8-13 and TIM9-10 of known quantities, in order to correlate observed integrals of these chromatograms to molar ratios between the two chaperones (not shown).

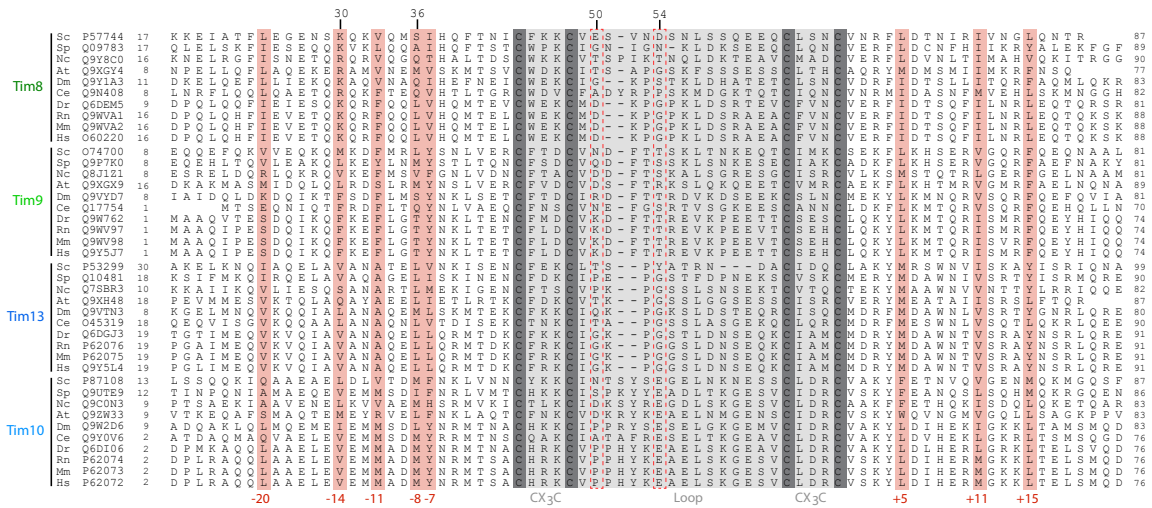
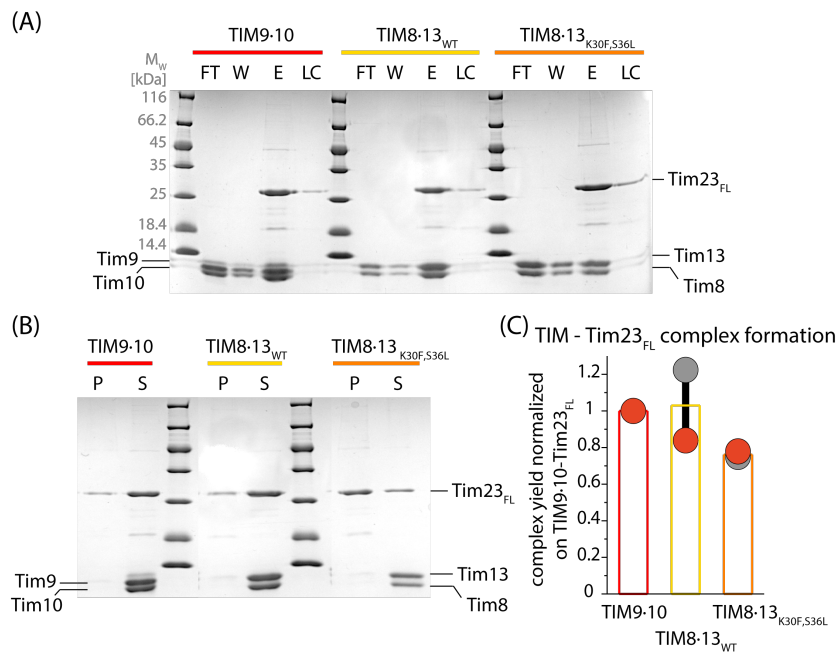


FIGURE S5: Sequence alignment of the small Tims across eukaryotes reveals a conserved hydrophobic motif. The shown sequences include *Saccharomyces cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp), *Neurospora crassa* (Nc), *Arabidopsis thaliana* (At), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), *Danio rerio* (Dr), *Rattus norvegicus* (Rn), *Mus musculus* (Mm) and *Homo sapiens* (Hs). The positions of the hydrophobic motif are highlighted in red, and was identified by analyzing the mean hydrophobicity across the alignment, as reporter earlier (Weinhäupl et al., 2018). The respective numbering of the hydrophobic position is provided below the sequence alignment. The strictly conserved cysteines in dark grey, and the top part of the chaperones in light grey. Positions discussed in the main text are numbered above the alignment, whereby the numbering refers to *Sc* Tim8.



**FIGURE S6: The Tim8 double mutant with increased hydrophobicity does not improve Tim23 binding.** Results of a pull-down binding experiment of Tim23<sub>FL</sub> with wild-type TIM9-10, wild-type TIM8-13 and TIM8-13<sub>K30F,S36L</sub>. **(A)** SDS-PAGE of the different fractions of the pull-down experiments. Lane descriptions are as in Figure 1H. **(B)** SDS-PAGE of the soluble (S) and insoluble (P) fraction of the Tim23<sub>FL</sub> complex with TIM9-10<sub>WT</sub>, TIM8-13<sub>WT</sub> or TIM8-13<sub>K30F,S36L</sub> showing that under the same conditions, Tim23<sub>FL</sub>-TIM8-13<sub>K30F,S36L</sub> complex is less stable. Immediately after the elution of the complex during the pull-down experiment, dialysis was performed. Dialysed sample was centrifuged in order to obtain the remaining complex in the soluble fraction and an insoluble Tim23<sub>FL</sub> in the pellet. **(C)** Quantification of the amount of complex obtained after imidazole-elution. The three pull-down experiments (TIM9-10, TIM8-13, TIM8-13<sub>K30F,S36L</sub>) were done in parallel, binding each time the same amount of Tim23<sub>FL</sub> to the column, and performing the same wash, elution and end-wash (denaturing conditions) steps. The experiment was repeated twice, with Tim23<sub>FL</sub> from two different production batches (shown in red and grey, respectively). As the protein was from different purifications, the amount of Tim23<sub>FL</sub> bound onto the column was not exactly identical; for this reason, the amount obtained in each experiment (of the order of 3 mg complex) was normalized to the amount of TIM9-10-Tim23<sub>FL</sub> complex.



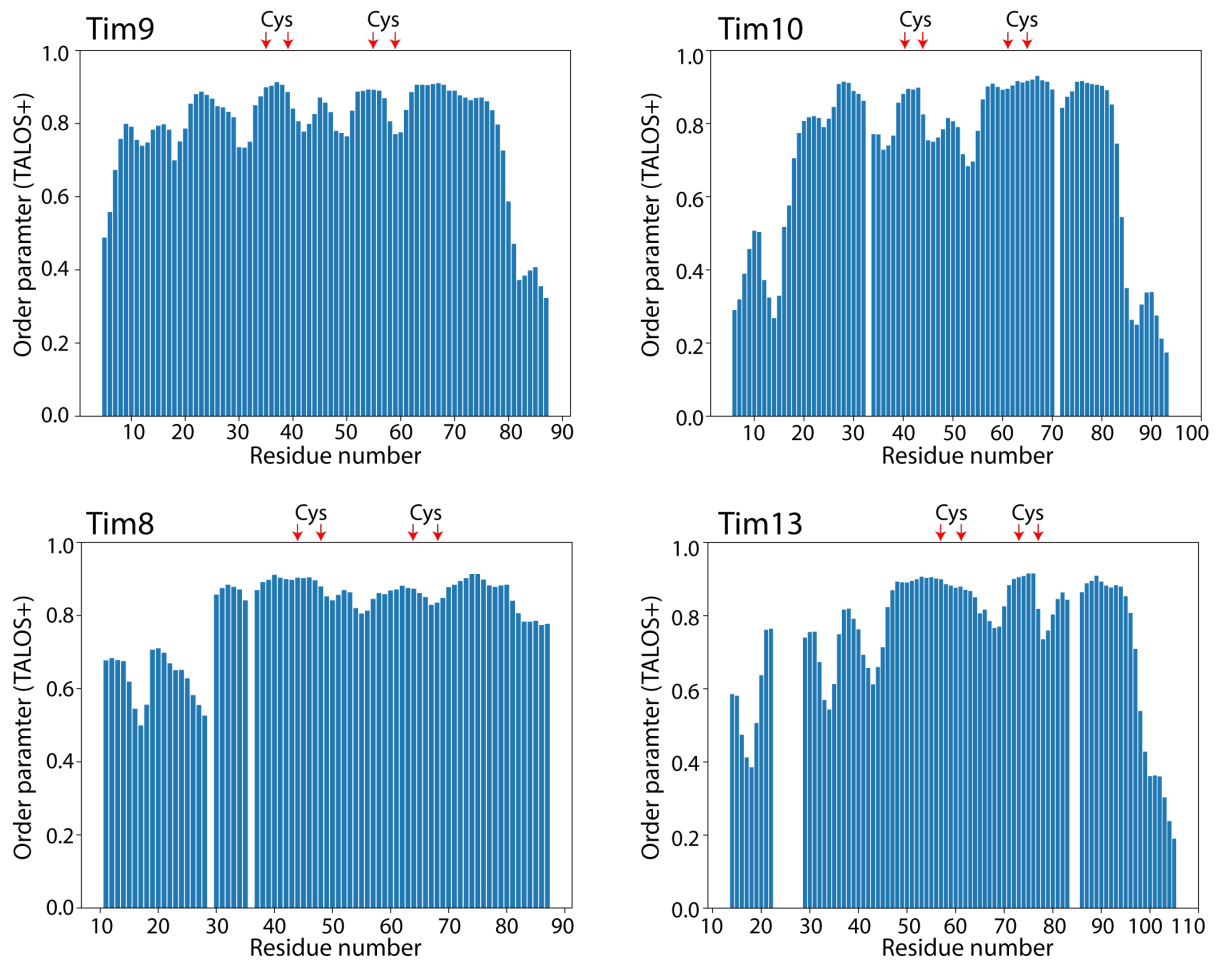


FIGURE S7: **TIM chaperones have a compact rigid core and flexible tentacles.** Local order parameters in TIM8-13 and TIM9-10, derived from assigned backbone chemical shifts and the TALOS-N (Shen and Bax, 2013) software. This data shows that the core of the chaperones is rather rigid, and the tentacles become increasingly flexible towards the N- and C-termini.

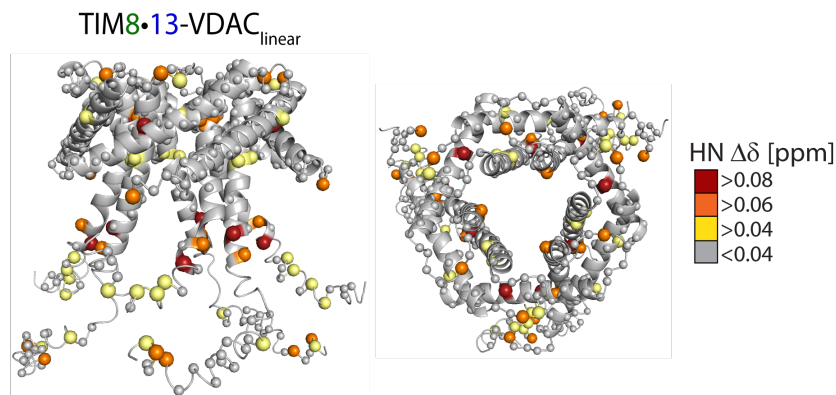
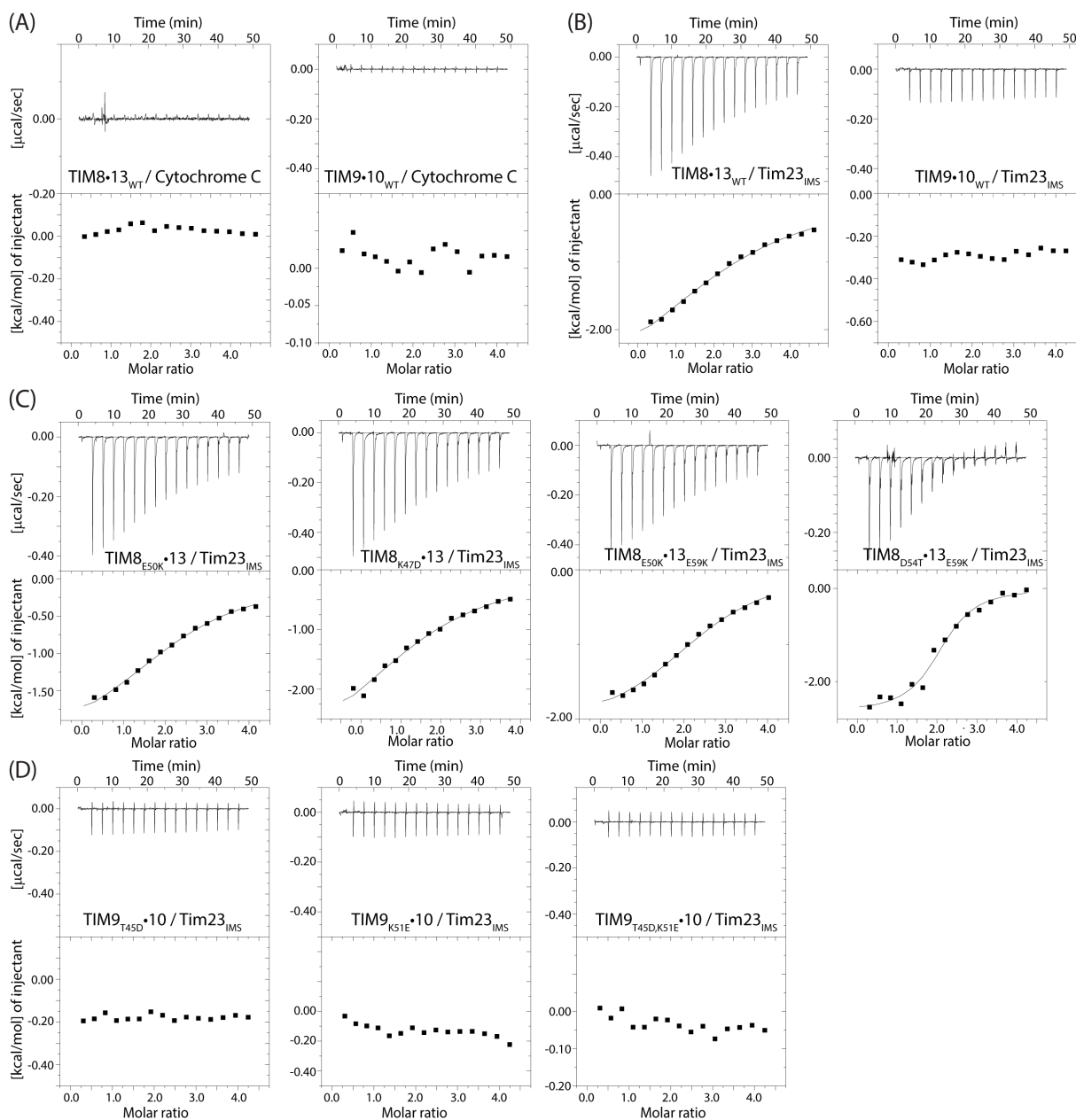


FIGURE S8: **The linear VDAC<sub>257-279</sub> peptide does not bind to a well-defined binding site of TIM8-13.** <sup>1</sup>H-<sup>15</sup>N chemical-shift perturbations on TIM8-13 upon addition of the linear VDAC<sub>257-279</sub> peptide, mapped onto the structure of TIM8-13. In contrast to the cyclic VDAC<sub>257-279</sub> peptide, we observe widely spread CSPs with the linear peptide, indicating unspecific binding across the chaperone.



**FIGURE S9: Isothermal titration calorimetry (ITC) experiments with TIM8-13 and TIM9-10 chaperones.** Thermograms (thermal power as a function of time) are displayed in the upper plots, and binding isotherms (ligand-normalized heat effects per injection as a function of the molar ratio,  $[\text{Tim}23_{\text{IMS}}]/[\text{chaperone}]$ ) are displayed in the lower plots. **(A)** Calorimetric titration of wild-type TIM8-13 and TIM9-10 with cytochrome c. No binding of chaperones with folded, soluble cytochrome c was detected, supporting that interaction of the chaperone with client proteins requires either a hydrophobic region, or an unfolded stretch of hydrophilic residues. **(B)** ITC of TIM9-10 and TIM8-13 with  $\text{Tim}23_{\text{IMS}}$ . **(C)** Calorimetric titration of TIM8-13 mutants with  $\text{Tim}23_{\text{IMS}}$ . Variants were obtained by charge inversion of negatively charged residues in the top region of the chaperone. The highly conserved Lys47 was exchanged for a negatively charged residue to investigate its relevance for binding of  $\text{Tim}23$ . All variants show similar isothermal binding curves, despite altered charges. This observation points to more complex binding interactions, which were not significantly perturbed in the chosen mutants. **(D)** Calorimetric titration of TIM9-10 variants with  $\text{Tim}23_{\text{IMS}}$ . Variants were obtained by inverting the charge of selected positively charged residues in the top region of the chaperone. No binding was detected between  $\text{Tim}23_{\text{IMS}}$  and any of the TIM9-10 variants.

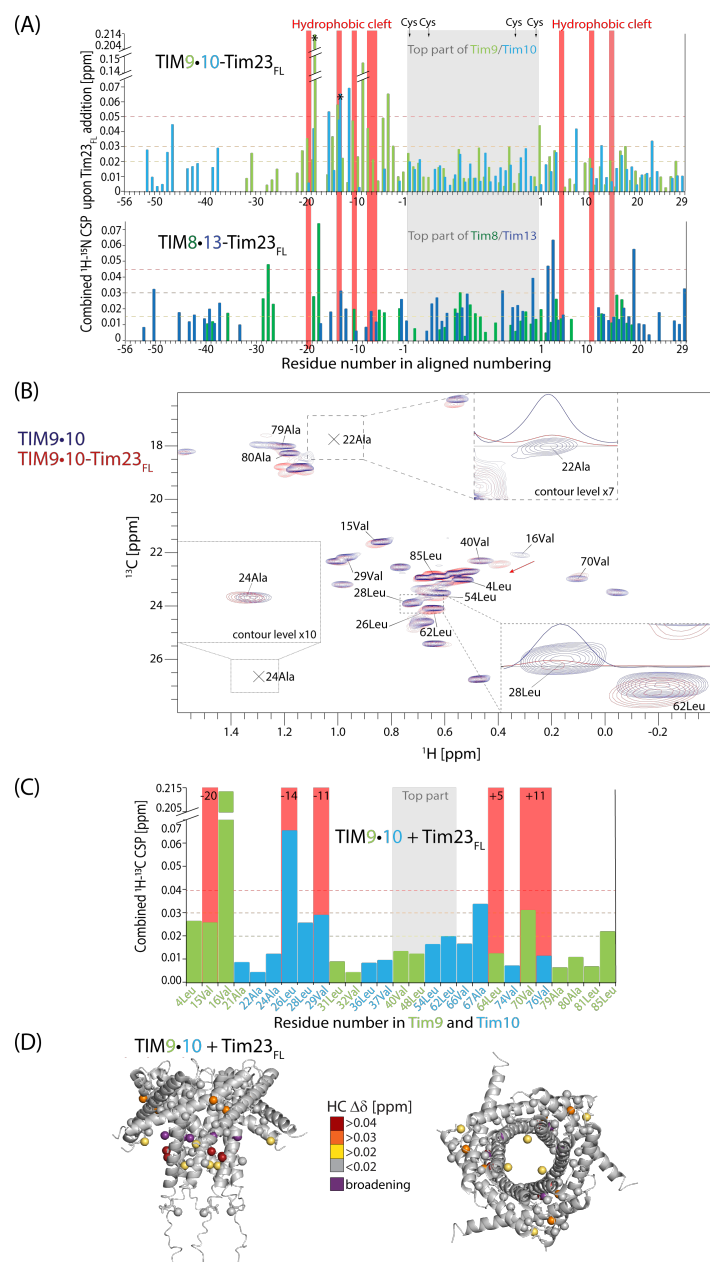


FIGURE S10: **Full-length Tim23 interactions with TIM9-10 and TIM8-13 chaperones.** (A) Chemical-shift perturbations observed upon addition of the Tim23<sub>FL</sub> to TIM9-10 (top) and TIM8-13 (bottom). Comparison of CSPs on TIM8-13 induced by Tim23<sub>FL</sub> and VDAC<sub>257-279</sub> (Fig. 5.2C), shows that for binding of full-length Tim23, besides conserved hydrophobic binding site, TIM8-13 employs its hydrophilic top region where 3 fold increase in CSPs can be seen (mapped on the structure in Fig. 5.4C). Asterisk marks residues 16Val and 26Leu which showed biggest methyl CSP upon adding Tim23<sub>FL</sub>. (B) Methyl spectra overlay of the ILVA-labeled apo-TIM9-10 chaperone sample (blue) and ILVA-TIM9-10 in complex with the full-length Tim23 (red). In the zoomed-in regions, showing residues 22Ala and 28Leu, difference in the peak intensity is shown with the 1D traces (hydrogen spectra) at the corresponding carbon frequency of the selected peak. The biggest methyl CSP for the residue 16Val is indicated with a red arrow. (C) Methyl-detected CSP of the ILVA-labeled TIM9-10 upon addition of the full-length Tim23. (D) Mapping of Tim23<sub>FL</sub>-induced methyl CSPs and biggest peak-broadening on TIM9-10, showing absence of interaction with the top part of TIM9-10.

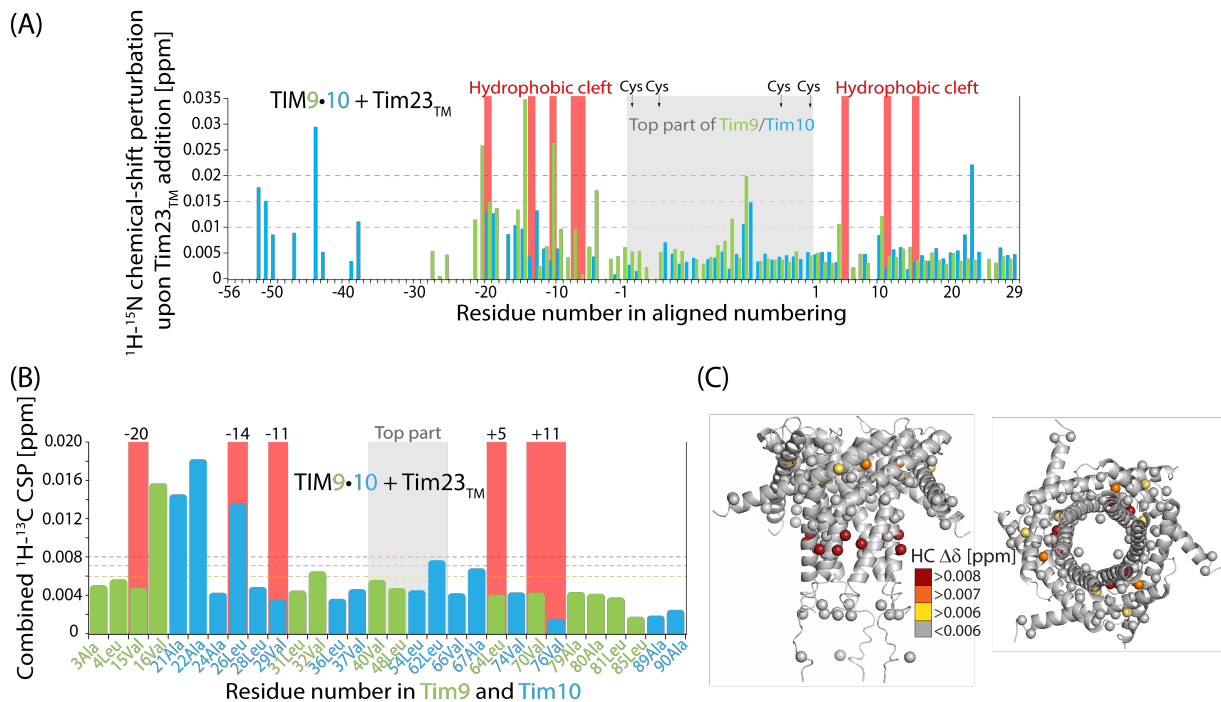


FIGURE S11: **Interaction of trans-membrane  $\text{Tim23}$  with the  $\text{TIM9-10}$  chaperone.** (A)  $^1\text{H}$ - $^{15}\text{N}$  chemical-shift perturbations observed upon adding the trans-membrane construct  $\text{Tim23}_{\text{TM}}$  to  $\text{TIM9-10}$ . (B) Methyl-detected chemical-shift perturbations of the ILVA-labeled  $\text{TIM9-10}$  upon addition of the trans-membrane  $\text{Tim23}$ . (C) Mapping of  $\text{Tim23}_{\text{TM}}$ -induced methyl CSPs on  $\text{TIM9-10}$ , showing binding in the hydrophobic cleft, similarly to the binding detected with  $\text{Tim23}_{\text{FL}}$ .

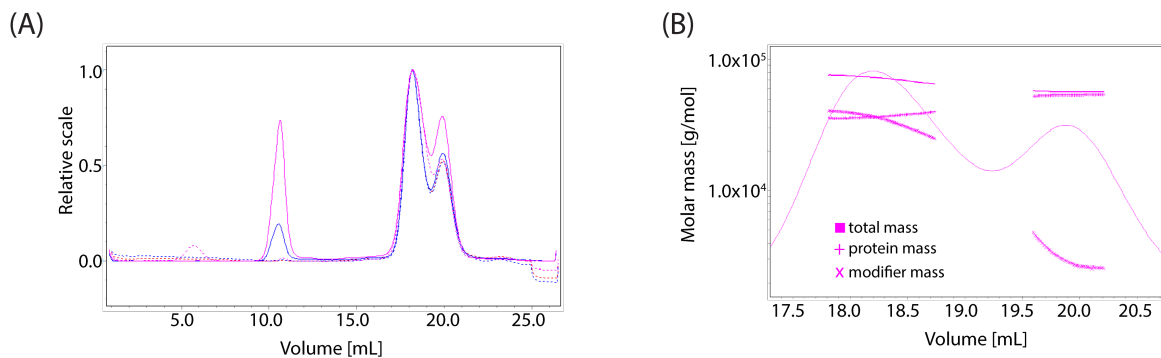
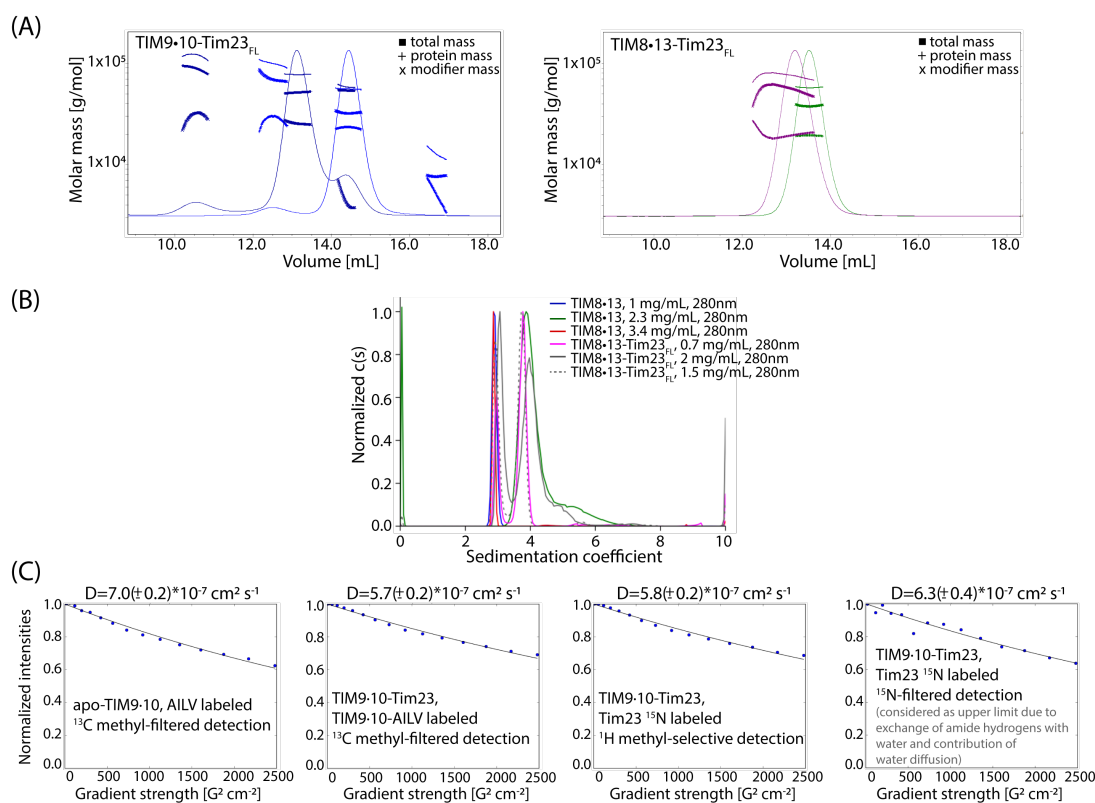


FIGURE S12: **In the simultaneous presence of  $\text{Tim23}$ ,  $\text{TIM8-13}$  and  $\text{TIM9-10}$ , no ternary complex is observed.** The presence of ternary complexes in a sample containing  $\text{TIM9-10}$ ,  $\text{Tim23}_{\text{FL}}$  and  $\text{TIM8-13}$  was investigated using multi-angle light scattering coupled to size-exclusion chromatography.  $\text{TIM9-10-Tim23}_{\text{FL}}$  complex was prepared using the pull-down method of Figure 5.1B, and  $\text{TIM8-13}$  was added. If  $\text{TIM8-13}$  binds to the soluble N-terminal half of  $\text{Tim23}$  with sufficient affinity, a ternary complex may be detected.

(A) Elution profiles of  $\text{TIM9-10-Tim23}_{\text{FL}}$  complex (blue) and  $\text{TIM9-10-Tim23}_{\text{FL}}$  after adding slight excess of  $\text{TIM8-13}$  (pink). The elution profiles of the two samples are similar, with no additional peak appearing at higher molecular weights, suggesting that there is no ternary complex present. Protein fraction eluting at 10 mL, in the column void volume, corresponds to the aggregates present in both samples. (B) Multi-angle light scattering analysis of two main peaks of the size-exclusion chromatography of the  $\text{TIM9-10-Tim23}_{\text{FL}}$  after adding slight excess of  $\text{TIM8-13}$ . A protein fraction eluting at 20 mL, corresponds for 36% of the injected mass and it corresponds to a co-elution of  $\text{TIM9-10}$  ( $M_{\text{w,theor}} = 60.8$  kDa) and  $\text{TIM8-13}$  ( $M_{\text{w,theor}} = 62.4$  kDa), while the fraction at 18.2 mL, corresponds for 49% of the injected mass and it corresponds to a co-elution of the complex of  $\text{TIM9-10-Tim23}_{\text{FL}}$  ( $M_{\text{w,theor}} = 85$  kDa) and the complex of  $\text{TIM8-13-Tim23}_{\text{FL}}$  ( $M_{\text{w,theor}} = 86.6$  kDa).



**FIGURE S13: Experimental characterization of the size of TIM chaperones and their precursor protein complexes.** (A) Size-exclusion-chromatography coupled to multi-angle light scattering (SEC-MALS) of TIM9-10-Tim23 and TIM8-13-Tim23. For TIM9-10 the dominant contribution (> 85 %) corresponds to 56 kDa, or ca. 5.5 subunits; for TIM9-10-Tim23, the major contribution (71 %) corresponds to a mass of 77.5 kDa, in reasonable agreement with the expected 23 kDa increase from Tim23-binding, with a clearly visible ‘shoulder’ at the position of the apo-TIM9-10. For TIM8-13, a peak corresponding to ca. 86 % of the signal is at a mass of 57.9 kDa; in TIM8-13-Tim23, the peak is at 74.2 kDa, in reasonable agreement with the expected 23 kDa increase. A control experiment with BSA yielded an observed mass of 62 kDa (theoretical: 66 kDa).

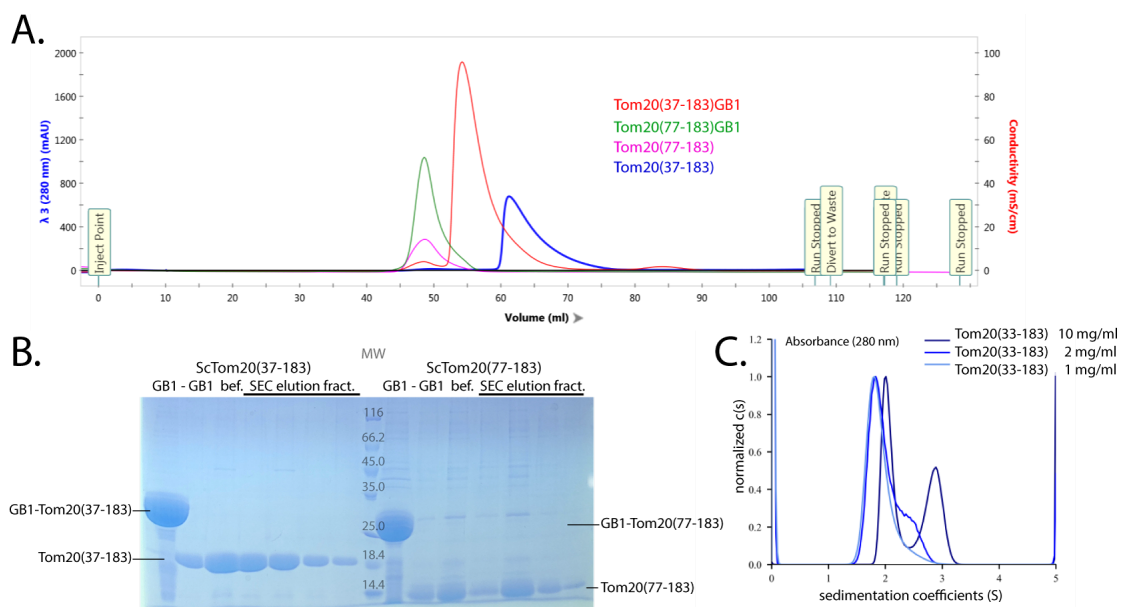
(B) Analytical ultra-centrifugation of TIM8-13 and TIM8-13-Tim23 complexes. For the TIM8-13 we observe a main contribution,  $98 \pm 1\%$  of the total signal, at  $2.89 \pm 0.01$  S ( $s_{20w} = 3.97 \pm 0.01$ S). The Non Interacting Species analysis gives  $M_w = 57 \pm 5$  kDa. For the TIM8-13-Tim23 sample we observe a main contribution ( $67 \pm 2\%$  of the total signal) at 3.76S ( $s_{20w} = 5.17$ S), at 0.7 mg/mL sample concentration, slightly shifted to  $s_{20w} = 5.67$ S, at 2 mg/mL. The Non Interacting Species analysis results in a  $M_w = 70 \pm 5$  kDa, similar to SEC-MALS. A further contribution is detected at  $3 \pm 0.05$ S ( $s_{20w} = 4.1 \pm 0.1$ S), for  $33 \pm 5\%$  of the total signal. The Non Interacting Species analysis gives  $M_w = 50 \pm 3$  kDa, could be imprecise. This contribution superimposes to the main one of TIM8-13 alone.

(C) NMR diffusion-ordered spectroscopy curves (DOSY), obtained as integrals of one-dimensional <sup>1</sup>H spectra over either methyls (first three) or amides (fourth panel) as a function of the gradient strength. The four panels correspond to three different samples, <sup>13</sup>CH<sub>3</sub>-ILVA-labeled apo-TIM9-10 (first panel), TIM9-10-ILVA methyl labeled TIM9-10-Tim23<sub>FL</sub> complex (second panel) and Tim23<sub>FL</sub>-<sup>15</sup>N labeled TIM9-10-Tim23<sub>FL</sub> complex (last two panels). The first two used a <sup>13</sup>C-filtered experiment, the third one uses a <sup>1</sup>H methyl-frequency selective scheme and the fourth uses a <sup>15</sup>N-filtered version. The latter thus detects amide protons; as these are water exchangeable, in particular in disordered proteins, the observed diffusion coefficient is likely over-estimated, as it contains contributions from the diffusion coefficient of water; thus, this value is to be seen as an upper limit. The diffusion coefficient for a spherical particle scales with the cubic root of the molecular weight, which would lead to an expected change in the diffusion coefficient of  $(71.1/94.1)^{1/3}$ , i.e. ca. 10 % lower in the complex compared to the apo state; additional domain motions, such as the flexibility of the Tim23 tail, decreases the diffusion coefficient, i.e. would lead to a larger relative change. This expected 10-15 % effect on diffusion coefficients is in good agreement with the experimental effect (11-20 %). A TIM9-10:Tim23 stoichiometry, as found for TIM9-10-Ggc1 would lead to an expected DOSY effect of 32 % (considering only mass), thus in worse agreement with the data.



## Appendix B

# Appendix B



**FIGURE S1: Oligomerization state of the yeast Tom20 constructs.** **A.** Size-exclusion chromatography (SEC) profiles of the two yeast Tom20 constructs, run before and after the cleavage of the GB1-solubility tag. The construct used in our studies, for chemical shift assignment and interaction assays, is ScTom20(37-183). Upon removal of the solubility tag, the protein elutes at higher volumes as expected for smaller construct however, the non-symmetrical shape of the elution peak (blue) indicates non-homogeneity of the size of the particles in the sample. The non-symmetrical peak is between the expected size of a monomer and expected size of a dimer indicating the exchange between these two states. The shorter construct ScTom20(77-183) is eluting at the same volume, before and after removal of the GB1-tag (comparing the pink and green chromatograms). This construct was not used in further experiments due to its oligomerization. **B.** The gel after sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) of Tom20 protein. The SEC elution fractions of both Tom20(37-183) and Tom20(77-183) are migrating at the size expected for a cleaved protein sample, indicating that the elution from the SEC column at lower volumes is not due to the unsuccessful removal of the solubility tag. **C.** Analytical ultracentrifugation (AUC) experiments of the yeast Tom20(33-183) protein samples. The exchange rate between the monomer and the dimer state was characterized as fast, however the attempts to determined exact rate with only three different protein concentrations was not successful. The Tom20(33-183) construct was not used further, it was replaced by the new GB1-Tom20(37-183) construct with improved production yield.

ScTom20(37-183) apo

ScTom20(37-183) : pSu9(1-25) = 1 : 5

ScTom20(37-183) : ATP1(iMTS-L) = 1 : 5

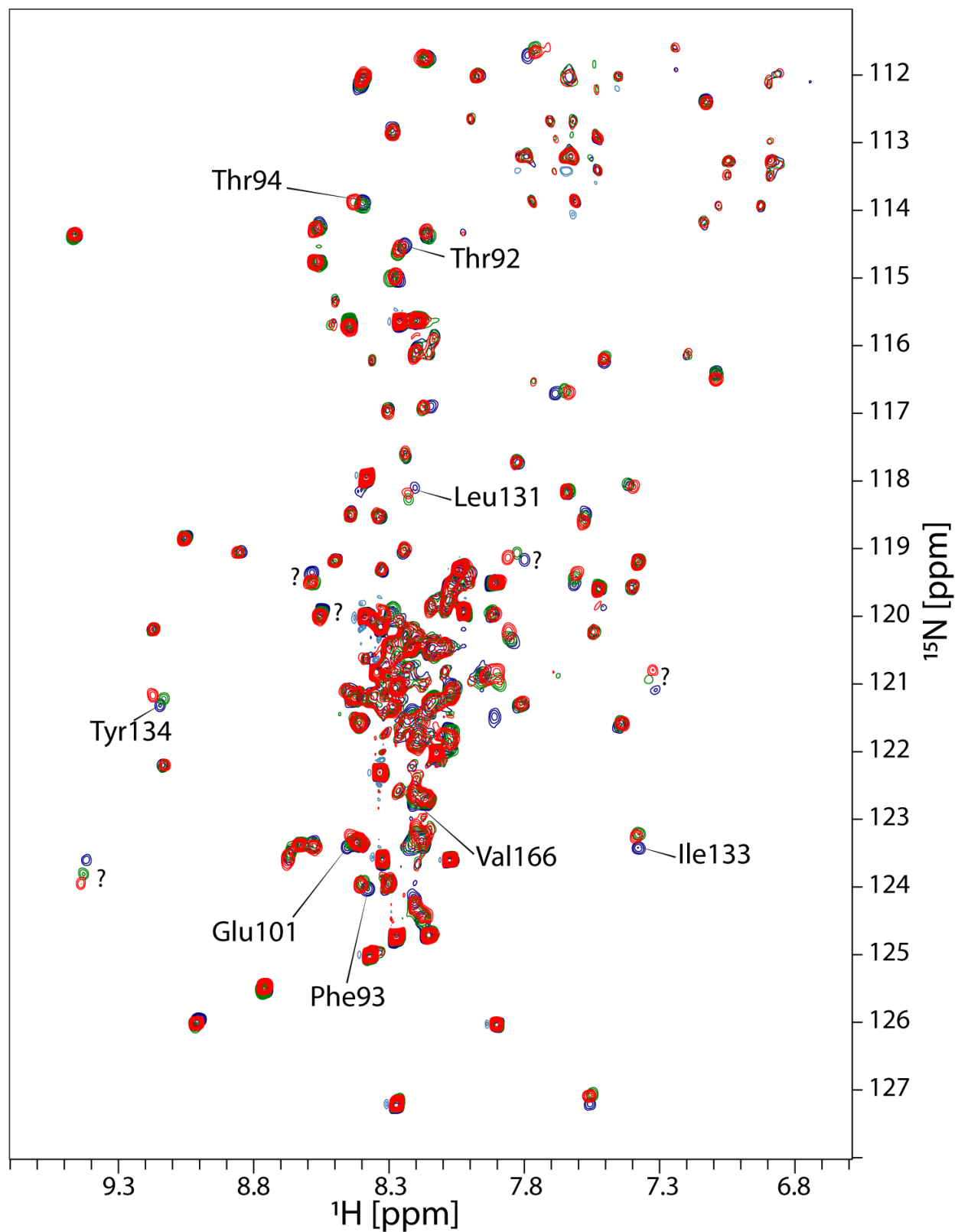


FIGURE S2: HN BEST-TROSY spectra of  $^2\text{H}^{15}\text{N}$ -labeled ScTom20(37-183). Spectrum of the  $^2\text{H}^{15}\text{N}$ -labeled ScTom20(37-183) alone is shown in blue. Spectrum upon adding pSu9(1-25) peptide to the  $^2\text{H}^{15}\text{N}$ -labeled ScTom20 is shown in green and upon adding ATP1(iMTS-L) peptide is shown in red. Both peptides were added in five-fold molar excess. Assigned residues of the presequence binding sites, for which chemical shift perturbations were observed are indicated.



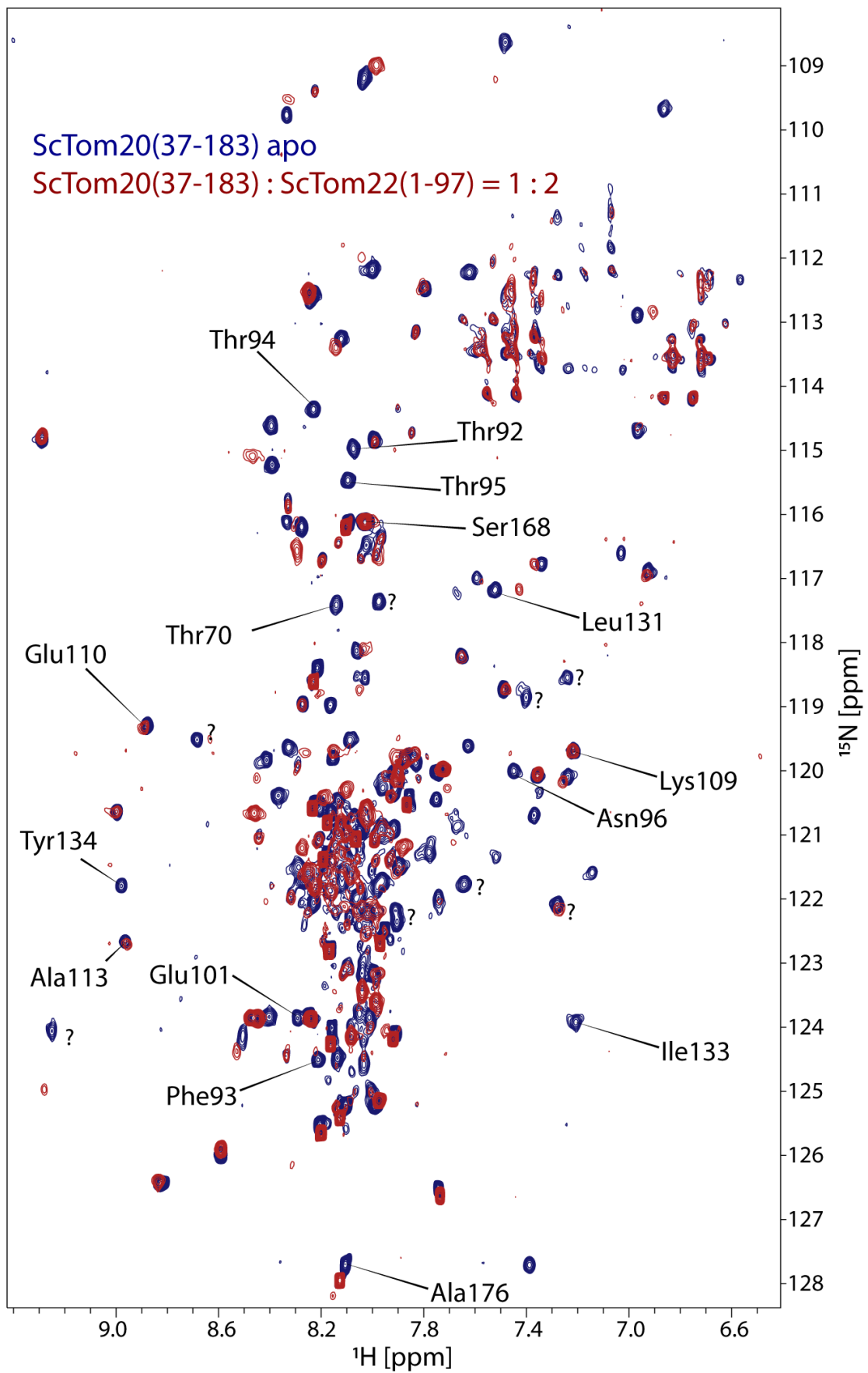


FIGURE S3: HN BEST-TROSY spectra of  $^2\text{H}^{15}\text{N}$ -labeled ScTom20(37-183) alone and upon adding Tom22(1-97). Spectrum of the  $^2\text{H}^{15}\text{N}$ -labeled ScTom20(37-183) alone is shown in blue and upon adding two-fold excess of Tom22(1-97) is shown in red.

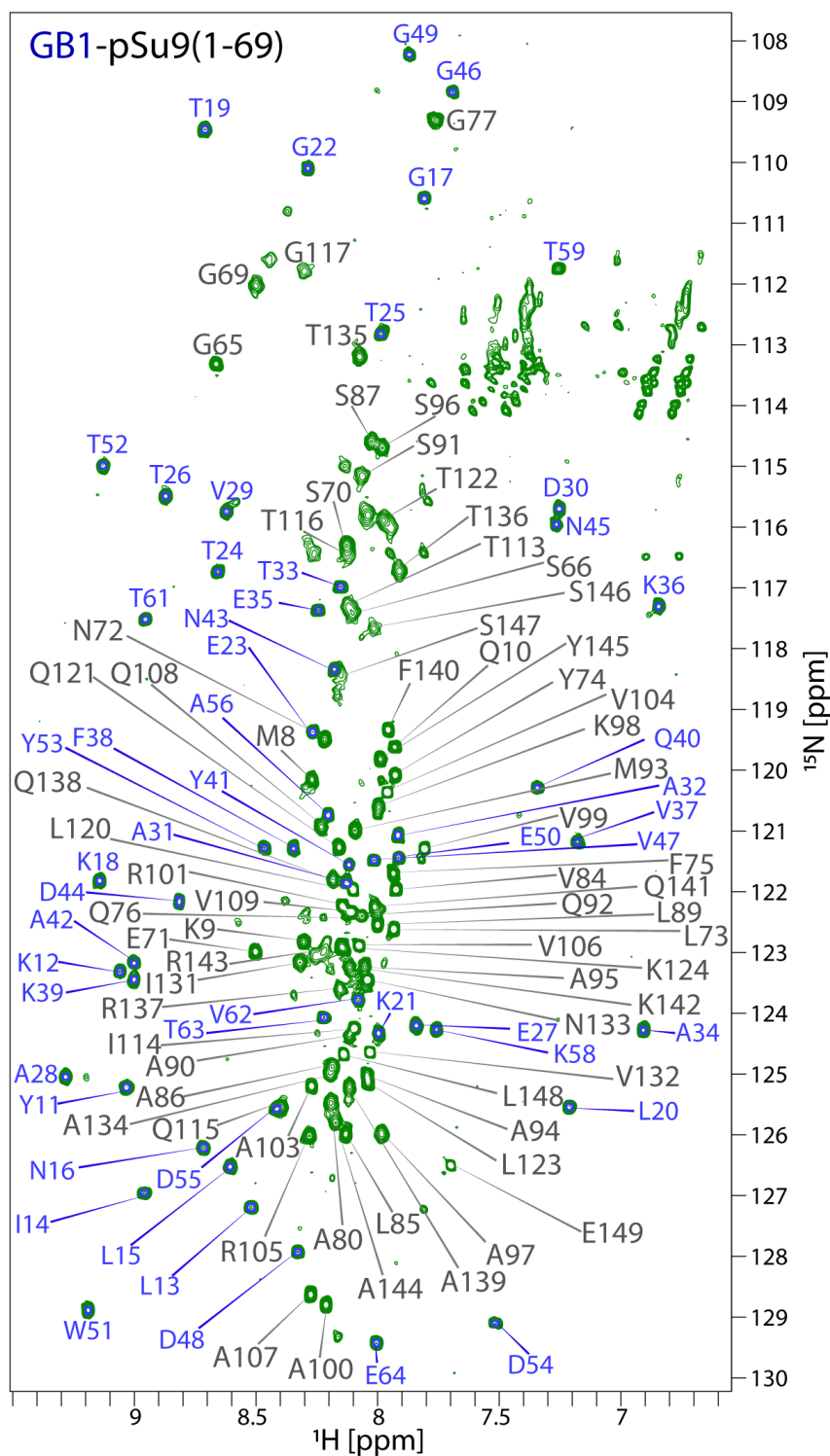


FIGURE S4: Assigned  $^{15}\text{N}^1\text{H}$  BEST-TROSY spectrum of  $^{15}\text{N}$ -labeled GB1-pSu9(1-69). Assigned spectrum of the  $^{15}\text{N}$ -labeled GB1-pSu9(1-69). Approximately 66% of the HN resonances of the pSu9(1-69) are assigned. The assignment is submitted to the BMRB database under entry number: 51093. Peaks corresponding to the pSu9(1-69) resonances have a low dispersion in the  $^1\text{H}$  dimension, characteristic for an IDP. The peaks of GB1 are well dispersed (shown in blue), as expected for a well folded protein domain.

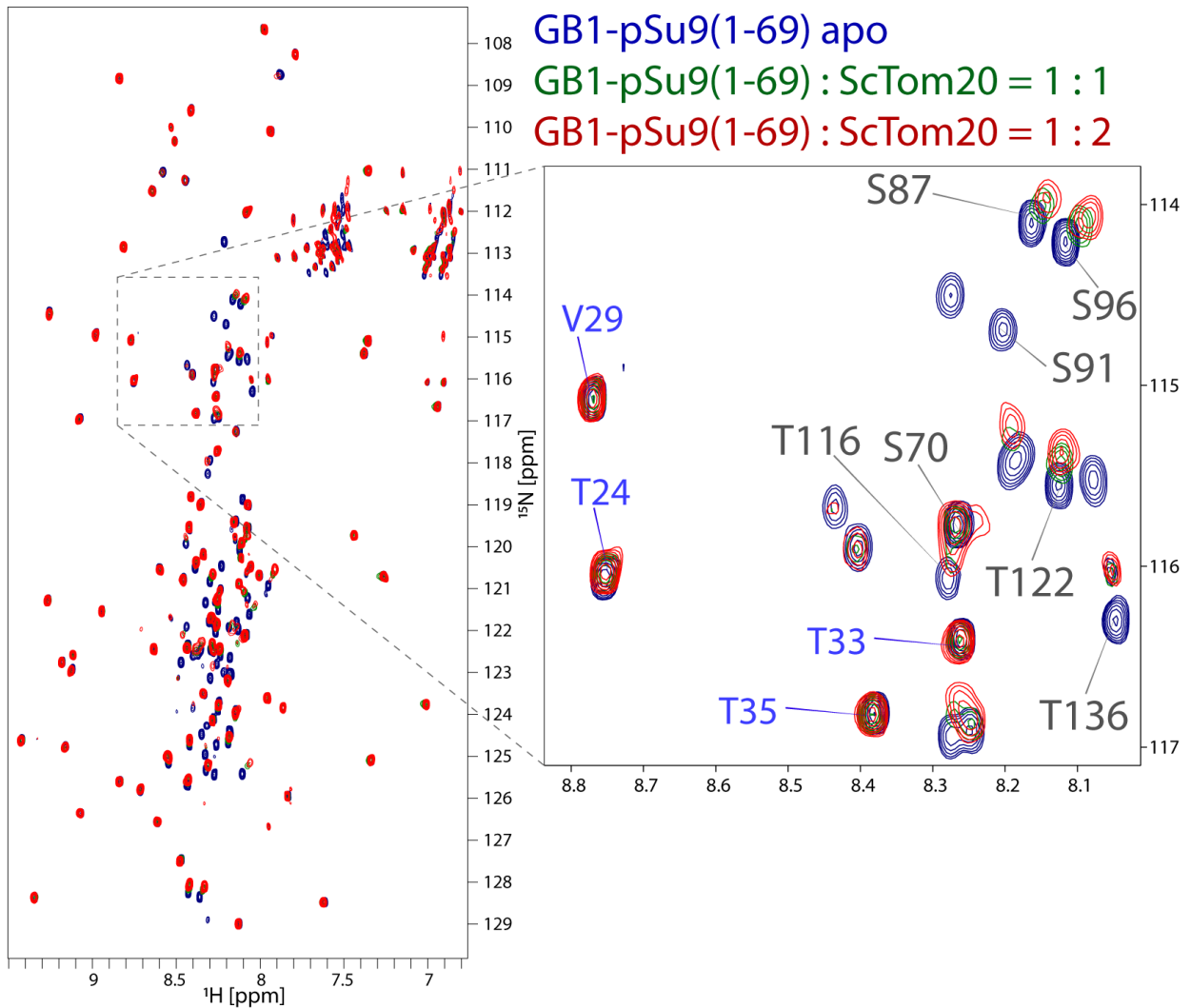
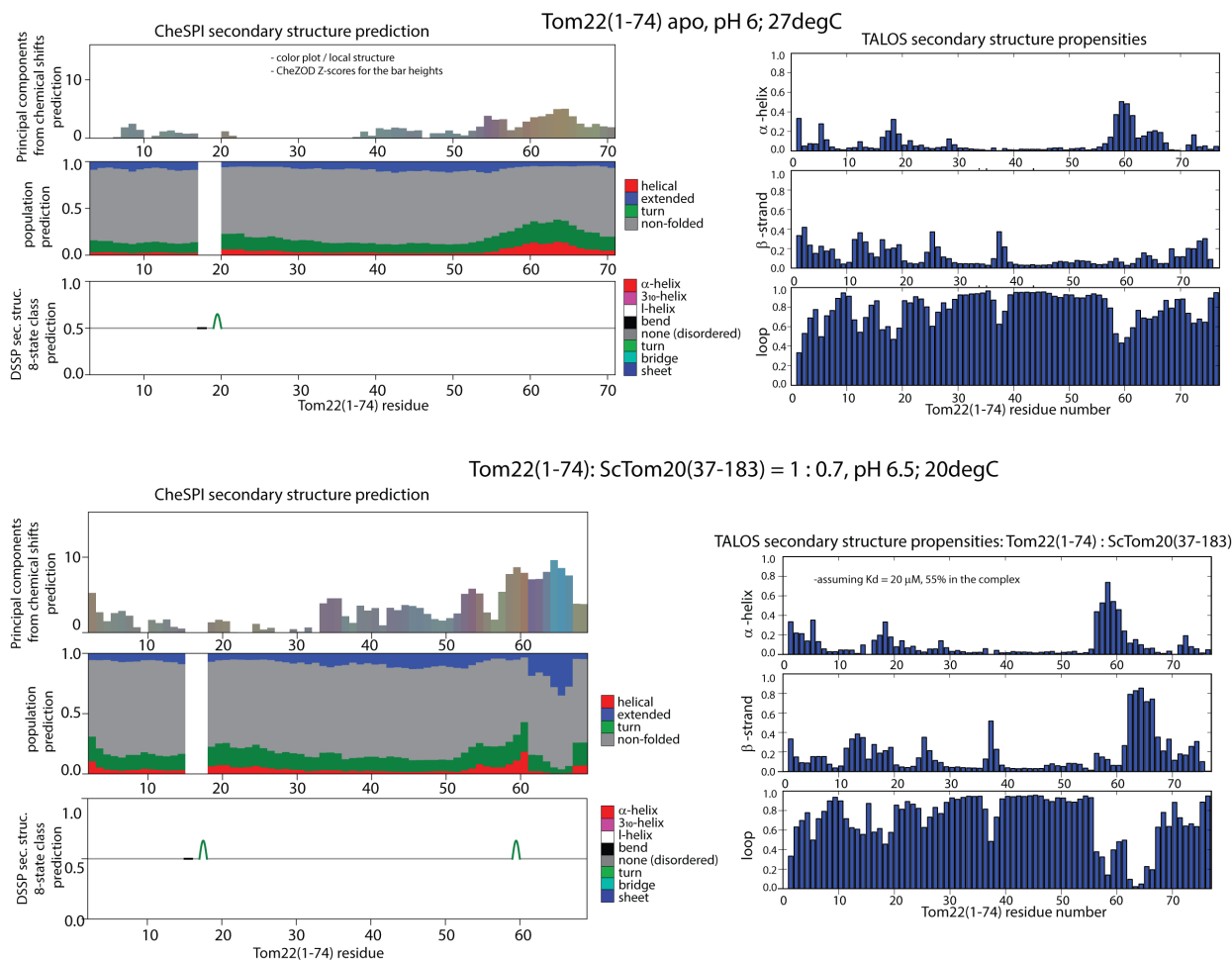


FIGURE S5:  $^{15}\text{N}^1\text{H}$  BEST-TROSY spectra of  $^{15}\text{N}$ -labeled GB1-pSu9(1-69) in the absence and presence of ScTom20. Spectrum of  $^{15}\text{N}$ -labeled GB1-pSu9(1-69) alone is shown in blue. Spectrum shown in orange is  $^{15}\text{N}$ -labeled GB1-pSu9(1-69) upon adding molar equivalent of the purified, unlabeled ScTom20(37-183) and in red is the spectrum upon adding two-fold molar excess of ScTom20. Enlarger area shows both GB1 residues (in blue: V29, T24, T33, T35) that do not change upon adding ScTom20, and pSu9 residues (S87, S96, T116, T122, T136) for which linear chemical shift perturbation and decrease in peak intensity could be observed with increasing concentration of ScTom20.



**FIGURE S6: Secondary structure propensities of Tom22 in presence and absence of Tom20.** Residue-wise secondary structure propensities for apo Tom22(1-74) (upper panels) and for yeast Tom20(37-183)-bound Tom22(1-74) (lower panel). Results obtained by CheSPI software are shown on the left hand side and by TALOS-N software on the right hand side. Based on provided H, N, C $\alpha$ , C $\beta$ , C $\gamma$  chemical shifts TALOS-N software predicted (20% more) increase in foldness of resn. 56-62 of Tom20-bound Tom22. For the Tom22 residues 62-69 the propensity for  $\beta$ -strand significantly increases. However, since the C $\alpha$ , C $\beta$  and C $\gamma$  resonances of residues 56, 57, 58 were not provided due to peak broadening when interacting with Tom20 TALOS-N provides propensities based only on the sequence and the matches in the database. Similar solution TALOS-N applied for the residues 61, 62, 64, 65 and 67 for which the consensus in database matches were not found. The results of CheSPI for the same region of Tom22 show small percentage of  $\alpha$ -helical secondary structure, comparable to the apo state of Tom22. Not all C $\beta$  and C $\alpha$  chemical shifts were provided due to severe peak broadening in this region of Tom22 upon interaction with Tom20 (for example peaks of Thr57, Leu58, Leu59, Val63).

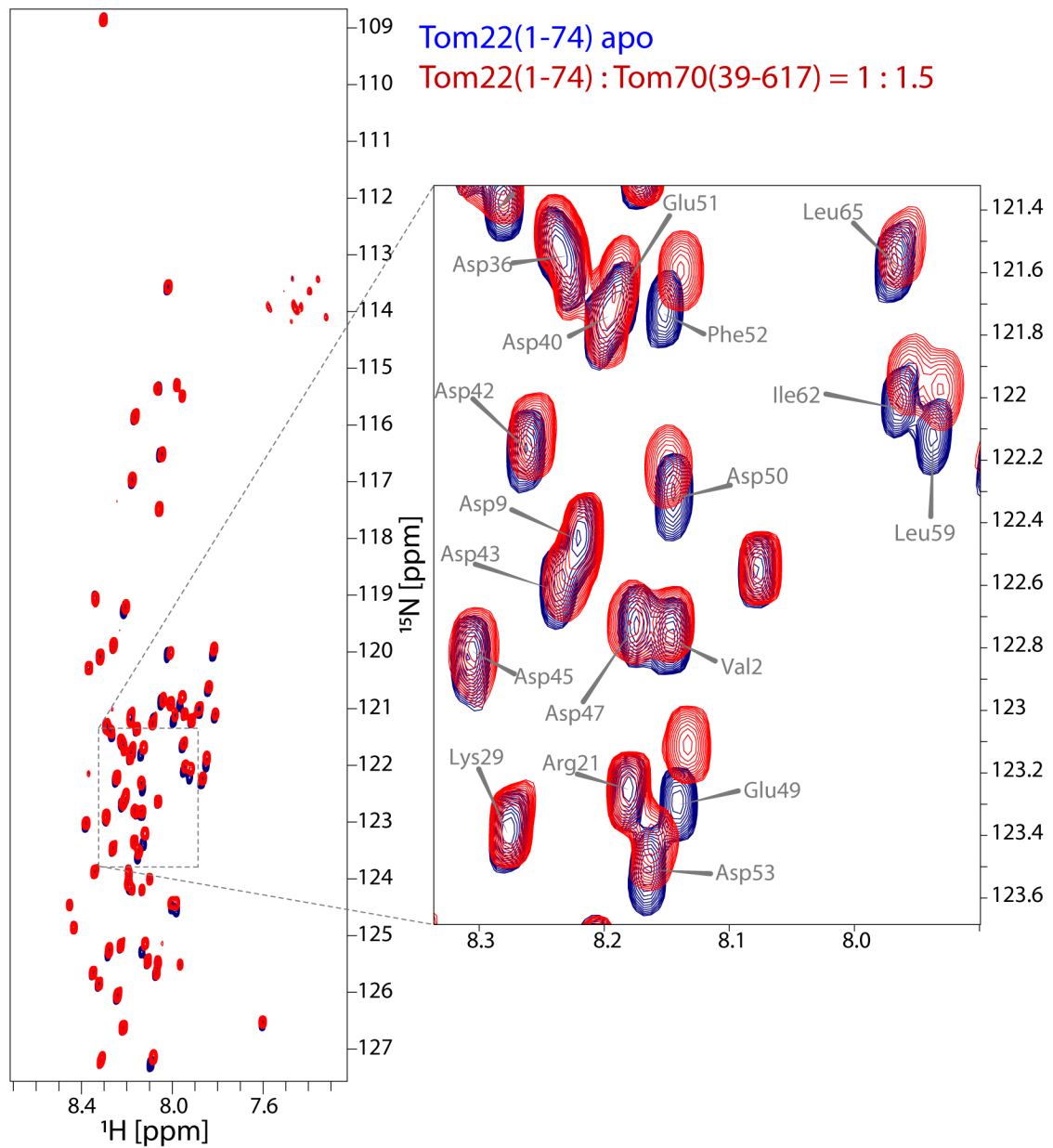
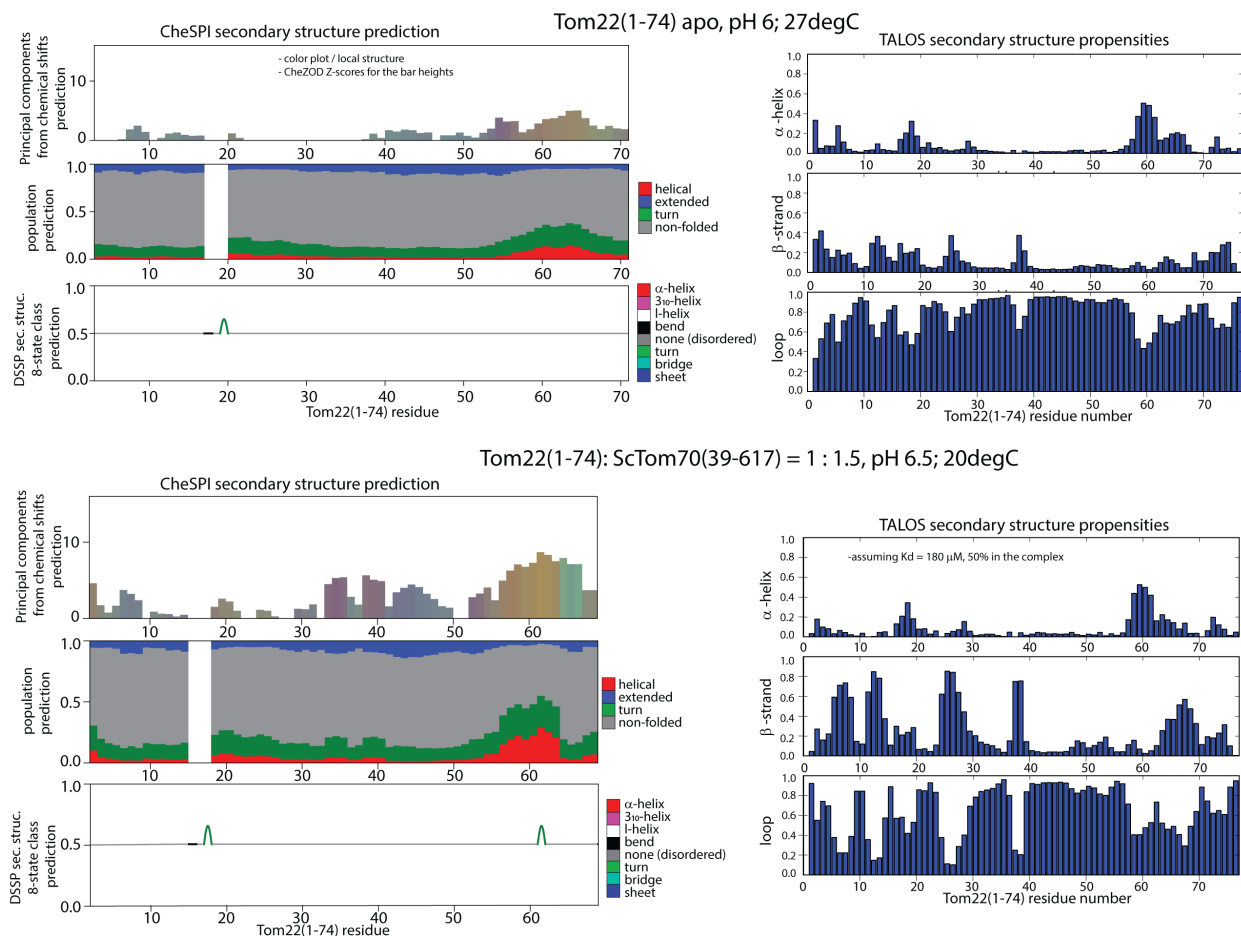


FIGURE S7: HN BEST-TROSY spectra of  $^{15}\text{N}$ -labeled Tom22(1-74). Spectrum of the  $^{15}\text{N}$ -labeled Tom22(1-74) alone is shown in blue. Spectrum upon adding 1.5 molar excess of Tom70(39-617) is shown in red.

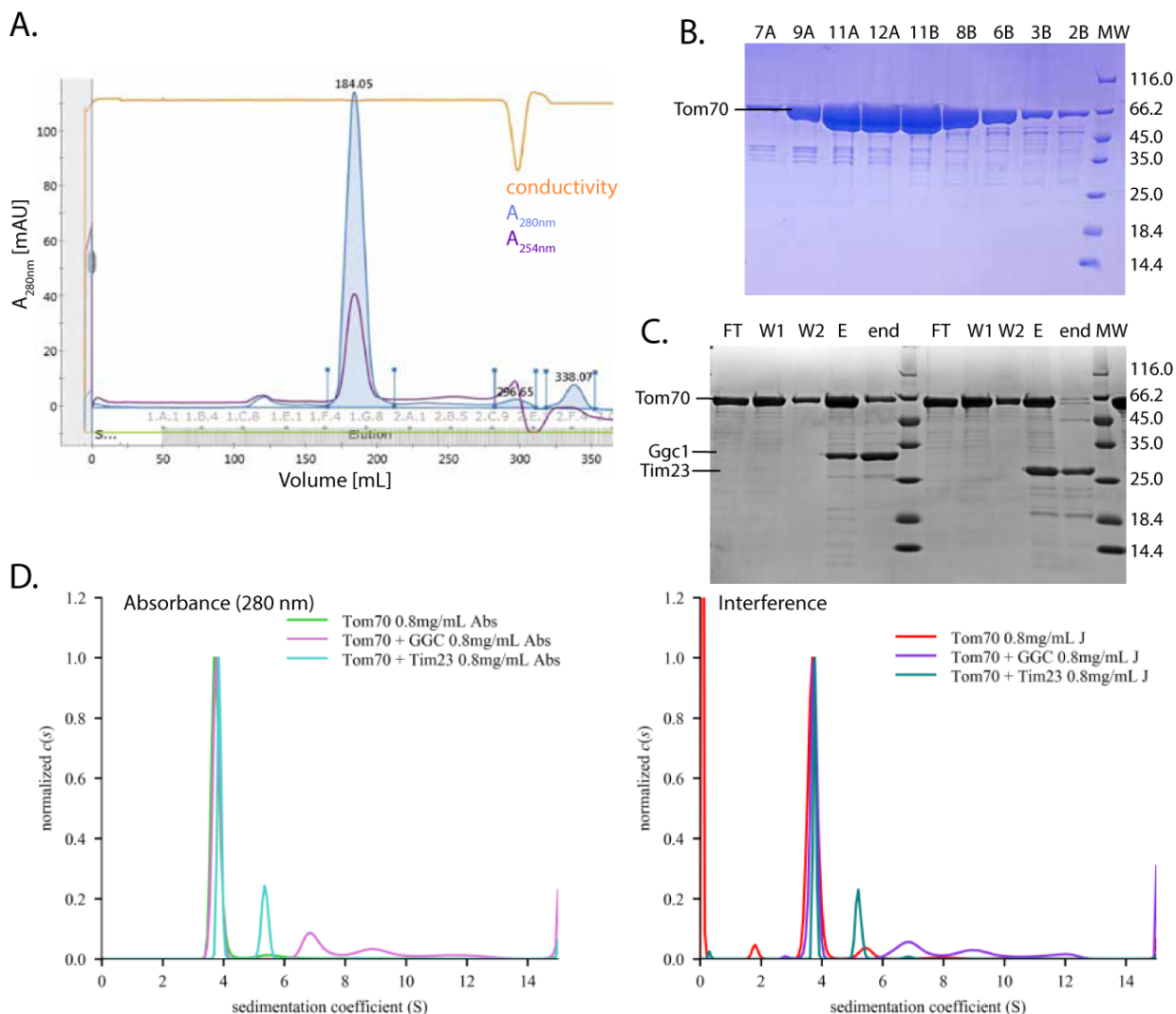


**FIGURE S8: Secondary structure propensities of Tom22 in presence and absence of Tom70.** Residue-wise secondary structure propensities for apo Tom22(1-74) (upper panels) and for Tom70(39-617)-bound Tom22(1-74) (lower panel). Results obtained by CheSPI software are shown on the left hand side and by TALOS-N software on the right hand side. Based on provided H, N,  $C\alpha$ ,  $C\beta$ ,  $C_O$  chemical shifts CheSPI software predicted slight increase in the helical and turn conformation population resn. 56-65 of Tom70-bound Tom22. The rest of the Tom70-bound Tom22 has comparable secondary structure propensities to the apo Tom22. TALOS-N software for Tom70-bound Tom22 predicts increase in the  $\beta$ -strand secondary structure propensities for the residues 7-15, with strong score for resn. 7, 8, 11, 12 and 13 while for the residues 14-19 which are not assigned the prediction is based on the sequence matches in the database. Similarly, the sequence based classification and predicted increase in the  $\beta$ -strand secondary structure propensities for the residues 25-30 could not be considered significant due to classification of the residues as dynamic based on TALOS-N analysis of the provided chemical shifts. Tom22-Tom70 sample was prepared with 150M isotopically labeled Tom22. Assuming dissociation constant of 180 M, 50% of Tom22 is in the bound state.

Residue-wise secondary structure propensities were obtained by CheSPI and TALOS-N software (Figure S8). Based on provided H, N,  $C\alpha$ ,  $C\beta$ ,  $C_O$  chemical shifts of the Tom70-bound Tom22 (same used for calculating the secondary chemical shifts), CheSPI software predicted slight increase in the helical and turn conformation population resn. 56-65 of Tom70-bound Tom22. The rest of the Tom70-bound Tom22 has comparable secondary structure propensities to the apo Tom22. TALOS-N software for Tom70-bound Tom22 predicts increase in the  $\beta$ -strand secondary structure propensities for the residues 7-15, with strong score for resn. 7, 8, 11, 12 and 13 while for the residues 14-19 which are not assigned the prediction is based on the sequence matches in the database. Similarly, the sequence based classification and predicted increase in the  $\beta$ -strand secondary structure propensities for the residues 25-30 could not be considered significant due to classification of the residues as dynamic based on TALOS-N analysis of the provided chemical shifts.

### B.0.1 Characterization of the yeast Tom70 receptor

During the size-exclusion (SEC) purification step of Tom70(39-617) minor fraction at higher molecular weight could be eluted, along with a major fraction corresponding to theoretical molecular weight of Tom70(39-617) construct of approximately 66 kDa (FigureS9A.). Parameters obtained from ProtParam Ex-pasy tool for used protein constructs could be seen in Table7.1, along with key production and purification parameters. To characterize protein sample of Tom70, such as estimate the size, approximate the shape (globular vs. elongated) and oligomeric state, we analysed the Tom70(39-617) protein sample after SEC purification by analytical ultra centrifugation (AUC) (FigureS9D.). A main peak corresponding to  $92 \pm 2\%$  of the total signal was detected at  $3.72 \pm 0.03$  S and the Non Interacting Species (NIS) analysis gives MW =  $63.8 \pm 0.1$  kDa which corresponds to an elongated monomer of Tom70(39-617). The NIS analysis of a minor peak at  $5.5 \pm 0.05$  S, which corresponds to 5% of the total signal, gave MW =  $100 \pm 10$  kDa. The ratio of sedimentation coefficients of major and this minor fraction  $5.5/3.72 = 1.48$  is close to the expected ratio  $s_{\text{dimer}}/s_{\text{monomer}}$  (1.6), for a monomer and a dimer with the same shape. This peak could correspond to an elongated dimer. A second minor peak at  $8.9 \pm 0.05$  S, for only 2% of the total signal, could correspond to an elongated tetramer based on the sedimentation coefficient ratios  $8.9/3.72 = 2.39$  and  $8.9/5.5 = 1.62$  which are close to the expected ratios  $s_{\text{tetramer}}/s_{\text{monomer}}$  (2.5) and  $s_{\text{dimer}}/s_{\text{monomer}}$  (1.6). However, the significance of this contribution, 2% of tetrameric state, could not be determined by this method. For our further NMR and biochemical assays 5% of dimer protein state in the sample does not impose any significant problems and based on the analysis of our samples by SDS-PAGE the purity of the sample is satisfactory (FigureS9B.).



**FIGURE S9: Purification and characterization of Tom70(39-617) protein construct.** **A.** Size exclusion chromatogram of cleaved (non His-tagged) deuterated  $^{15}\text{N}^{13}\text{C}$  labeled Tom70(39-617) on Superdex S200 26/600 column. Most of the protein sample elutes as monomer in a single peak. **B.** Gel after sodium dodecyl sulfate—polyacrylamide gel electrophoresis (SDS-PAGE) of main elution fraction/peak after SEC. Purity of Tom70 elution fraction is satisfactory for further NMR and biochemical studies. **C.** Pull-down experiment of Ggc1 and Tim23 with Tom70. The lanes correspond to flow-through after applying Tom70 (FT), additional washes (W1 and W2), imidazole elution (E) and final wash of anything that aggregated on the resin with denaturant (end). **D.** Analytical ultracentrifugation (AUC) analysis of apo Tom70(39-617), Tom70(39-617) in a complex with Ggc1<sub>FL</sub> or Tim23<sub>FL</sub>.

Complex formation of Tom70(39-617) with Tim23 and Ggc1 precursor proteins was achieved with a pull-down experiment described in Methods7.3. Briefly, precursor proteins were loaded in denaturing conditions on the Ni-NTA affinity resin and after replacing denaturant with binding buffer, Tom70(39-617) was added to the resin. After washing the excess of Tom70, the complex was eluted in imidazole and any unbound and therefore aggregated precursor was removed from the resin with a final (end) wash step by applying solution of denaturant and imidazole. Complex Tom70-Tim23 and Tom70-Ggc1 were successfully formed (FigureS9C.). The AUC analysis of the Tom70-Tim23 complex showed that the complex is not very stable at higher sedimentation velocity, where  $\sim 25\%$  of the initial signal was lost probably due to fast sedimentation of aggregates at 42000rpm. A main AUC peak,  $68 \pm 1\%$  of the remaining signal, with sedimentation coefficient of  $3.82 \pm 0.08$  S corresponds to the elongated monomer of Tom70 in its apo state ( $MW = 57.9 \pm 0.2$  kDa) (FigureS9D.). Only 30% of the total signal corresponds to a globular compact complex of one Tom70 and one Tim23 subunit (from NIS analysis:  $MW = 82.4 \pm 1.3$  kDa) with a sedimentation



coefficient of  $5.28 \pm 0.08$  S.

For Tom70-Ggc1 complex the loss of initial signal due to aggregates was higher, up to 50%. A main peak with sedimentation coefficient of  $3.76 \pm 0.02$  S for  $67 \pm 1\%$  of the total signal. The Non Interacting Species (NIS) analysis gives  $M_w = 66.2 \pm 1.1$  kDa. This peak corresponds reasonably to the elongated monomer of Tom70 in its apo state. For 15% of the total signal a peak with  $6.93 \pm 0.05$  S may correspond to a globular compact complex of one Tom70 with two Ggc1 preproteins. For 11% of the total signal, a peak at  $9.01 \pm 0.01$  S analysed by NIS analysis ( $MW = 179 \pm 33$  kDa) could correspond to a globular compact complex of 2 Tom70 with 2 Ggc1, however this could not be concluded with certainty since it could also correspond to some oligomeric state of the two proteins in our sample.

#### Assignment of the Ile<sup>δ1</sup> Met<sup>ε</sup> Thr<sup>γ</sup> methyl groups of the Tom70(39-617).

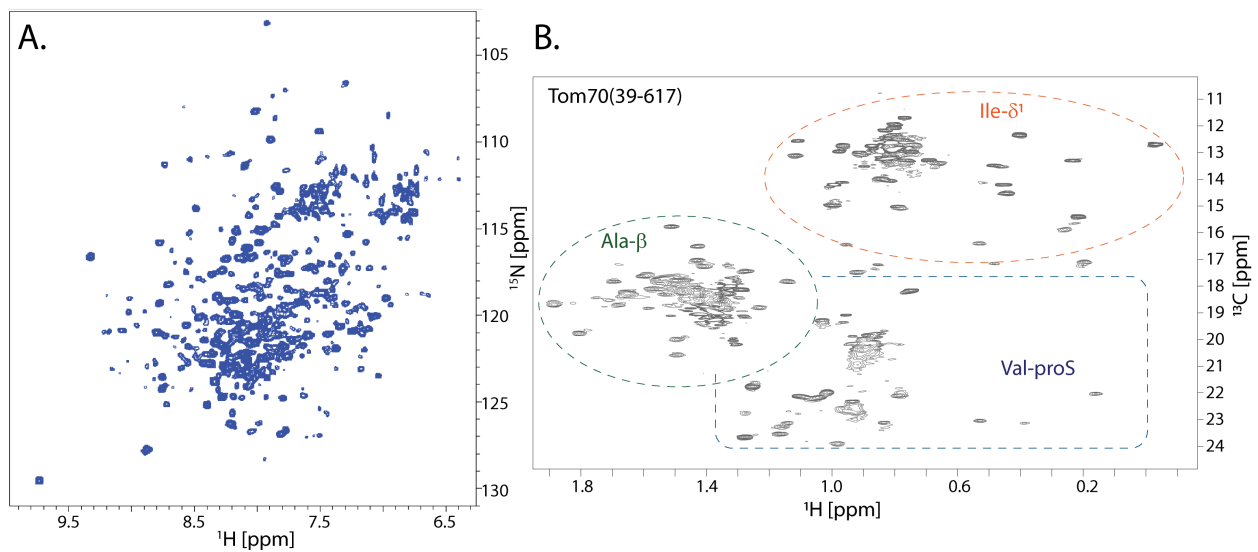


FIGURE S10: NH backbone (BEST-TROSY) and methyl ( $^1\text{H}$ ,  $^{13}\text{C}$  SOFAST-HMQC) spectrum of Ile<sup>δ1</sup>Val<sup>γ1</sup>Ala<sup>β</sup> groups of Tom70(39-617).

**A.** 2D NH spectrum of  $^2\text{H}$ ,  $^{15}\text{N}$ -labeled Tom70(39-617) shows wide peak dispersion as expected for a folded helical protein. Certain degree of peaks overlapping is expected for this 578 residue long construct.

**B.**  $^{13}\text{C}^1\text{H}_3$  spectrum of the Tom70(39-617) specifically labeled on the Ile, Val and Ala methyl groups. Twenty-six peaks were expected in the isoleucine region (orange ellipse), however we observe more than 30. Due to low resolution in the alanine region (green ellipse; 55 peaks expected) and valine region (blue square; 17 peaks expected) we did not continued with the assignment of these methyl groups.

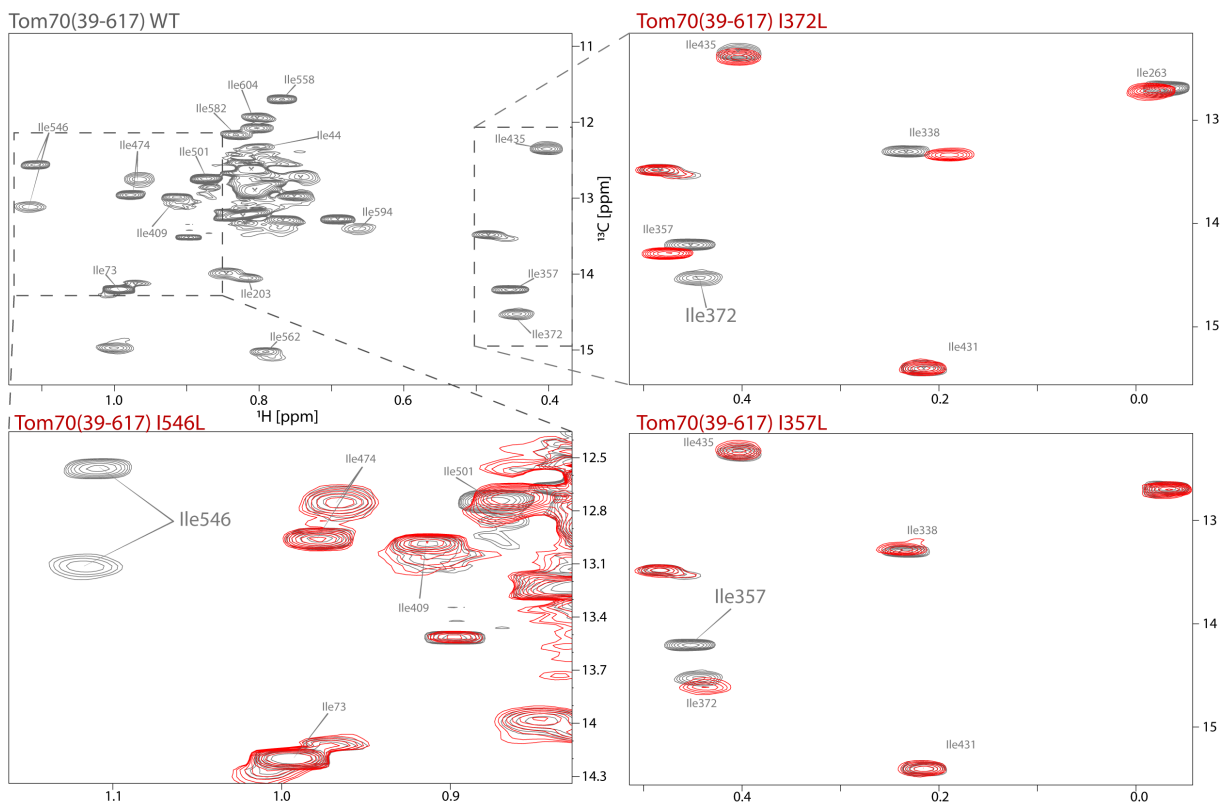


FIGURE S11: **Example of the isoleucine assignment process of Tom70.** Shown in gray is part of the isoleucine spectrum of the wild-type Tom70(39-617). Shown in red are some of the spectra of Tom70 variants that were used for the assignment of the Tom70 methyl groups. The same process was used for the methionine and threonine region.

The chemical shift assignment of the Met, Ile and Thr methyl groups of Tom70(39-617) construct was performed by the site specific mutagenesis approach where each Met, Thr and Ile amino acid was individually exchanged and from the difference in the mutant and wild-type  $^{13}\text{C}^1\text{H}_3$  spectra the resonances were assigned. For the residues that were hard to unambiguously assign by mutagenesis approach I used ShiftX software to predicted chemical shifts of Tom70 (input was structural model with Tom71 as a template as mentioned previously in the Introduction 1.4). This part of work was done with help from U. Guillermin and J. Schneider.

## MIT Tom70(39-617) 4XTS,2xIV

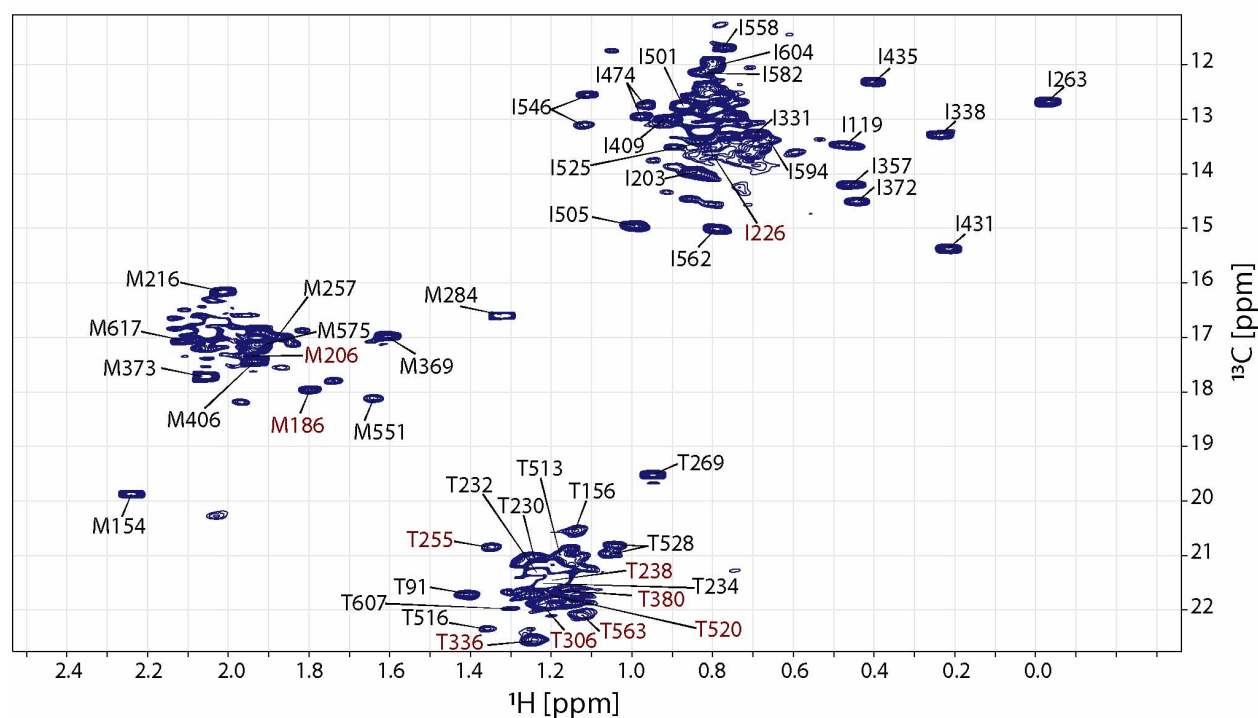


FIGURE S12: **Chemical shift assignment of the Ile<sup>δ1</sup> Met<sup>ε</sup> Thr<sup>γ</sup> <sup>13</sup>C<sup>1</sup>H<sub>3</sub> groups of Tom70(39-617).** <sup>13</sup>C<sup>1</sup>H<sub>3</sub> spectrum of the Tom70(39-617)4xTS,2xIV. This variant of Tom70 was used as a template for all the other variants used for the assignment of Met and Thr. The six amino acids that were exchanged are in the flexible N-terminal tail of the cytosolic Tom70(39-617) construct: T43S, T52S, T59S, T66S, I44V and I73V. These six amino acid substitutions (slightly) improved the resolution of the observed methyl spectra. The chemical shifts assigned with ShiftX are shown in red. Assignment of 100 % of Met, 87 % of Ile (δ1) and 77% of Thr methyl groups is accomplished.

## B.0.2 Kinetic parameters obtained with ITC experiments

TABLE S1: **Thermodynamic parameters of Tom receptor interactions from isothermal titration calorimetry (ITC) measurements.** For the ITC experiments yeast Tom22(1-97), yeast Tom70(39-617), yeast Tom20(37-183) and GB1-pSu9(1-69) were used. All samples were dialysed in the same 20 mM Tris, 100 mM NaCl, pH 7 buffer. The fitted parameters were: N (number of sites), K (binding constant in M<sup>-1</sup>) and ΔH (heat change in cal/mol) and the dissociation constant K<sub>d</sub> (in μM) was calculated as the reciprocal of the binding constant.

protein in the cell	titrant in the syringe	K <sub>d</sub> [μM]	N	ΔH [cal mol <sup>-1</sup> ]	final ratio	comment
Tom20 (70 μM)	pSu9 (700 μM)	18.28	1.48	-7375	1 : 3	
Tom20 (55 μM)	pSu9 (700 μM)	31.8	1.07	-9366	1 : 3.7	
Tom22 (50 μM)	pSu9 (700 μM)	error too big	-	-	1 : 3	aggregation
Tom20 (40 μM)	Tom22 (646 μM)	not detectable	-	-	1 : 3.2	
Tom20 (70 μM)	Tom22 (646 μM)	35.7	1.73	-6560	1 : 1.8	final ratio too low
Tom20:pSu9 1:1 (70 μM)	Tom22 (646 μM)	110	2	6104	1 : 1.8	sequential fit test
Tom20:pSu9 1:1 (50 μM)	Tom22 (646 μM)	as buffer				
Tom20 (70 μM)	Tom70 (993 μM)	4.04	1.42	-6988	1 : 2.8	
Tom20 (55 μM)	Tom70 (993 μM)	4.46	1.43	-6785	1 : 3.6	
Tom22 (40 μM)	Tom70 (650 μM)	30.9	1.28	-6821	1 : 3.25	
Tom22 (50 μM)	Tom70 (993 μM)	cannot be fitted				
Tom70 (50 μM)	pSu9 (764 μM)	weird jumps in signal	2x same behavior		1 : 3	

### B.0.3 Experimentally determined diffusion coefficients using NMR diffusion-ordered spectroscopy (DOSY).

TABLE S2: Experimentally determined diffusion coefficients using NMR diffusion-ordered spectroscopy (DOSY). Diffusion coefficient values extracted from the DOSY curves, obtained as integrals of one-dimensional <sup>1</sup>H spectra over either methyls, or amides as a function of the gradient strength.

NMR diffusion experiments (DOSY)					TRACT experiment
protein sample	conditions	D [cm <sup>2</sup> s <sup>-1</sup> ]	DOSY exp.	protein labeling	τ <sub>c</sub> [nsec]
Rn Tom20(51-145)	pH 7, 20°C	(10.2 ± 0.0) × 10 <sup>-7</sup>		monomer	5.7 ± 0.2
Rn Tom20(51-145)	pH 6.5, 20°C	(9.7 ± 0.0) × 10 <sup>-7</sup>			5.4 ± 0.2
Tom22(1-74)	pH 7, 27°C	13.6 × 10 <sup>-7</sup>	methyl, 13C filtered		
Tom22(1-74)GB1		10.2 × 10 <sup>-7</sup>			
Tom22(1-97)GB1	pH 7, 20°C	(7.9 ± 0.6) × 10 <sup>-7</sup> / (9.0 ± 0.2) × 10 <sup>-7</sup>	NH sel / 1H sel	<sup>15</sup> N Tom22	
Tom22(1-97)GB1+Tom70(39-617)		(5.7 ± 0.2) × 10 <sup>-7</sup> / (5.4 ± 0.4) × 10 <sup>-7</sup>	1H sel / NH sel		
Tom22(1-97)		(9.0 ± 0.1) × 10 <sup>-7</sup>	methyl DOSY	<sup>15</sup> N	
Tom22(1-97)+ScTom20(37-183)		(8.4 ± 0.2) × 10 <sup>-7</sup>	methyl		
Tom22(1-74)	pH 7, 25°C	(11.8 ± 0.2) × 10 <sup>-7</sup>	methyl	<sup>15</sup> N	
Tom22(1-74)+ScTom20		(12.5 ± 0.7) × 10 <sup>-7</sup>	methyl		
Tom22(1-97)GB1		(7.9 ± 0.6) × 10 <sup>-7</sup> / (9.0 ± 0.2) × 10 <sup>-7</sup>	NH sel / 1H sel	<sup>15</sup> N Tom22	
Tom22(1-97)GB1+Tom70(39-617)		(5.7 ± 0.2) × 10 <sup>-7</sup> / (5.4 ± 0.5) × 10 <sup>-7</sup>	1H sel / NH sel		
Tom22(1-74)	pH 7, 30°C	(9.9 ± 0.2) × 10 <sup>-7</sup> / (9.3 ± 0.0) × 10 <sup>-7</sup>	NH sel / 1H sel	<sup>15</sup> N Tom22s + residual GB1	
Tom22(1-74)+Xdj1 (1:2)		(8.8 ± 0.2) × 10 <sup>-7</sup>	NH sel		
Tom22(1-74)+Xdj1 (1:1)		(8.9 ± 0.2) × 10 <sup>-7</sup>	NH sel		
Tom22(1-74)	pH 6.5, 25°C	(10.0 ± 0.1) × 10 <sup>-7</sup>	methyl	DCN Tom22	
Tom22(1-74):Tom70(39-617) (1:2)		(6.0 ± 0.1) × 10 <sup>-7</sup>	methyl		
Tom22(1-74):Tom70(39-617) (1:0.5)		(7.7 ± 0.1) × 10 <sup>-7</sup>	methyl		
Tom20		(8.7 ± 0.3) × 10 <sup>-7</sup>	1H sel		
Tom20:Tom22(1-97)gb1		(5.1 ± 0.0) × 10 <sup>-7</sup> / (5.2 ± 0.0) × 10 <sup>-7</sup> / (5.3 ± 0.0) × 10 <sup>-7</sup>	1H sel / 13C filt / NH sel		

## B.1 Published work

The publications attached here are not presented in the main manuscript.



# How do Chaperones Bind (Partly) Unfolded Client Proteins?

Iva Sučec<sup>1\*</sup>, Beate Bersch<sup>1</sup> and Paul Schanda<sup>1,2\*</sup>

<sup>1</sup>CEA, CNRS, Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, Grenoble, France, <sup>2</sup>Institute of Science and Technology Austria, Klosterneuburg, Austria

## OPEN ACCESS

### Edited by:

Miguel Arbesú,  
Leibniz-Institut für Molekulare  
Pharmakologie (FMP), Germany

### Reviewed by:

Rina Rosenzweig,  
Weizmann Institute of Science, Israel  
James Bardwell,  
Howard Hughes Medical Institute  
(HHMI), United States

### \*Correspondence:

Iva Sučec  
iva.sucec@ibs.fr  
Paul Schanda  
paul.schanda@ist.ac.at

### Specialty section:

This article was submitted to  
Molecular Recognition,  
a section of the journal  
Frontiers in Molecular Biosciences

**Received:** 23 August 2021  
**Accepted:** 06 October 2021  
**Published:** 25 October 2021

### Citation:

Sučec I, Bersch B and Schanda P  
(2021) How do Chaperones Bind  
(Partly) Unfolded Client Proteins?  
*Front. Mol. Biosci.* 8:762005.  
doi: 10.3389/fmolb.2021.762005

Molecular chaperones are central to cellular protein homeostasis. Dynamic disorder is a key feature of the complexes of molecular chaperones and their client proteins, and it facilitates the client release towards a folded state or the handover to downstream components. The dynamic nature also implies that a given chaperone can interact with many different client proteins, based on physico-chemical sequence properties rather than on structural complementarity of their (folded) 3D structure. Yet, the balance between this promiscuity and some degree of client specificity is poorly understood. Here, we review recent atomic-level descriptions of chaperones with client proteins, including chaperones in complex with intrinsically disordered proteins, with membrane-protein precursors, or partially folded client proteins. We focus hereby on chaperone-client interactions that are independent of ATP. The picture emerging from these studies highlights the importance of dynamics in these complexes, whereby several interaction types, not only hydrophobic ones, contribute to the complex formation. We discuss these features of chaperone-client complexes and possible factors that may contribute to this balance of promiscuity and specificity.

**Keywords:** conformational ensemble, holdase, entropy, enthalpy, fuzzy complex, chaperone-client complexes, NMR spectroscopy

## 1 INTRODUCTION

Molecular chaperones are the essential components to ensure the protein homeostasis of the cell. Their importance is highlighted by their abundance in the cell: the family of 70 kDa heat-shock proteins (Hsp70) on its own, for example, is estimated to correspond to up to 3% of the total protein mass in eukaryotic cells under non-stress conditions Finka and Goloubinoff (2013). There are many types and isoforms of chaperones in each cell, and they generally are organized in cooperating networks Balchin et al. (2016). A central question in understanding chaperone function is how they interact with the polypeptides they bind, i.e., with their “client” proteins. How do chaperones achieve their ability to interact with many different client proteins efficiently while also retaining some kind of specificity? And how do the interactions between chaperones and their clients enable the clients to be refolded, safely transported or even disaggregated from insoluble forms? During the last few years, several complexes of chaperones with their full-length client proteins have been characterized at the atomic level, and have thereby shed light onto the underlying interaction patterns. In this review, we discuss the features of more than ten different chaperone systems, and provide insight into the interactions of these (predominantly folded) chaperones with their (predominantly unfolded) clients, and on how the balance of different types of interactions (hydrophobic, hydrophilic, electrostatic) may lay the basis for achieving some degree of promiscuity and some specificity. We invite the reader who wants to

quickly read only about the general common features that emerge from these examples to jump directly to **section 6**. We also refer to reviews on various aspects of chaperone-client complexes, e.g., those by Kim et al. (2013), Skjærven et al. (2015), Craig and Marszalek (2017), Hiller and Burmann (2018) or Rosenzweig et al. (2019).

## 2 BASIC CHAPERONE FUNCTION AND BINDING PROPERTIES

A basic property of a molecular chaperone is its ability to bind to partially or fully disordered client proteins. When bound to chaperones, these latter proteins are generally not in their native functional 3D structure. All proteins in the cell are, at some stage(s) of their life cycle, in such non-native states. This is obviously the case at the very start of a protein's presence in the cell: when being translated as an unfolded chain on the ribosome, the nascent chains may bind to chaperones before reaching their native fold. Other instances where chaperones are essential are when proteins unfold or misfold. Spontaneous unfolding may arise due to the fact that folded proteins are often only marginally stable, and the disruption of a few interactions within the crowded cellular environment can favor unfolded states [Christiansen et al. (2013), Gershenson et al. (2014)]. Moreover, some inherently insoluble proteins are produced in the aqueous environment of the cytosol, and need to be transported to a different cellular compartment; for example, membrane proteins destined to the bacterial outer membrane or the mitochondrial or chloroplast membranes rely on a suite of chaperones for their transport and insertion into the respective membranes (discussed in **section 5.1**, **section 5.2.1** and **section 5.2.2**).

The basic ability to interact with disordered proteins is common to all chaperones, at least in some of the conformational states that a given chaperone can adopt. This property is often referred to as “holdase” activity. Some chaperones are assumed to possess primarily, if not exclusively, a holdase activity; this is the case for prefoldin [Vainberg et al. (1998), Arranz et al. (2018)], mitochondrial TIM chaperones [Höhr et al. (2015), Becker et al. (2019)] and bacterial periplasmic chaperones [Skp, SurA and Dsb Goemans et al. (2014), Thoma et al. (2015)]. The role of these holdases is to safeguard their client from aggregation, and hand the protein to other downstream factors, such as other chaperones or insertases or degradation machineries (proteases). The case of J-domain proteins is somewhat particular: they exhibit a holdase function, and relay their clients to Hsp70 chaperones, and they also act upon the Hsp70 chaperone by enhancing the ATP hydrolysis of Hsp70 [Silver and Way (1993), Kampinga and Craig (2010)].

A chaperone that is able to assist the protein folding to its native state, is often described as having an additional “foldase” activity. The notion of foldase comes with the idea that the chaperone plays an active role, exerting some kind of force on its client Nunes et al. (2015). However, the distinction between a holdase and a foldase is not straightforward. Some proteins that

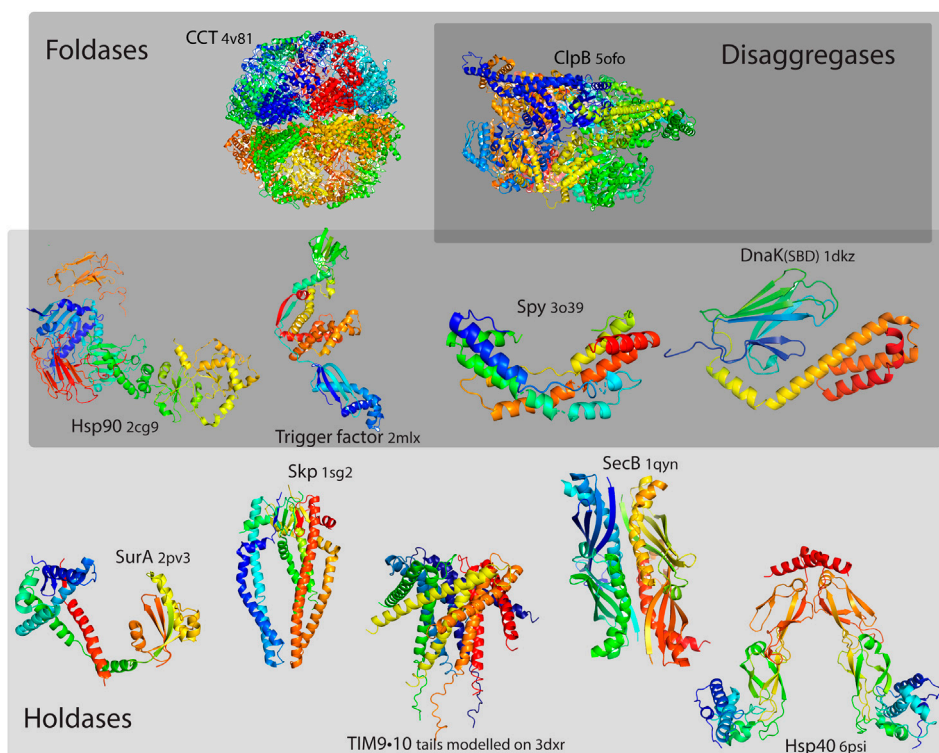
are often assigned a foldase function may be rather passive: a client protein may exploit the properties of the chaperone surface to facilitate its refolding (on the chaperone surface), or its unfolding, followed by spontaneous refolding [He et al. (2016), Stull et al. (2016)]. Few selected examples of molecular chaperones classified by their chaperoning properties could be seen in **Figure 1**.

### 2.1 ATP-Driven Chaperones: Using ATP to Alter the Chaperone While in Action

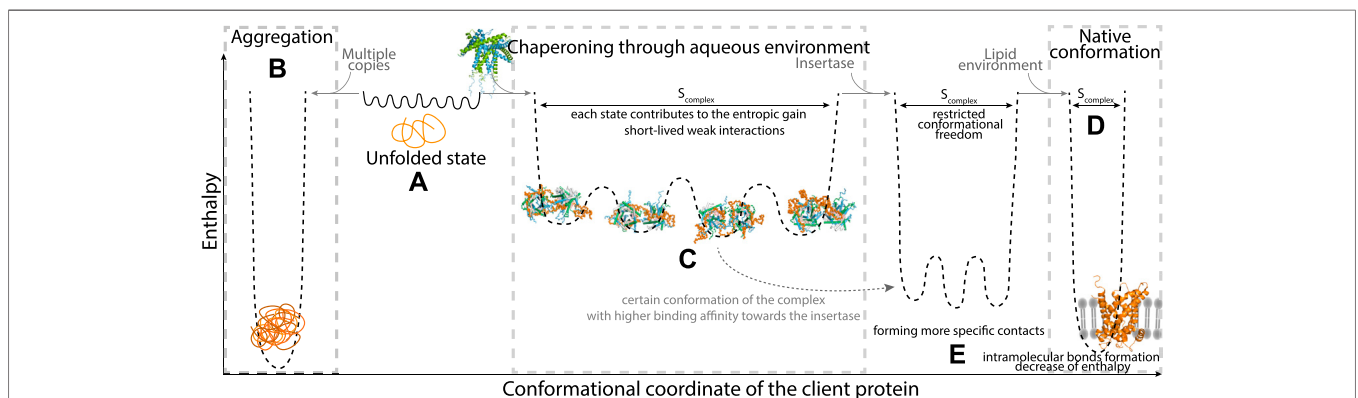
Some chaperones exploit furthermore an ATPase activity for their function. As in many motor proteins, ATP hydrolysis is exploited to drive conformational transitions in chaperones. A prominent example is Hsp70, in which ATP hydrolysis triggers a conformational transition which leads to a strongly increased client-protein affinity [Zhang and Zuiderweg (2004), Mayer and Bukau (2005)]. The ATP-driven structural transitions that some chaperones undergo can be seen as switches, by which a given chaperone changes its properties and, thus, its ability to act as a holdase. For example, in the large, barrel-shaped group II Hsp60 chaperonins (thermosome, TRiC), client binding occurs in the open state, at hydrophobic binding sites close to the barrel entry; it has been proposed that in the course of the ATP-driven structural transition, these sites become partly buried, which changes the environment of the client and presumably favors its refolding within the (predominantly hydrophilic) chamber, and its subsequent release to the cytosole [Kafri and Horovitz (2003), Meyer et al. (2003), Bigotti and Clarke (2005), Spiess et al. (2006), Reissmann et al. (2007), Nakagawa et al. (2014), Jin et al. (2019)]. Note that this is not the only mechanism by which Hsp60 assists refolding of its clients: refolding outside the chamber is another possible mechanism (Priya et al., 2013). Also other chaperones can employ different mechanism of chaperoning depending on the client and the clients folding pathway (reviewed in Koldewey et al., 2017 on the examples of Hsp60, Hsp70 and Spy).

It is worthwhile noting that it is a common mis-conception that the energy released upon breaking the bond to the  $\gamma$ -phosphate group in ATP is what drives large-scale conformational change. In fact, as any bond breaking, cleaving of  $\gamma$ -phosphate group in ATP does not release energy but rather requires it. This reaction is thermodynamically possible because the phosphoanhydride bonds are relatively weak and require less energy to break them than the energy released when stronger covalent bonds are formed in the product(s). Breaking the bond of the  $\gamma$ -phosphate group in ATP by nucleophilic attack from water (hydrolysis) or some electron-rich species, most commonly enables the energetically unfavorable reaction to occur by reaction coupling; for example, a phosphorylated product of one reaction is used as a reactant in the second reaction (Berg et al., 2002). In biological systems there are several ways ATP drives the conformational change and subsequently the activity of certain proteins. One of the examples is the sodium-potassium pump that undergoes its first conformational change upon ATP binding, the second conformational





**FIGURE 1** | Selected examples of molecular chaperones classified by their chaperoning properties. While all chaperones could be considered as holdases, with the function of binding its structurally unstable client and preventing its aggregation, foldases have an additional function of assisting the client protein folding to its native state, and disaggregases have an additional function in dissociating protein aggregates. Certain chaperones could have multiple functions. Delineating strictly between these functions is hardly possible, as chaperones may have different functions depending on their clients; thus, the frontiers are not to be seen as strict boundaries.



**FIGURE 2** | Schematic view of the thermodynamic properties of a client protein exemplified for a membrane-protein client (represented in orange). In the unfolded state of the client (**A**), the polypeptide exists in an ensemble of multiple rapidly-interconverting conformations. Depending on the concentration of the client (presence of multiple copies) or the presence of a chaperone, the conformational landscape carries different energetic properties. Aggregation, shown on left (**B**) is enthalpically favourable due to multiple intra- and inter-molecular hydrophobic interactions. In the presence of a dedicated chaperone, shown in the middle (**C**), the chaperone and client form favorable interactions (enthalpic contribution), and as the client generally stays highly dynamic there is no (or little) entropic cost, i.e., the entropy ( $S_{\text{complex}}$ ) remains large, comparable to the unfolded state shown in section A of this figure. For the membrane-protein clients, the pathway to the fully folded state (**D**) involves engaging with an insertase, which promotes folding of the client (**E**). Certain conformations out of the complex ensemble may have higher affinity towards the insertase. The interaction of client and insertase may lead to a step-wise dissociation of the client from the chaperone and formation of more specific contacts with the insertase. The lower conformational entropy ( $S_{\text{complex}}$ ) may be compensated by favorable enthalpic interaction, or entropy gain from release of structured water molecules.

change once the protein gets phosphorylated by the  $\gamma$ -phosphate group upon cleavage of ATP and the last one induced by proteins de-phosphorylation. [reviewed in Jorgensen et al. (2003)]. What drives structural transitions of other ATP-fueled machines, such as the ATP-driven chaperones, is the fact that upon hydrolysis two new species, ADP and phosphate, are generated. Their binding properties and charges differ from those of ATP, and these altered properties of the complex drive a conformational rearrangement of the protein [Bagshaw and Trentham (1973), Hwang and Karplus (2019)].

In the present review we focus on ATP-independent chaperones, and their properties as holdases; as stated above, the holdase activity is common to all chaperones, including ATP-fueled ones.

### 3 CHAPERONE-CLIENT COMPLEXES AND THE BALANCE OF STABILITY VS. EASE OF RELEASE

The complexes of chaperones with their clients need to fulfill contradicting requirements: on one hand the complexes, at least those of some holdase chaperones, need to be at least somewhat stable, such that the client protein does not spontaneously detach from the chaperone. This property is important particularly for “transfer chaperones”, which accompany highly aggregation-prone clients, such as membrane proteins, to their target insertase or translocase. Spontaneous detachment of the client before the complex reaches its destination may lead to aggregation of the client. On the other hand, release of the client, for example the handover to a downstream insertase or the release of the natively re-folded client, should proceed without a significant energy barrier. From some chaperones (e.g., Spy, see below), the clients detach spontaneously once they reached a conformation allowing the detachment; other chaperone–client complexes dissociate once they reach, e.g., a membrane-protein insertase or translocase, and need to detach without significant energy barrier. (The relay often proceeds without energy input from ATP hydrolysis, e.g., in the bacterial periplasm or the mitochondrial intermembrane space.) Chaperones, therefore, must meet the contradicting requirements of stability and absence of significant energy barriers for dissociation [Hartl et al. (2011), Burmann et al. (2013), Hiller and Burmann (2018)].

As discussed here, using recent examples that have been characterized at the structural level, dynamics within the bound state is the way how this apparent contradiction can be resolved. In this sense, chaperone–client complexes may be seen as “fuzzy” complexes. This term, introduced by Fuxreiter and Tompa Tompa and Fuxreiter (2008), refers to protein–protein complexes in which at least one of the two proteins remains dynamic while bound. In many reported cases the bound client is disordered, and bound to the chaperone as an ensemble of inter-converting states.

Another example of importance of dynamics in chaperone–client interaction is increase in flexibility of the linker loop at the

substrate interface of a chaperone Spy, which is proposed to increase promiscuity of Spy Horowitz et al. (2016).

## 4 SPECIFICITY VS. PROMISCUITY

The complexity of living organisms relies on the promiscuity of proteins: enzymes capable of processing a range of substrates or receptors able to recognize different molecules, and also the molecular chaperones, which are able to bind to a range of client proteins. Promiscuity is essential, because if each chaperone was highly specific to a small set of client proteins, the energetic cost of maintaining many different regulatory networks would be very high Cumberworth et al. (2013).

Promiscuity is often assumed to be the rule for chaperones; trigger factor, for example, has a substrate proteome with more than 170 members Martinez-Hackert and Hendrickson (2009); the TIM9.10 chaperone binds at least 40 different proteins; the family comprising the 70 kDa heat-shock protein (Hsp70) and Hsp90 family has a very wide clientome that covers at least 20% of the yeast proteome, for example Taipale et al. (2010).

Despite the presence of many chaperones in the cell, each capable to bind a broad range of clients, cellular experiments generally reveal preferences, and not all chaperones bind a given client. This can be nicely illustrated with the example of mitochondrial protein import. Herein, proteins which are destined to the mitochondria but produced in the cytosol need to be guided along the entire way; this is particularly important for mitochondrial membrane proteins, because of their strong tendency to aggregate. A central question in the mitochondrial import field is which chaperones are responsible for the transport of the newly synthesized mitochondrial precursor proteins from the cytosolic ribosomes to the mitochondria.

Insightful studies Jores et al. (2018) have revealed, for example, that newly synthesized outer-membrane  $\beta$ -barrel porins associate with Hsp70 and Hsp90 chaperones as well as with a certain number of Hsp40 chaperones (Ydj1, Sis1, and Djp1) and Hsp104, but not other general chaperones such as Hsp60 or 14-3-3. When doing the same assay with mitochondrial inner-membrane proteins, Hsp70 and Hsp90 are again found to associate, but the levels of associated Hsp40 chaperones differ from the ones found to associate with the  $\beta$ -barrel forming porin. This kind of experiments suggests that Hsp70 and Hsp90 interact with client proteins mainly by hydrophobic interactions, but that the Hsp40 association is based on more subtle differences in their substrate proteins, allowing also to fine-tune the specificity.

Given such kind of experimental findings, the central question is: what are the structural or sequence properties of the client proteins which make one chaperone bind but not another one? And how does the chaperone recognize these differences?

Unfortunately, these questions are not well understood. A few rather rare cases are known where a chaperone is highly specific, with only one Szolajska and Chroboczek (2011) or few clients Kuehn et al. (1993); in these cases, the recognition is achieved by



complementary surfaces and a well defined binding site on the client. For example, the periplasmic holdase chaperones PapD and FimC from pathogenic bacteria, are specific for the pilus forming sub-units Sauer et al. (2000) and they interact with their clients via the donor-strand complementation mechanism. Such surface-/strand-complementarity resembles the situation of complexes formed between folded globular proteins. These cases are rather rare and not representative of most chaperones.

For most chaperones, identifying the binding motif, or even clarifying the clientome, is not as easy. In the clients of the Hsp70 chaperone family, a binding motif has been identified. It consists of a hydrophobic core and two flanking regions with basic residues Rüdiger et al. (1997); this motif is indeed very abundant in most proteins. However, for the very important Hsp90 family and its very large clientome, bioinformatic analyses have not been able to identify a specific binding motif Taipale et al. (2010). In some cases, such as the small TIMs discussed in **section 5.1**, preferences for binding one rather than another client is based on a combination of interaction types within the same binding groove or on different binding sites: a hydrophobic interaction with one binding interface, and a polar/charge-based contact at a separate binding interface.

The specific recognition of clients is related to the process by which clients are targeted to organelles within the cell; for example, precursor proteins destined to chloroplasts or mitochondria need to be recognized by a set of chaperones/receptor domains for their import. In some cases, these targeting signals are very well defined amino acid sequences von Heijne (2002). However, many targeting sequences are less well defined and scattered over the primary structure, that share certain physical properties rather than exact amino acid sequence. For example, mitochondrial preproteins targeted for the matrix carry an amphiphilic helix as a cleavable recognition signal, however the only similarity between the targeting signal of different proteins is that one side of the helix is positively charged whereas the other is hydrophobic Maduke and Roise (1996). Chloroplast outer membrane proteins carry a targeting signal, rather than an exact targeting sequence, recognized by the cytosolic AKR2 (Ankyrin repeat protein) chaperone. In addition to the importance of moderate hydrophobicity of the targeting signal, positively charged residues flanking transmembrane domains are important for specificity, and if these positive charges are missing, the preproteins are targeted to the plasma membrane rather than to the chloroplast Lee et al. (2011).

The mechanisms by which a client is selected by a chaperone and not bound by another one are far from being solved. Deciphering the recognition mechanisms is complicated by the fact that these complexes are often highly dynamic. Therefore, specific contacts with which folded proteins recognize each other, e.g., salt bridges, tend to be short-lived. It is currently poorly understood how specific recognition is compatible with the highly dynamic character of the chaperone–client complexes. The general mechanisms that underlie these complexes, described in the examples below,

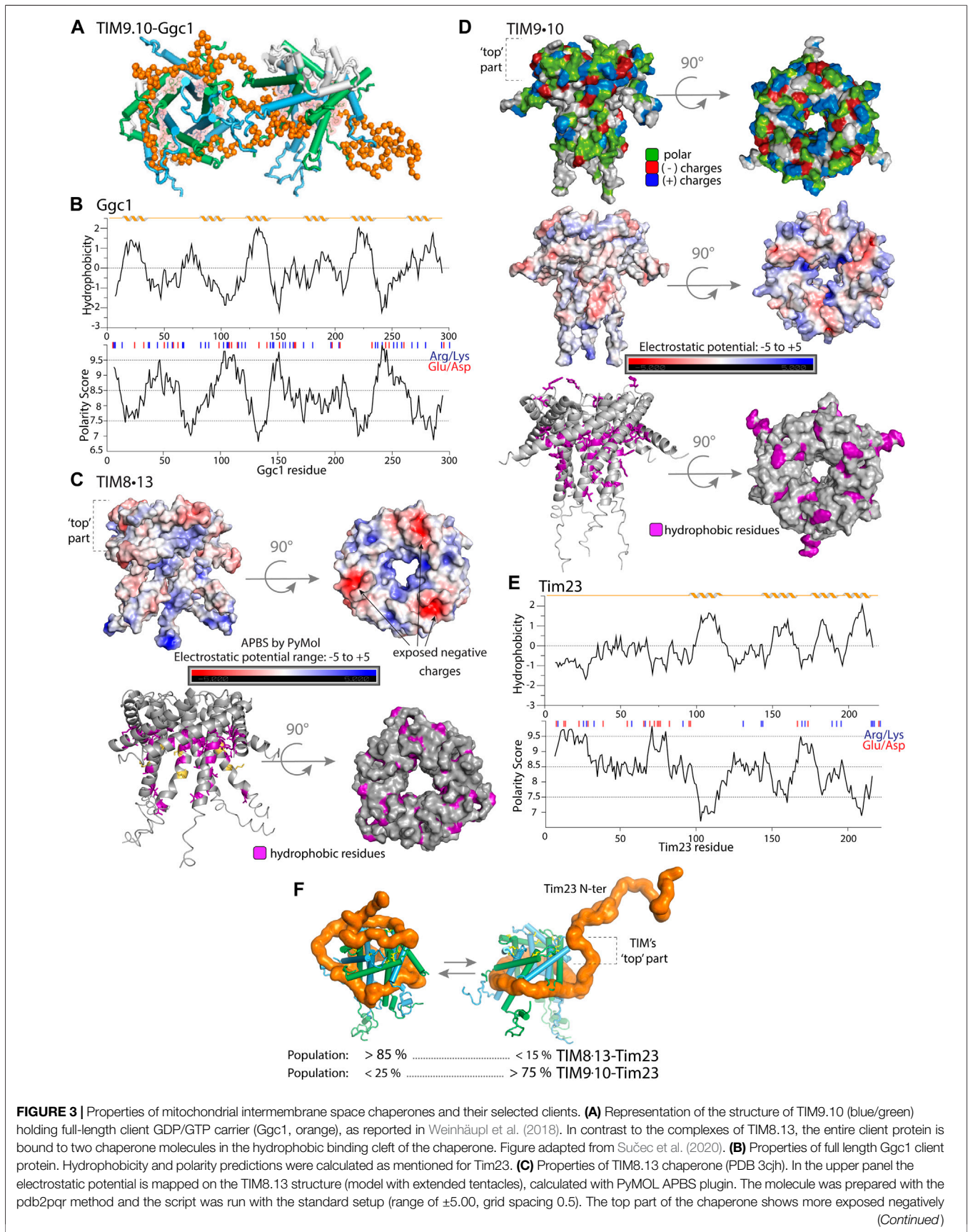
may nonetheless provide possible routes how specificity may be achieved.

## 5 LESSONS LEARNED FROM ATOMIC-LEVEL STUDIES OF CHAPERONE COMPLEXES

### 5.1 Chaperoning in the Mitochondrial Intermembrane Space

The vast majority of the mitochondrial proteins are imported into the organelle in a post-translational manner as precursor proteins, that are recognized either by a cleavable pre-sequence or by an internal targeting sequence. Cytosolic chaperones transport these precursor proteins to the mitochondrial entry gate, the translocase of the outer membrane (TOM) complex [reviewed in Becker et al. (2019)]. Depending on their final destination, the precursor proteins are then either inserted into the outer membrane from the mitochondrial outside, or directly relayed to the translocase of the inner membrane (TIM23 complex), or transferred across the intermembrane space (IMS) to an insertase in the inner or outer membrane. Here we focus on this latter process and its associated chaperones, namely the translocation of membrane-protein precursors in the intermembrane space. The so-called small TIM chaperones are responsible for this safe translocation of the highly aggregation-prone membrane-protein precursors across the aqueous IMS compartment [Koehler et al. (1998a), Koehler et al. (1998b), Wiedemann et al. (2001), Vergnolle et al. (2005)]. In their apo state, these chaperones form hetero-hexameric complexes of ~ 65 kDa, composed of alternating subunits of Tim8 and Tim13 [TIM8.13, Beverly et al. (2008)] or Tim9 and Tim10 [TIM9.10, Webb et al. (2006)], or Tim9, Tim10 and Tim12 [TIM9.10.12, Gebert et al. (2008)]. It is interesting to note that the TIM chaperones are in continuous exchange, whereby subunits co-exist between the hexameric state (predominant at ambient temperature) and monomeric forms. At room temperature, approximately 10% of the Tim9 and Tim10 subunits coexist as monomers, together with the hexameric TIM9.10 complex. NMR experiments have established that the integration of subunits into the hexamer (and release of subunits from the hexamer) is slow, on time scales of many tens of minutes Weinhäupl et al. (2021).

As the only known chaperone system of the mitochondrial IMS, TIM chaperones are crucial for the recognition and transfer of most of the mitochondrial membrane precursor proteins Morgenstern et al. (2017). They recognize a broad range of membrane proteins and transfer them in an unfolded state from the mitochondrial outer membrane pore (TOM, with its central pore formed by Tom40), through the aqueous mitochondrial intermembrane space, towards the insertases of the inner membrane (TIM22) or outer membrane (SAM) [Koehler et al. (1999), Bauer et al. (2000), Rehling et al. (2003), Paschen et al. (2003), Kozjak et al. (2003), Gentle et al. (2004), Lithgow and Schneider (2010)]. As for many chaperone-substrate pairs, where the substrate is often aggregation prone,



**FIGURE 3** | charged residues, lower electrostatic potential, compared to the rest of the protein and compared to TIM9.10 (panel E). Mapped conserved hydrophobic residues on TIM8.13 model (purple, **lower panel**). Yellow: polar residues in the conserved hydrophobic positions leading to weaker binding of all hydrophobic clients. **(D)** Properties of TIM9.10 chaperone (PDB: 3dxx). Upper panel: positive (blue), negative (red) charges and polar (green) residues. Electrostatic potential mapped on TIM9.10 model with extended tentacles (**middle panel**). Lower panel: conserved hydrophobic residues (in the binding cleft and on the “top”) on TIM9.10 model, shown in purple. In the top view, all hydrophobic residues of TIM9.10 are shown. **(E)** Properties of full length Tim23 client protein. Upper plot: hydrophobicity prediction (Kyte-Doolittle scale); lower plot: polarity prediction (Grantham scale, Expsy Bioinformatics Resource Portal). Red and blue bars: negatively and positively charged residues. NMR spectra of the soluble Tim23 N-terminal fragment (residues 1–98) in isolation show the hallmark features of a highly flexible intrinsically disordered protein Sućec et al. (2020). **(F)** Representative conformations of TIM chaperones bound to Tim23 client protein in which the hydrophilic N-tail of Tim23 is either bound or unbound on the ‘top’ part of the chaperone. The best-fit populations of the two classes of states are shown for either TIM9.10 or TIM8.13 as derived from SAXS/MD. Figure adapted from Sućec et al. (2020).

protecting the clients from misfolding and aggregation requires that these complexes do not dissociate spontaneously, i.e., the overall affinity of chaperone and client needs to be relatively strong. It is difficult to experimentally determine dissociation constants ( $K_d$ ) because one cannot obtain these complexes by simply titrating solutions of chaperone and client due to the insolubility of the membrane-protein precursors. For the small TIM chaperones, no experimental  $K_d$  values are available. However, it has been experimentally shown that the transfer of client proteins (membrane proteins of the inner membrane, so-called mitochondrial carriers) from one TIM chaperone to another takes approximately 4 h Weinhäupl et al. (2018), indicating high stability of the client-TIM complex in the absence of the downstream insertase of the inner membrane. In the cell, release at the insertase is presumably energetically more favorable (**Figure 2**), as the time scale for import into mitochondria is rather in the minutes time scale Wiedemann et al. (2006). It is likely that interactions of a part of the (highly dynamic) precursor protein with the insertase complex lowers the energy barrier for release (**Figure 2**).

When TIM chaperones bind clients, they do not undergo significant changes of their structure, nor of their backbone or sidechain dynamics, as revealed by NMR methods Weinhäupl et al. (2018). Interestingly, also the above-mentioned exchange of subunits between monomers still exists when the client is bound to the hexameric TIM9.10. Subunits still enter and exit the hexamers, with a time scale similar to the apo chaperone Weinhäupl et al. (2021). The very different size and thus NMR properties of the chaperone-client complexes and monomers makes it difficult to quantify the populations with precision, but the data suggest that even the ratio (monomeric subunits vs. hexamer) is similar to the apo state. This is different to the Skp chaperone or DegP (see below), which assemble upon client binding.

The  $\alpha$ -helical inner-membrane protein clients are wrapped around the TIM chaperones, using a cleft that is formed by highly conserved hydrophobic residues. The fact that these hydrophobic patches are in a cleft presumably helps to shield them, such that TIM chaperones do not aggregate by intermolecular hydrophobic interactions. Depending on the client length, a single client can recruit up to two TIM9.10 chaperones for the transfer (**Figure 3A**), or possibly even more than two chaperones (although this has not been shown yet). The clients are, thus, in extended conformation, which is quite different from the compact “fluid-globule” state that OMPs adopt in the bacterial membrane-protein chaperone Skp (see below). Interestingly, the

clients, both  $\alpha$ -helical and  $\beta$ -barrel-forming proteins, have some degree of secondary structure; some residual  $\alpha$ -helical propensity was detected for inner-membrane proteins Weinhäupl et al. (2018), and for  $\beta$ -barrel clients it was shown that only clients with pre-formed  $\beta$ -turns bind efficiently Jores et al. (2016). For the latter case, this appears intuitive, because in a  $\beta$ -turn of  $\beta$ -barrel outer-membrane proteins one face is hydrophilic and the other is hydrophobic. A  $\beta$ -hairpin conformation would ensure exposure of an entire hydrophobic patch for efficient binding to the hydrophobic cleft of the chaperone.

The bound clients are highly dynamic, adopting multiple inter-converting, short-lived conformations, while staying bound on the chaperone surface, in an unfolded, extended state Weinhäupl et al. (2018). These inter-conversion dynamics occur on a time scale of ca. 1 ms: specific NMR methods probing this time window (relaxation-dispersion NMR) have highlighted extensive conformational exchange in the complex Weinhäupl et al. (2018). The overall high affinity is thus achieved by multiple contributions from weak interactions, primarily within the conserved hydrophobic cleft. The authors have proposed that the dynamics of the client protein in the bound state enable the successful transfer of the client protein to the insertase without significant energy barrier: in some of these inter-converting client-chaperone conformations certain parts of the client protein, those with a higher affinity for the insertase, are transiently detached from the complex. Upon interaction with the insertase the clients are gradually released from the chaperone without significant energetic barrier. The favorable enthalpic contribution of the client’s folding makes this process energetically favorable (**Figure 2**). The dynamics of these chaperone-client complexes are the key to reconcile two apparently contradicting requirements: high overall stability and a low energy barrier for release. Here, the avidity of many individually weak interactions ensures high overall complex stability, whereas transfer to the downstream insertase proceeds step-wise, breaking only a few, weak interactions at a time, and therefore without the need for a large activation energy to be overcome. This kind of mechanism has been proposed for the functionally similar but structurally very different Skp complexes Burmann et al. (2013).

It has also been investigated which parts of the client interact with and are important for binding to the chaperones. Regarding the transmembrane part of the clients, e.g., the ADP/ATP carrier (Aac) or the GDP/GTP carrier of the inner membrane, the above-mentioned integrated NMR study has provided only limited information on a residue-by-residue basis. The millisecond



dynamics of the client within the binding cleft leads to severe line broadening and relatively poorly resolved spectra of the clients Weinhäupl et al. (2018), and has hampered the identification of specific residues of the clients that bind to the cleft. Some information about the relative importance of different parts of the client Aac comes from interaction studies with peptide fragments Curran et al. (2002). Briefly, a peptide scan with 13-residue-long fragments along the sequence of Aac, each overlapping with its predecessor by 10 residues, was performed; the peptide fragments were covalently linked to a cellulose membrane, incubated with TIM9.10, and the quantity of bound chaperone was assessed by immunodetection ( $\alpha$ -Tim10). By far the highest complex yield was achieved for fragments derived from the transmembrane helices, while only very small amounts were detected for fragments from the hydrophilic loop regions. Although this data does not provide residue-specific information, it establishes that hydrophobic fragments are important for binding. This finding is expected, given the hydrophobic character of the binding groove on the TIM chaperone.

Analysis of the structural properties of the TIM-client complexes have shed light onto the types of interactions that are crucial for complex formation. The two TIM chaperones, TIM8.13 and TIM9.10, are structurally very similar, and they both have highly conserved hydrophobic residues located in the cleft formed between the inner (N-) and outer (C-terminal)  $\alpha$ -helices. However, they have different specificity towards the mitochondrial precursor proteins [Weinhäupl et al. (2018), Sućec et al. (2020)]. For membrane-protein clients consisting essentially of transmembrane spanning parts, such as mitochondrial solute carriers of the inner membrane like the ATP/ADP carrier (Aac) Curran et al. (2002) and the outer membrane  $\beta$ -barrel proteins [Hoppins and Nargang (2004), Habib et al. (2005)], TIM9.10 has higher binding affinity than TIM8.13 [ca. 10-fold higher Weinhäupl et al. (2018), Sućec et al. (2020)]. However, TIM8.13 has higher affinity for binding membrane-protein clients with an additional soluble and more hydrophilic domain, such as Tim23, the translocase of the inner-membrane (TIM23) complex Davis et al. (2007), and the aspartate-glutamate carriers Roesch et al. (2004). TIM9.10's native clients are all-transmembrane mitochondrial precursor proteins which are highly hydrophobic, such as the mitochondrial carriers, of which one representative is shown in **Figure 3B**, although TIM9.10 is also able to bind to e.g., Tim23.

How can this somewhat different client specificity of the two overall very similar chaperones be explained? Two regions on TIM chaperones have distinct properties, and are, thus, presumably responsible for the client specificity. Firstly, certain residues within the conserved hydrophobic cleft are less hydrophobic (Lys, Ser) in TIM8.13 compared to the corresponding positions in TIM9.10 (see the orange residues in **Figure 3C**, lower panel). As a consequence, TIM8.13 might be less able to hold the transmembrane, hydrophobic parts of mitochondrial preproteins than TIM9.10. In a recent study it was shown that a TIM8.13 mutant, in which these more hydrophilic residues were changed to hydrophobic ones, was

much more capable of holding all-transmembrane (TIM9.10) clients Sućec et al. (2020).

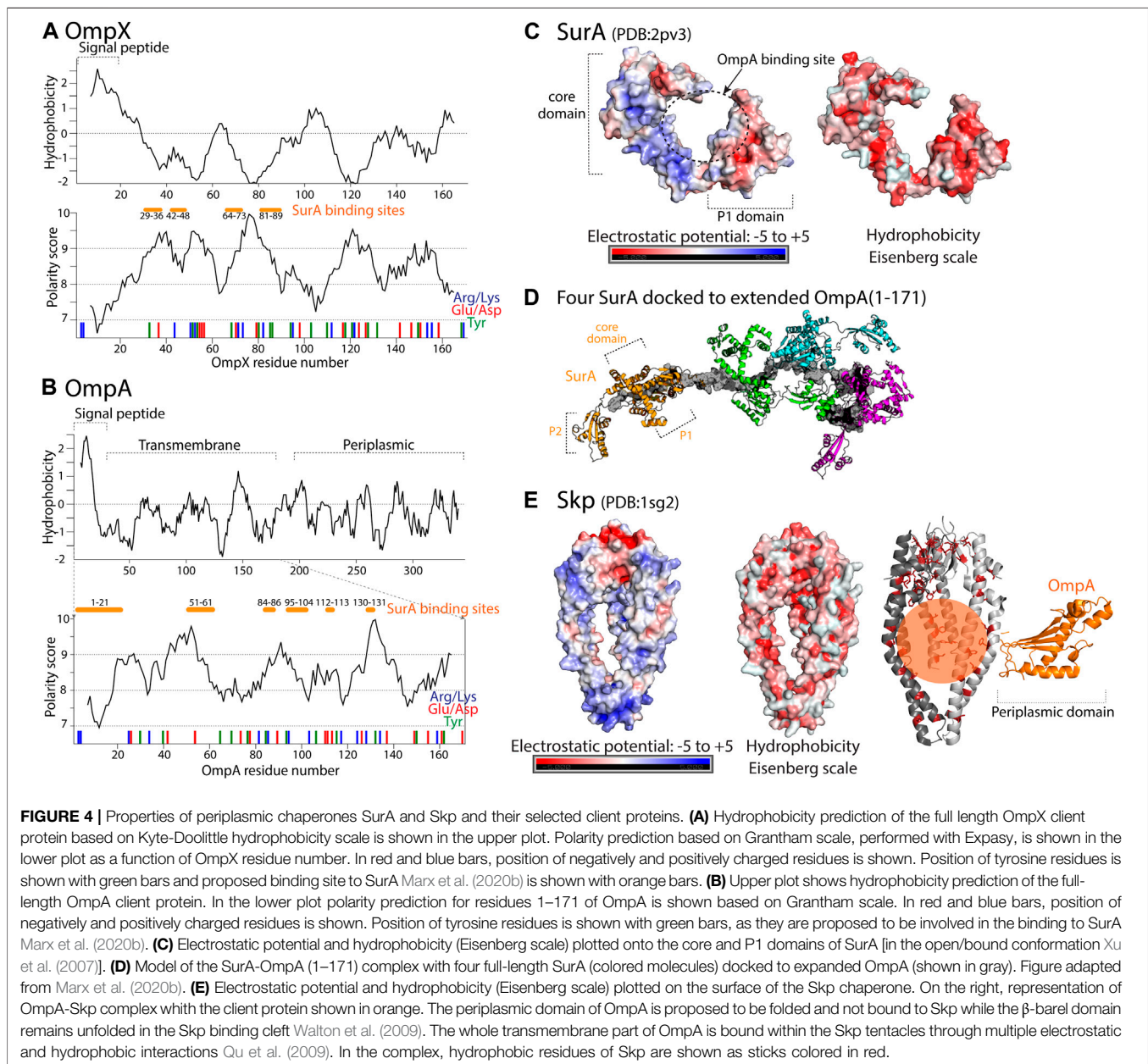
A second difference is found in the top part of the chaperones, where TIM9.10 differs from TIM8.13 in polarity and charge (**Figures 3C,D**). TIM8.13 uses additional hydrophilic interactions for binding the N-terminal region of Tim23 via the top part of the chaperone (**Figure 3C**). It is noteworthy, however, that single-point mutations in this top part of the chaperones did not swap the client affinities of TIM9.10 and TIM8.13, unlike in the above-described case where the hydrophobic cleft of TIM8.13 was rendered more hydrophobic. This suggests that the interaction with the hydrophilic part of Tim23 involves a more complex pattern than could be resolved by the few mutations introduced Sućec et al. (2020).

NMR data of the soluble Tim23 fragment show that TIM9.10 hardly interacts with this predominantly polar fragment; in fact the detected interaction involves only a patch of hydrophobic residues at the N-terminus of Tim23 (MSWLFG) and a further stretch with increased hydrophobicity Sućec et al. (2020). TIM8.13 interacts much more strongly with the soluble fragment of Tim23, and the interaction involves a stretch of at least 35–40 residues of Tim23.

Taken together, the current view is that the hydrophobic binding cleft of small TIMs enables (promiscuous) binding to the hydrophobic transmembrane parts of the clients, whereby TIM8.13 is less hydrophobic and thus less performing in this binding; additional hydrophilic interactions compensate to some degree for this lower affinity, depending on the client. In the case of Tim23 (**Figure 3E**), the resulting ensembles of states of the TIM9.10 or TIM8.13 complexes have different population levels, as revealed by SAXS/MD data. **Figure 3F** recapitulates this situation for Tim23-binding to the two chaperones. The states in which the hydrophilic tail of the client interacts with the hydrophilic top part of the chaperone are much more populated in the case of TIM8.13 than TIM9.10, and these hydrophilic interactions compensate for the inherently lower ability of TIM8.13 to interact via its hydrophobic cleft with the transmembrane part.

## 5.2 Chaperoning in the Bacterial Periplasm

In Gram-negative bacteria, outer membrane proteins (OMPs) are synthesized on the cytoplasmic ribosome and translocated in an unfolded form across the inner membrane by the Sec machinery [reviewed in Komarudin and Driessen (2019), Oswald et al. (2021)]. To reach their final destination, they have to cross the periplasm, an aqueous compartment. At the exit of the Sec machinery, OMPs are taken in charge by periplasm-specific chaperones. The bacterial periplasm is a special and somewhat demanding environment for chaperones, because 1) it lacks ATP, 2) it is an oxidizing environment and 3) it is separated from the outside only by a porous membrane, and is therefore particularly susceptible to changes in the outside. The periplasm of gram-negative bacteria contains ATP-independent chaperones that contribute to the biogenesis of OMPs Sklar et al. (2007): the holdases SurA, Skp, FkpA and PpiD as well as DegP, which has



chaperone and protease functions [Arié et al. (2001), Matern et al. (2010), Merdanovic et al. (2011), Ge et al. (2014)]. The membrane anchored chaperone PpiD contributes to the efficient detachment of newly secreted OMPs from the Sec machinery Fürst et al. (2018). Other periplasmic chaperones such as SurA, DegP, and Skp are likely to take over newly translocated proteins from PpiD on their way into the periplasm or to the outer membrane. Outer membrane protein biogenesis in Gram-negative bacteria is reviewed in detail, e.g., in Rollauer et al. (2015), Schiffrin et al. (2017), Tomasek and Kahne (2021).

The bacterial periplasmic chaperones Skp (Seventeen Kilodalton Protein; Chen and Henning (1996)) and SurA [Survival factor A, Behrens et al. (2001), Bitto and McKay

(2002)] share the pool of OMP clients; initially, they appeared to have redundant function in escorting the OMPs to the BAM complex of the outer membrane [Sklar et al. (2007), McMorran et al. (2015)]. Recent work Wang et al. (2021) indicates that the SurA chaperone, with the PPIase (peptidyl-prolyl *cis-trans* isomerase) activity Sklar et al. (2007), has a role in targeting OMPs to the BAM complex, while the chaperone Skp delivers unintegrated OMPs to the DegP for their degradation. The client proteins that have been extensively used as models for interactions, OmpX and OmpA, share similar hydrophobic and polar properties of their amino-acid sequence in the transmembrane parts (Figures 4A,B); OmpA has an additional soluble periplasmic domain on the C-terminus (but several studies used only the TM part as client).

### 5.2.1 SurA

SurA is the key chaperone for insertion of OMPs into the OM by the help of the BAM complex Sklar et al. (2007). It is composed of three domains: the core domain which is formed by the N- and C-terminal regions and two central PPIase domains (P1 and P2). The crystal structure of SurA has been determined in its apo form Bitto and McKay (2002). In this structure, P1 is bound to the core domain whereas P2 is connected to the core domain by an extended linker. In solution, SurA appears to be monomeric [Calabrese et al. (2020), Marx et al. (2020a)] and samples at least two different conformational ensembles that are distinct from the crystal structure Calabrese et al. (2020). The different conformations are in dynamic interconversion on a submillisecond time scale. The hydrodynamic properties of SurA reveal a radius of gyration compatible with one of the two PPIase domains being spatially separated from the core structure. SurA switches between a major P1-closed and a minor P2-closed state [Marx et al. (2020a), Jia et al. (2020)]. It has been shown that the two PPIase domains compete in a non-allosteric manner with each other for binding to the core domain Marx et al. (2020a).

The presence of either P1, P2, or both PPIase domains was shown not to be required for chaperoning activity of SurA, as all three mutants fully complement  $\Delta$ surA OMP assembly phenotypes in an otherwise unmodified genetic background Soltes et al. (2016). Whereas other chaperones (Skp, DegP) assemble to form a cage-like cavity for client binding (see below), SurA does not oligomerize. Client binding and the client complexes have been studied using different techniques including NMR Thoma et al. (2015), single molecule fluorescence Chamachi et al. (2021) and cross-linking mass spectrometry [Calabrese et al. (2020), Marx et al. (2020b)] experiments. It has been shown that SurA recognizes substrates with a preferential Ar-X-Ar motif (Ar: aromatic amino acid, X: any type of residue) Bitto and McKay (2003). Such a tripeptide is found at the C-terminus of many OMPs and this so-called  $\beta$  signal has been proposed to play a role in the recognition of OMP clients by the BAM complex [Hennecke et al. (2005), Wang et al. (2021)].

NMR and single-molecule fluorescence have shown that the SurA-bound OMP client proteins form a rapidly exchanging conformational ensemble with exchange rate constants on the microsecond timescale [Thoma et al. (2015), Chamachi et al. (2021)]. Crosslinking mass spectrometry experiments [Calabrese et al. (2020), Marx et al. (2020b)] revealed that SurA binds the OMPs in a groove formed between the core and P1 domains. This groove forms when the two PPIase domains are simultaneously dissociated from the core Marx et al. (2020b) and the role of the P2 domain is regulating the interaction between the P1 and the core domains Calabrese et al. (2020). The OMP binding site on SurA is large enough to accommodate an entire transmembrane  $\beta$ -strand or  $\beta$ -hairpin. The bottom of the groove is formed by a 30 Å-wide hydrophobic stretch and is positively charged (Figure 4C, left panel). Regions of the core and the P1 domain outside of the groove bear negative charges, suggesting that electrostatic interactions contribute to the complex formation between SurA and its

clients. These features may be required to accommodate the alternating hydrophobic–hydrophilic patterns of extended OMP transmembrane domains. Interestingly, the non-client protein OmpLA showed much less cross-linking products than the SurA clients OmpA (transmembrane domain) and OmpX, suggesting that SurA shows substrate specificity in solution, independently of co-chaperones Marx et al. (2020b). SurA-bound OMPs are in an extended conformation [Marx et al. (2020b), Chamachi et al. (2021)]. Several binding regions of SurA, on both OmpX and transmembrane domain of OmpA, were detected (Figures 4A,B, shown as orange bars); the importance of multiple conserved Tyr residues in the binding site has been suggested Marx et al. (2020b). It is noteworthy that the preferred contact sites do not occur at the site of the  $\beta$ -signal. This suggests that this fragment, important for the recognition by the BAM complex is free, flexible and thus available for protein-protein interactions.

SurA binds the unfolded OMPs with a dissociation constant ( $K_d$ ) of a few hundred nanomolar [Bitto and McKay (2003), Bitto and McKay (2004), Wu et al. (2011), Chamachi et al. (2021)]. This overall strong binding of SurA and the client is achieved through many weak and not very site-specific interactions He et al. (2020). A recently proposed model Marx et al. (2020b) is that multiple SurA can bind to the transmembrane domain of OmpA, keeping it in an extended, insertion-competent state (Figure 4D). Such higher order complex formation may depend on the length of the client protein, as also observed for the small TIM chaperones [Figure 3A; Weinhäupl et al. (2018)]. FhuA, which forms a 22 strand  $\beta$  barrel showed a close to 1:2 stoichiometry in complex with SurA Thoma et al. (2015). Effective transfer of the client to the BAM complex requires the BAM interacting region of the OMP to be free and flexible [Wang et al. (2016), Marx et al. (2020b), Wang et al. (2021)]. Since the client samples multiple conformations during its interaction with the chaperone He et al. (2020), it can be expected that some of those client conformations have higher binding affinity towards the BAM complex and the dissociation from SurA and insertion into the membrane can continue in a similar way as seen for TIM chaperones.

It has been proposed that SurA may not hold the OMP client continuously, but rather bind and unbind rapidly and repeatedly. Marx et al. (2020a) proposed that the binding to the OMP may be faster than the collapse of the OMP to a molten-globule state, as the OMPs are not very hydrophobic. This “kinetic trapping” model awaits further experimental investigation.

### 5.2.2 Skp

The 17 kDa protein (Skp) chaperone represents a pathway for OMP transport across the bacterial periplasm that is parallel to SurA. The crystal structure of Skp shows a trimeric oligomeric state with a “jellyfish”-like architecture [Korndörfer et al. (2004), Walton and Sousa (2004)]. Within the trimer, a nine-stranded  $\beta$ -barrel is formed to which each monomer contributes three  $\beta$ -strands (the trimerization interface) and a long,  $\alpha$ -helical “tentacle,” made of two  $\alpha$ -helices in a coiled-coil arrangement. These helices are highly dynamic in solution and sample a large



conformational space Holdbrook et al. (2017), thereby allowing a drastic increase of the cavity in an ATP-independent manner. Interestingly, a recent study proposes an Skp activation mechanism that involves a monomer to trimer transition induced by the unfolded client protein. wtSkp has been shown to exist in a monomer-trimer equilibrium [Mas et al. (2020), Sandlin et al. (2015)], the monomer being intrinsically disordered Mas et al. (2020). Fully monomeric Skp mutant proteins were unable to bind unfolded client proteins, whereas NMR Mas et al. (2020) and smFRET Pan et al. (2020) studies showed substrate-induced trimer formation for Skp. It seems that simultaneous contacts of the unfolded client protein with all three Skp subunits stabilizes the trimer by avidity. This coupled folding and oligomerization mechanism may ensure the tight regulation of Skp activity in the periplasm Mas et al. (2020).

Skp-OMP complexes have been extensively studied by NMR, and it represented the first such atom-specific study of a full-length client bound to a chaperone Burmann et al. (2013). NMR spectra acquired on trimeric Skp show only a single set of NMR resonances for the three subunits, demonstrating that on time average each subunit samples the same conformational space. Binding of client protein to the trimer induces a transition from a very flexible state Holdbrook et al. (2017) to a more rigid state of the long  $\alpha$ -helical tentacles Burmann et al. (2013). The decrease in dynamics within the helices forming the substrate-binding cavity is thought to keep the uOMP within the cavity. The three-fold symmetry of the Skp subunits remains intact in the Skp-OMP complexes, which shows that the complex must be in a dynamic equilibrium of multiple states on a sub-millisecond time scale; non-dynamic binding would break the symmetry. The interaction strength of Skp with client proteins has been determined experimentally, and lies in the nanomolar  $K_D$  range [Qu et al. (2007), Wu et al. (2011), Chamachi et al. (2021)]. Skp-client protein complexes were shown to have global lifetimes of several hours *in vitro* Burmann et al. (2013). Although it may at first sight appear counter-intuitive to have such a high affinity in such a dynamic complex, it is the avidity of the multiple interactions, each individually weak and thus short-lived, which allows for the high overall affinity.

NMR data, including paramagnetic NMR that probe distances from a paramagnetic tag, suggest that, despite its high dynamics, the client protein adopts a more compact state within the Skp cavity than expected for a urea-denatured protein. From such NMR data, an average radius of 21 Å for OmpX in the Skp cavity has been found, whereas in 8 M urea the corresponding radius is more than two-fold larger Burmann et al. (2013). The authors proposed the term “fluid globule” to describe this very compact, yet highly dynamic nature of the client.

Recently, smFRET studies have revealed that Skp-bound uOMP had a lower energy transfer efficiency than aqueous OmpX Chamachi et al. (2021) or OmpC Pan et al. (2020), suggesting chaperone-induced extension of the aqueous, unfolded (but not denatured) OMPs. Pan et al. have estimated the radius of the OmpC-Skp complex to 39 Å Pan et al. (2020) which, assuming a radius of 6 Å for the Skp helices surrounding the binding cavity, seems compatible with the radius of the (shorter) OmpX client of 21 Å, deduced from NMR data

Burmann et al. (2013). Skp therefore binds its client protein in a highly dynamic state, which is more compact than the urea denatured protein but more extended than the collapsed, aqueous OMP at pM concentration. Using single molecule FRET spectroscopy, Chamachi et al. found intra-chain dynamics on the  $\mu$ s time-scale for Skp-bound OmpX Chamachi et al. (2021). As stated above, fast inter-converting client conformations (fuzziness of the client) are providing multiple protein-protein interaction sites. These are required for the formation of the holdase-competent Skp trimer and may also allow the ATP-independent transfer to downstream factors, such as BamA or the degradase DegP.

There is no specific motif in the transmembrane domain of OMPs that is known to be the recognition site of Skp. Instead, it is thought that the entire unfolded transmembrane domain is engaged in the interaction. By adapting the size of its binding cage or by recruiting more homotrimer chaperone complexes [Korndörfer et al. (2004), Schiffrin et al. (2016)] Skp is able to bind a broad range of clients of different sizes, including OMPs and periplasmic proteins Jarchow et al. (2008). This ability of clients with different lengths to recruit more or less chaperones is reminiscent of the case of the mitochondrial TIM chaperones or SurA (see above). Positively charged Skp tentacles (Figure 4E) bind the entire transmembrane domain of the OmpA client (Figure 4B) through multiple electrostatic and also hydrophobic interactions, encapsulating it completely, while the periplasmic domain of OmpA is soluble, outside of the Skp binding cavity, and according to its NMR signature in a folded state that resembles the final native state of this domain [Qu et al. (2009), Walton et al. (2009)]. This ability of leaving soluble parts of the client outside the binding cleft/cavity was also found in the case of the mitochondrial TIM8.13/TIM9.10 interacting with the Tim23 client (see section on mitochondrial TIM chaperones, Figure 3F).

Interestingly, different OMPs show very similar behavior in Skp Burmann et al. (2013); furthermore, a given OMP shows a very similar random-coil behavior in different chaperones [Skp and SurA; Thoma et al. (2015)].

### 5.2.3 DegP

The stress-induced DegP belongs to the High Temperature Requirement A (HtrA) protein family in the bacterial periplasm, where it is important for quality control of outer-membrane proteins. It has an established function in the degradation of proteins via its serine protease activity; it has also been shown to exhibit chaperone properties [Spiess et al. (1999), Clausen et al. (2002), Jiang et al. (2008), Subrini and Betton (2009), Clausen et al. (2011), Sawa et al. (2011)]. The protease function may be the more important one. Binding of unfolded OMPs to SurA or Skp is at the rate 1000-fold higher than binding to the DegP. It has been proposed that this difference in kinetics favors OMP binding to the former two chaperones, thus preventing degradation Wu et al. (2011).

DegP is found to exist in an inactive hexameric form (presumed to be a resting state) in which the so-called LA loop of the PDZ1 domain interacts with the active site L1 and L2 loops from a neighbouring subunit, thereby blocking the

access to the catalytic side Krojer et al. (2002). The inactive hexameric form of DegP is converted into 12- and 24-mers upon interaction with client proteins [Krojer et al. (2008b), Jiang et al. (2008)]. This transition requires binding of unfolded substrate protein, where the C-terminus binds to the PDZ1 domain while the cleavage site is presented to the protease domain Krojer et al. (2008a). It has been shown that simultaneous binding of covalently linked PDZ1-binding and cleavage-site degrons is required for efficient formation of the active, dodecameric protease complex Kim et al. (2011). The crystal- and cryo-EM structures of the 24-mer of DegP have been obtained from the proteolytically inactive DegP<sub>S210A</sub> mutant that lacks the catalytic Ser; the protease-inactive 12- and 24-mer DegP variants were purified in presence of substrate protein Krojer et al. (2008b). The crystal structure of the 1.13 MDa large 24-mer of DegP Krojer et al. (2008b) shows the formation of a large cavity interior of the 24-mer forming wide pores (up to 35Å) and 24 proteolytic sites that could be accessed only by the encapsulated client protein. It is interesting to note that oligomerization, in this case to 12- or 24-mers, plays an important role for activating the chaperone; dynamic oligomerization has been found already in Skp and TIM chaperones (see above).

In the context of this review, it is particularly interesting to consider the state of the encapsulated protein and the way it may interact with the chaperone. For the DegP case, the information available about the state of the encapsulated client protein and the binding motif(s) is somewhat indirect, as there are no NMR or smFRET studies available to date. Intriguingly, Krojer et al. (2008b) have proposed that the OMPs may be present in a folded state inside the cavity. This proposition comes from several observations. The co-purified substrates with wild-type DegP were stable over tens of minutes; as unfolded model proteins are digested over this time scale, their interpretation is that the OMPs may be folded. Furthermore, a characteristic shift in SDS-PAGE mobility, often used as a signature of folded OMPs, suggested that at least 50% of the encapsulated OMP clients had some residual tertiary structure. In the same study, cryo-EM reconstitution of co-purified OMP and protease-inactive DegP 12-mer variant showed an electron density in the center of the cavity, which was interpreted as belonging to a folded OMP encapsulated. However, due to low resolution of the electron density map (28 Å), it is difficult to make definite statements on the folding state, or to characterize the interactions formed between chaperone and client.

#### 5.2.4 DegQ

A study of the related HtrA protein DegQ, another degradase of the bacterial periplasm with dual protease and chaperone functions, provides interesting further insight into this family. Like DegP, DegQ exists in a resting hexamer state and can form 12- and 24-mers upon substrate interaction. DegQ does not play a role in OMP biogenesis but targets rather soluble proteins. Malet et al. used chemically unfolded and reduced lysozyme, or a short peptide that binds to the PDZ1 domain, to trigger the oligomerization to the 12-mer or 24-mer state Malet et al. (2012). The cryo-EM structure (at ca. 12–14 Å resolution) and mass-spectrometry analysis of a protease-

deficient DegQ mutant in its 12-mer state shows that it harbors simultaneously five to six lysozyme molecules. At this resolution it remains difficult to make definite statements about the state of the client or its interaction mode with the chaperone; nonetheless, the identification of rather well-defined lobes is compatible with a lysozyme molecule that is close to its native conformation Malet et al. (2012). Six lysozyme molecules could be fitted into the electron density without clashes. Tryptophan fluorescence measurements suggested a folded state, although a slight shift relative to the Trp spectrum of isolated folded lysozyme suggested that the fold may be altered.

By comparison of the lysozyme-filled DegQ with DegQ that was triggered to form the 12-mer with the short peptide, the authors could localize the lysozyme molecules, and thus infer information about the interacting parts of the chaperone Malet et al. (2012). Several loops have been found in the vicinity of the density that was ascribed to lysozyme [see **Figure 1D** of ref. Malet et al. (2012)]; these regions (a helix corresponding to residues 251–257, and loop residues 408–413, 31–33 and 58–62) contain predominantly methyl-bearing residues and only few charges (a single Lys is oriented towards the lysozyme).

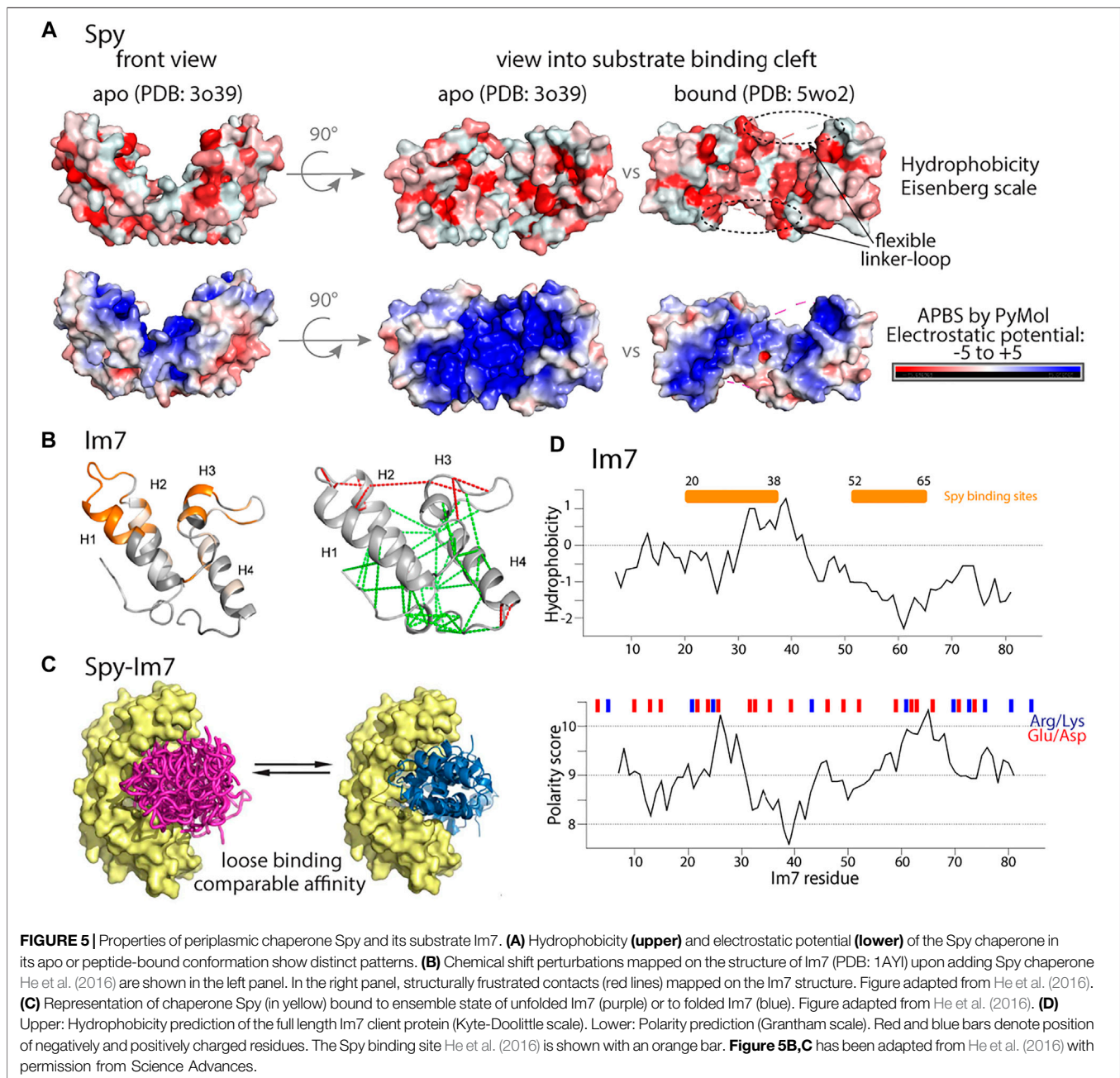
To summarize, both DegP and DegQ seem to support folded clients within their cavities, although the low resolution of the available data hampers precise statements about the interaction modes.

#### 5.2.5 Spy

The ATP-independent periplasmic chaperone Spy (spheroplast protein  $\gamma$ ) can bind to both native (folded) and non-native proteins, but with higher affinity for the latter. Ability to bind both unfolded and folded clients enables client folding while bound to the chaperone Mitra et al. (2021). It can be imagined that this ability could disrupt normal cell function by interacting with folded proteins. However, in the cell, levels of Spy are well controlled and Spy is only up-regulated in the stress conditions induced by protein aggregation Quan et al. (2011).

This stress-induced dimeric chaperone forms a cradle-like structure through an anti-parallel coiled-coil interaction of two 16 kDa monomers [Quan et al. (2011), PDB ID: 3o39]. Spy binds unfolded periplasmic or outer-membrane proteins with its cradle-shaped binding site, formed mostly by positively charged residues, and allows for their full folding while they are bound Stull et al. (2016). The structural and dynamical properties of Spy with a client, the small helical protein Im7, have been investigated independently by several groups, using NMR spectroscopy, MD simulations, crystallography and other biophysical methods [Salmon et al. (2016), Horowitz et al. (2016), He et al. (2016)]. Im7 is an interesting case because in the absence of its cognate binding partner (colicin E7), several residues at its binding interface are in energetic conflicts, i.e., they are restrained such that they cannot engage in the energetically favorable interactions with other residues, a situation termed local frustration Ferreiro et al. (2014). The Spy-Im7 complex is an instructive example for understanding chaperone-client interactions. It is also an interesting case that highlights the difficulty of studying such complexes by crystallography: the





high flexibility of the client and accordingly low electron density renders interpretation of the data difficult, possibly leading to mis-interpretation Wang (2018), which might have challenged one such study Horowitz et al. (2016). This controversy regarding Spy-Im7 crystal structures has been addressed (Rocchio et al., 2019), where the low occupied conformational ensembles of Im7 were reconstituted using selective anomalous labeling and residual electron and anomalous density (READ) method.

When Spy binds client proteins (as tested with Im7), it does not undergo large structural alterations. The main change is an increased flexibility of the linker loops; this increased loop

flexibility may facilitate the interaction with different substrate conformations Salmon et al. (2016). Remarkably, it was found that a more flexible mutant of Spy, with mutations in client-binding cradle shown up to seven-fold higher chaperone activity Quan et al. (2014), suggesting that increased flexibility is important for tight client binding. Furthermore, the polar and charged surface of Spy's binding cradle is changed (Figure 5A), forcing the conformational changes of the client while the client is still bound Horowitz et al. (2016).

In the complex, the client interacts primarily through a part that forms one face of the folded state (residues ca. 20–38 and 52–65), and the remainder of Im7 adopts essentially a native-

like conformation in the complex He et al. (2016), as judged by NMR chemical-shift perturbations (CSPs) (Figure 5B left panel). The client protein was found to be destabilized, i.e., more dynamics, as evidenced by hydrogen/deuterium exchange. Spy spatially compacts the conformational ensemble of its substrates. The binding is entropy (rather than enthalpy) driven, presumably due to the release of water molecules that are ordered in the apo state He et al. (2016).

He et al. have also prepared a triple mutant of Im7 which is unable to fold, and observed how this unfolded client interacts with Spy He et al. (2016). The mutant binds to Spy with  $K_d = 0.3 \mu\text{M}$  affinity (the wt Im7 has a somewhat higher  $K_d = 2 \mu\text{M}$ ; Figure 5C). The Spy-bound Im7 mutant shows the hallmark features of a random-coil polypeptide, and the interacting part of this Im7 includes the residues that are most involved in complex formation in the wild-type Im7, but also residues beyond this stretch, i.e., the binding sequence on the unfolded Im7 mutant is less well defined than on the wild-type Im7.

What drives the interaction between Im7 and its chaperone? When inspecting the biophysical properties of the residues of Im7 that are involved most in the interaction (20–38; 52–65) no clear relationship with the hydrophobic character or the presence of charges appears (Figure 5D). The interaction between the chaperone and the clients involves both hydrophilic and hydrophobic residues. Interestingly, the sites of Im7 that interact most strongly with Spy (residues ca. 20–38 and 52–65) correspond well to the sites for which the local frustration is highest in the native state (Figure 5B). Similar observations have been made with a different client, SH3, bound to Spy He and Hiller (2019). Thus, simple sequence properties do not seem to be the driving force for interactions, and it is rather the properties of the folded state, in particular the local frustration that plays an important role for the interaction of wild-type Im7 and its chaperone. Of note, the client protein in this case, wild-type Im7, appears to be already close to its native fold when it interacts with Spy. For the Im7 mutant that is unable to fold, the argument on structurally frustrated sites does not hold; in this case, a combination of hydrophobic and charged residues appears to be required for binding.

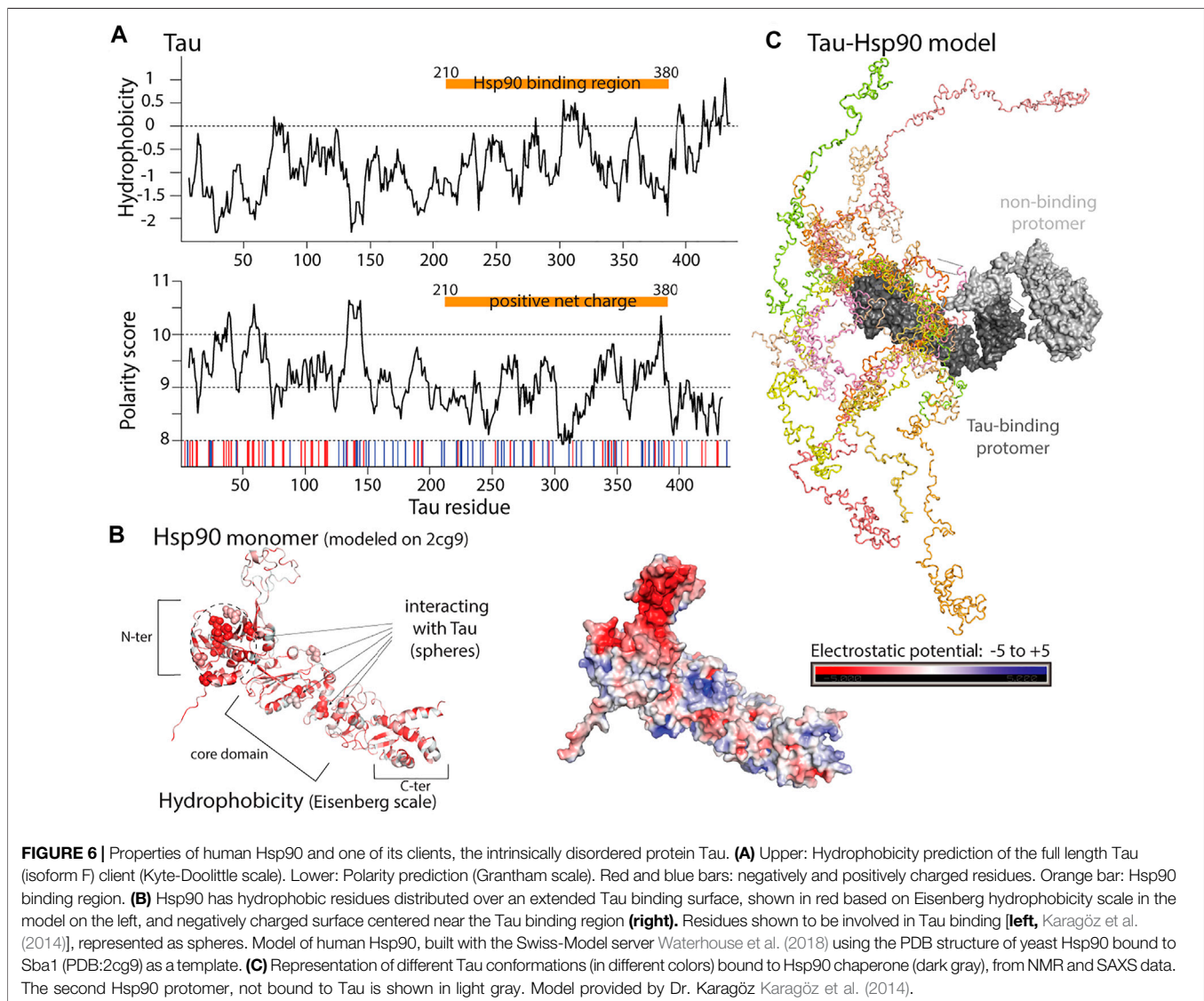
Do these client proteins behave differently when bound to different chaperones? This question was investigated by binding Im7 or SH3 to Skp or SurA He and Hiller (2018). Interestingly, a very similar pattern of CSPs is observed, both for the predominantly folded wild-type Im7 and the unfolded triple mutant, as well as for SH3, showing that the interaction with chaperones is rather independent of the exact details of the chaperone. One may argue that these three chaperones are similar to some extent, an amphipathic binding surface of both hydrophobic and polar residues. Further investigations with chaperones of more divergent properties (e.g., a highly hydrophobic surfaces) would allow clarifying how binding depends on polar/charge/hydrophobic properties.

### 5.3 Hsp90 Interaction With an Intrinsically Disordered Client

The ATP-dependent 90 kDa heat-shock protein family (Hsp90) has been extensively studied [see e.g., reviews by Taipale et al. (2010), Biebl and Buchner (2019)], and represents in itself a field that is far too vast to grasp in this review. The Hsp90 chaperones act in the late stage of protein folding (after Hsp70) and take care of a diverse substrate pool, including intrinsically disordered proteins (IDPs). We focus here only on one insightful complex of cytosolic Hsp90 with the IDP Tau, which has been investigated by NMR and small-angle X-ray scattering (SAXS) methods Karagöz et al. (2014) and EPR methods Weickert et al. (2020). Even though this review focuses on ATP-independent chaperones, we find it insightful to discuss this particular complex, formed by Hsp90 and Tau in the absence of ATP, where the ATP-independent holdase function is of main interest. Hsp90 aids Tau's association with microtubules or its degradation, and in that way plays a protective role against Tau's aggregation which is present in certain neurodegenerative diseases Medeiros et al. (2011).

Tau binds Hsp90 with a dissociation constant in the low micromolar range through its microtubule-binding part, including the aggregation-prone repeat region Karagöz et al. (2014). Although the authors have not determined the life time explicitly, the Tau-concentration-dependent NMR and fluorescence signal shows a behavior characteristic of the so-called fast exchange regime. This observation points to on/off rate constants in the  $\mu\text{s}$ -ms regime, i.e., the Hsp90-Tau complex forms transiently (as opposed to e.g. the complexes of small TIM chaperones with membrane-protein precursors, section 5.1).

The parts of Tau that appear to be the most important motifs (as seen from NMR data) contain hydrophobic residues (amino acid types Leu, Ile, Val, Phe and Tyr), and these sequences have a positive net charge, comprising several Arg and Lys residues (Figure 6A). This part of Tau binds to an extended surface of the chaperone (Figure 6B). The interaction surface on the chaperone has mixed properties: it comprises hydrophobic residues, but they are scattered, rather than forming a continuous hydrophobic patch; the authors claim that this scattered nature may ensure that it can make a large number of low-affinity contacts, and that it may also prevent Hsp90 from self-aggregation. The domains that are involved in the binding (N-ter and middle domains) have an overall negative charge, which shall complement the positive net charge of its client, but the binding site itself has a mixed positive/negative potential. Thus, it appears that the Hsp90/Tau interaction is based on a mixture of rather scattered hydrophobic and charge interactions. Hsp90-bound Tau is in an extended, unfolded and dynamic ensemble (Figure 6C). Electron paramagnetic resonance data show that Tau in isolation is also very dynamic, but has a tendency to fold back on itself ("paper-clip like"); when bound to Hsp90, Tau is extended further than in isolation, and the Hsp90-Tau interaction exposes Tau to oligomerization involving the two C-terminal pseudorepeats Weickert et al. (2020).



## 5.4 The Cytosolic Chaperone SecB

In Gram-negative bacteria, secretory proteins are synthesized on the cytoplasmic ribosome and targeted for post-translational translocation through the inner membrane. Nascent polypeptide chains of these proteins are recognized and transported to the Sec machinery by certain cytosolic chaperones [reviews on Sec machinery by Lycklama a Nijeholt and Driessen (2012), Chatzi et al. (2013), Tsirigotaki et al. (2017)]. The current model is that the ATP-independent SecB chaperone carrying an unfolded client protein binds to SecA in the cytosol and that this ternary complex interacts with the membrane-bound Sec machinery, after which the SecB dissociates, i.e. the client protein is handed over from SecB to the Sec complex [Hartl et al. (1990), Suo et al. (2015)]. [SecB also has other roles reviewed by Sala et al. (2014)].

An early model of client-protein recognition by SecB chaperone based on peptide-binding assays and stopped-flow fluorimetric experiments suggests that the primary interaction is

of ionic nature; initial binding could cause a conformational change in SecB, exposing its hydrophobic areas which would then further strengthen the interactions with the substrate protein (Randall, 1992; Stenberg and Fersht, 1997). It was shown that SecB exhibits a preference for unstructured stretches of polypeptides, that contain both basic and aromatic residues (Rüdiger et al., 1997; Knoblauch et al., 1999) and that it binds its clients in a 1:1 ratio [1 SecB tetramer per 1 unfolded client; Lecker et al. (1989), Hardy and Randall (1991), Stenberg and Fersht (1997)] with a 30  $\mu\text{M}$  affinity [obtained by calorimetric titration of SecB into a solution of maltose-binding protein (MBP) at 7°C Randall et al. (1998)]. Huang et al. (2016) reported affinities for binding of unfolded MBP and phosphatase A (PhoA), which are of the order of 0.05  $\mu\text{M}$  (MBP) to 0.5  $\mu\text{M}$  (PhoA); for shorter fragments of the clients (of the order of 30–80 residues) the  $K_d$  are in the 1–70  $\mu\text{M}$  range.

SecB forms homo-tetramers (dimer of dimers) of ~68 kDa [PDB: 1qyn, Dekker et al. (2003)]. Huang et al. (2016) have used



two different client proteins, the periplasmic maltose binding protein (MBP, 396 amino acids) and alkaline phosphatase A (PhoA, 471 amino acids) as client proteins, and investigated the formed complexes by NMR spectroscopy. MBP and PhoA were chemically unfolded (urea), and while they fold upon removal of urea in the absence of chaperone, SecB keeps them in an unfolded state.

NMR spectroscopy experiments showed that both MBP and PhoA remain in an unfolded conformation upon binding SecB Huang et al. (2016): NMR spectra of the SecB-bound clients strongly resemble those of the urea-denatured unfolded state. The authors performed a structure calculation of the complex based on distance measurements between the chaperone and client [nuclear Overhauser effect (NOE) and paramagnetic relaxation enhancement data]. Interestingly, the authors performed a structure calculation akin to what one generally does when determining the structure of a folded protein, i.e. attempting to determine a single state. Conceptually, their approach of structure determination implies that the energy landscape has a well-defined minimum (cf. **Figure 2**), which contrasts e.g. the cases of TIM and Skp, where an explicit ensemble is the only way to realistically represent the complex. Whether this implicit single-structure assumption is justified is not within the scope of this review. Based on the resulting structural models, the authors draw the following conclusion on the complex. Distinct parts of the client proteins bind to grooves on the four subunits as well as to a secondary binding site. A slight structural rearrangement of SecB is reported, involving rotation of a helix, which increases the hydrophobic surface; rearrangement of side chains of SecB increases the space available to bulky hydrophobic side chains of the client. The interaction surface is maximized by the client protein being wrapped around the chaperone; the authors state that the simultaneous binding of the multiple sites enhances affinity, but that the binding synergy is not strong.

The interacting regions on the clients are enriched in hydrophobic and aromatic residues; while hydrophobic contacts appear to be the driver of the interaction, several hydrogen bonds line the hydrophobic groove, according to the structure that was determined. Simultaneous substitution of three hydrophobic amino-acids of the SecB binding site were shown to be sufficient for defective binding of the MBP Huang et al. (2016). This observation is remarkably similar to the disruption of client-protein binding of TIM9.10 which can be caused by a single mutation in each Tim subunit Weinhäupl et al. (2018).

In conclusion, the SecB chaperone appears to employ primarily hydrophobic contacts for client binding. As an interesting difference to the other examples, distinct NOE signals are detected and interpreted in the framework of a well-defined binding pose for each of the binding fragments.

## 5.5 The Cytosolic Chaperone Trigger Factor

Trigger factor (TF), an ubiquitous chaperone that forms dimers in the cytosol and binds to the ribosome as a monomer, commonly functions in facilitating co-translational folding of cytoplasmic proteins or in handing them over to downstream foldases for post-translational folding. The structure and function

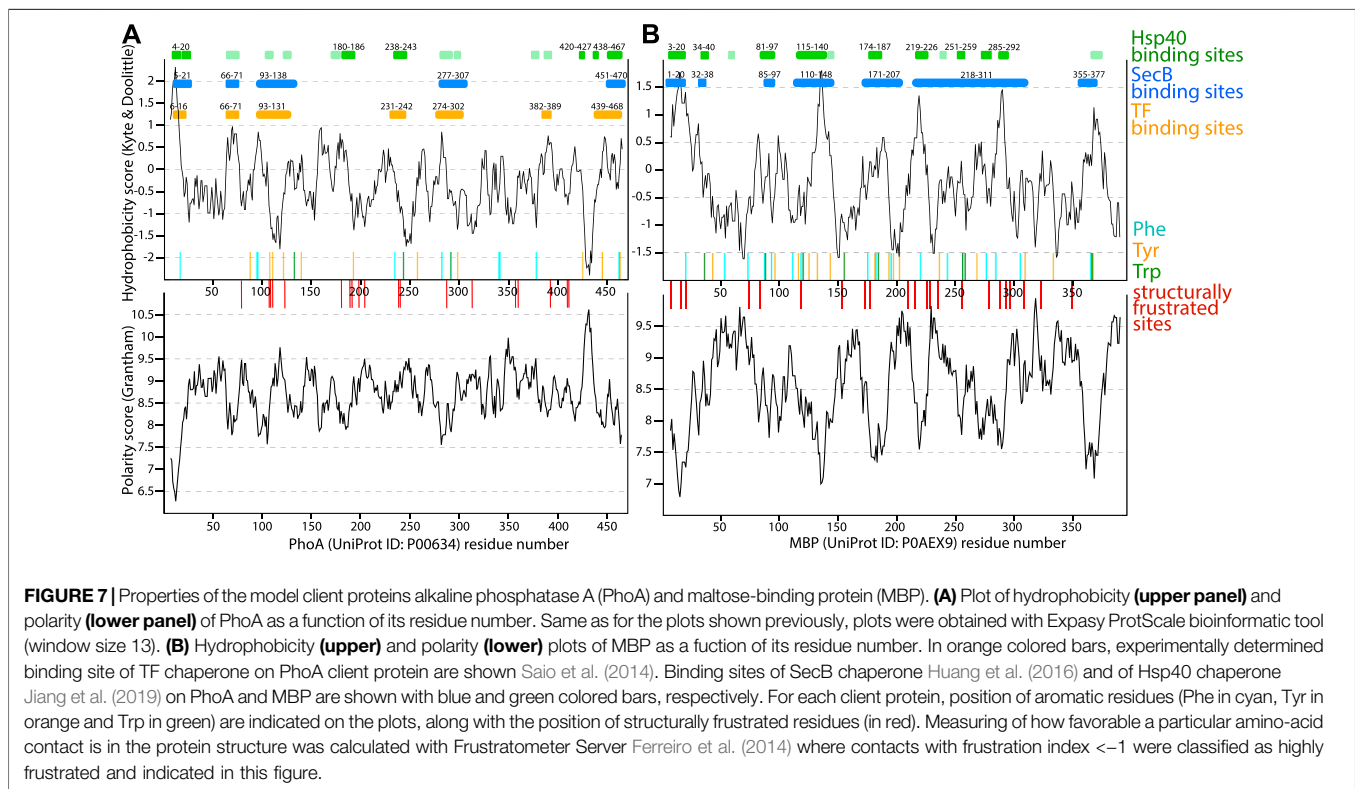
of TF are reviewed, e.g., by Hoffmann et al. (2010). For the purposes of this review we focus on the holdase function of TF and in particular on one complex, the one with model client proteins, PhoA, which was also discussed as a client in SecB studies, and the aggregation-prone G32D/I33P variant of MBP Saio et al. (2014). The complex formation with PhoA was achieved under reducing conditions where PhoA is unfolded, as PhoA requires oxidizing conditions to fold. The affinity is in the low  $\mu\text{M}$  range ( $K_d$ ), whereby increasing the length of client (full-length PhoA or fragments) leads to enhanced affinity, which points to some degree of binding synergy (avidity). Due to its large size (471 amino acids), PhoA recruits 3 TF monomers Saio et al. (2014). Similar characteristics as for SecB emerged from NMR experiments on TF-bound PhoA and complexes of TF with PhoA fragments: PhoA shows the hallmark features of an unfolded protein when bound to TF, in particular NMR spectra and NMR relaxation parameters characteristic of disordered proteins. It uses only ca. one third of its sequence to bind. A combination of NMR experiments with PhoA fragments allowed structure calculation, which was again performed with methods that aim for a single structure, as discussed in the SecB section above.

The important interactions identified in the resulting complex structures are predominantly hydrophobic in nature, with aromatic residues being mostly involved. A single amino-acid substitution at the hydrophobic substrate-binding sites in TF resulted in a significant decrease in the affinity for PhoA Saio et al. (2014), similarly to TIM9.10 or SecB.

## 5.6 J-Domain Proteins (Hsp40) and Client Specificity

J-domain proteins (JDP) are a component of the important Hsp70-Hsp40/NEF (nucleotide-exchange factor) system. They are often referred to as 40 kDa heat-shock proteins (Hsp40), arguably a misleading name as most JDPs have a mass rather different from 40 kDa. In this tripartite chaperoning system, Hsp70 is the ATP-driven holdase/foldase which binds to a very broad range of substrates. Hsp70 is very promiscuous: the binding sequence has a hydrophobic core (4–5 residues) flanked by basic residues. In the *E. coli* proteome, Hsp70-binding motifs occur statistically every 36 residues Rüdiger et al. (1997). A number of crystal structures of Hsp70 in complex with peptides have revealed a well-defined binding pocket [Mayer and Gierasch (2019), Mayer (2021)]. NEF assists the release of ADP from Hsp70 after ATP hydrolysis Kampinga and Craig (2010). The JDPs have several roles: 1) they stimulate the ATP-hydrolysis of Hsp70, which in turn leads to tighter binding of Hsp70 to its client, and 2) they bind to clients and hand them to Hsp70. The latter generally has higher client affinity; for example, a fragment of a mitochondrial outer-membrane protein binds to yeast Hsp70 about ten-fold stronger than it does to the JDP Ydj1, suggesting the directional transport, from JDP to Hsp70 Jores et al. (2018).

The ability of a given Hsp70 (highly promiscuous) to interact with different JDP co-chaperones brings the outstanding versatility to the Hsp70 system, engaged in a myriad of



cellular processes [Kampinga and Craig (2010), Craig and Marszalek (2017), Barriot et al. (2020), Mayer (2021)]. Consequently, most species have many more JDP genes than Hsp70 genes; e.g., the cytoplasm of *Saccharomyces cerevisiae* contains 13 different JDPs (Apj1, Xdj1, Ydj1, Caj1, Djp1, Hlj1, Sis1, Cwc23, Jjj1, Jjj2, Jjj3, Swa2, and Zuo1) but only two classes of Hsp70s, Ssa (SSA1–4) and Ssb (SSB1–2) (Sahi and Craig, 2007). JDPs by themselves can suppress protein aggregation, and in some cases they function in the cell without the involvement of Hsp70 [Kampinga and Craig (2010), Craig and Marszalek (2017)].

Common to all JDPs is the presence of a J-domain, a ca. 80-residue helix-bundle domain; besides this defining common feature, there is a wide variety of additional domains. The most ubiquitous classes of JDPs (type I and II) comprise  $\beta$ -barrel C-terminal domains I and II (CTD-I, CTD-II) and a dimerization domain at the extreme C-terminus; thus, these JDPs act as dimers, which increases the number of interaction sites to clients (as compared to monomers). Type I JDPs carry an additional zinc-finger domain Kampinga and Craig (2010). The JDPs that do not fulfill these definitions are called type III, with a large variety of domains. Some JDPs, especially of type III, are highly client specific, and bind only a single protein [Fotin et al. (2004), Vickery and Cupp-Vickery (2007)]. To do so, however, they use a dedicated domain for recognition; however, most type I and type II JDPs have a larger clientome. Nonetheless, specific involvement of certain JDP chaperones in binding different clients is common. For example, How do JDPs recognize their clients? Interesting

insight comes from an NMR study of several JDPs (type I and type II) with the model client proteins PhoA and MBP, which were introduced in the SecB and TF sections above Jiang et al. (2019). Akin to their behavior on SecB and TF, the two client proteins bind the JDP chaperones in a largely unfolded dynamic conformation devoid of secondary structure. Making use of the changes observed for NMR signals of the clients, the authors identified regions of PhoA and MBP which bind to the JDP. These are enriched in hydrophobic residues and comprise at least one aromatic residue; a negatively charged residue preceding the aromatic residue increases the affinity. These distinct stretches are shown in **Figures 7A,B**. The native-state secondary structure of these interacting stretches varies, suggesting that the final secondary structure is unimportant for binding.

The binding sites on the JDPs are located to the C-terminal domains I and II, within hydrophobic grooves in these domains; the sequence preferences of these two grooves slightly differ. However, not all the tested JDPs use both CTDs, and the authors propose how client specificity in the JDP family may arise, namely through the fact that not all JDPs use both CTDs, and the different preferences of the two CTDs within a JDP.

Structure calculation in that study involved a complex protocol based on calculations with small fragments; it is challenged by the inherent dynamics of the complex and the difficulty to identify with certainty which part of the client binds to which part of the JDP (Jiang et al., 2019). From NMR relaxation-dispersion (RD) data and bio-layer interferometry (BLI) data the authors state that the life time of the complex is in the millisecond range. To us, this estimation is not entirely

clear, as their BLI data rather point to slower on/off kinetics [seconds; Figure S10 of Jiang et al. (2019)], and the RD data were obtained only for a small fragment of PhoA, not representative of the entire client.

Other studies reported that when a given JDP engages with different client proteins, it may use different parts to do so. For example, both DNAJB6 and DNAJB8 bind stretches of glutamine (polyQ) in an amyloid-forming polypeptide (GAMKSFQ<sub>45</sub>F) that is largely disordered with a tendency to form  $\beta$ -turns and aggregate. For the interaction with these polyQ stretches, a Ser/Thr-rich region of the JDP is essential Kakkar et al. (2016b). The same JDPs also bind to another target, a mutant of the E3 ubiquitin ligase Parkin, a protein comprising globular folded ubiquitin-like and RING domains in its native state. Interestingly, for the interaction with this Parkin mutant, the Ser/Thr-rich region of the JDP is dispensable Kakkar et al. (2016a). Thus, this JDP has different modes of client interaction. Of note, the two described clients are structurally different, one is intrinsically disordered and the other one folds to a well-defined 3D structure, which may be responsible for recruiting different parts of the JDP.

## 6 CONCLUSION: EMERGING PATTERNS OF CLIENT-CHAPERONE INTERACTIONS

The function of chaperone proteins and the mechanisms by which they recognize, bind and fold their client proteins has been of great interest to structural biologists for more than 3 decades [Ellis (1990), Neupert et al. (1990), Hartl et al. (1992), Kelley and Georgopoulos (1992), Jakob and Buchner (1994), Horowitz (1998), Feldman and Frydman (2000)]. Crystallography has been crucial to obtain structures of the apo states but it has so far turned out to be very limited when it comes to characterizing structures of chaperones with full-length client proteins, due to the heterogeneity and dynamics. Only recently, and by combined efforts from several techniques (NMR, SAXS, SANS, FRET, MD, often integrated), we are getting atomic-level insights into the interactions that underlie chaperone–client interaction.

It is often assumed that the mechanisms by which chaperones hold polypeptides is by binding to hydrophobic stretches, thereby protecting them from self-aggregation in the aqueous environment. The selection of chaperone complexes that we have reviewed here shows that chaperone–client interaction is more complex than being limited to hydrophobic interactions, and we summarize here the emerging view.

The kind of interactions that is important for the complex formation depends on the nature of the client protein and the degree at which it is folded when encountering the chaperone. Among the clients we have encountered here were proteins which are unable to fold (in the environment where they bind the chaperone), while others can fold, or are even already close to the folded state. Along this continuum from disordered to folded proteins the way how chaperones and client proteins interact with each other necessarily differs. For proteins that are close to their native state already, the interaction sites are

essentially those sites that are least stable. Structurally frustrated sites generally correspond to these least stable parts of proteins (Ferreiro et al., 2014), therefore there is a good correspondence between the sites with highest structural frustration and the chaperone-binding sites. This was found for Im7 and SH3, two proteins which are known to populate partially unfolded intermediate folding states in solution, and which are close to their native state when bound to Spy (see [section 5.2.5](#)). The exact nature of the amino acids or the structural motif does not appear to be the determining criterion in such cases; in the Im7 and SH3 cases it is a mix of charged, polar and hydrophobic residues which binds the chaperone.

On the opposite extreme of this spectrum are proteins which are unable to fold, and which comprise parts that need to be shielded from the solvent to prevent aggregation. The mitochondrial membrane proteins are such a case ([section 5.1](#)), which bind to hydrophobic groove of the chaperone with their hydrophobic transmembrane parts, retaining only a small helical tendency in an otherwise elongated conformation. Reducing the hydrophobic nature of the chaperone by a single hydrophobic-to-charged mutation can totally abrogate the ability to hold the client. However, even for such unfolded clients, not only hydrophobic contacts are important, as highlighted by the cases of the complex formed by Hsp90 and the intrinsically disordered (and aggregation-prone) Tau ([section 5.3](#)), or the SurA-Omp and Skp-Omp cases ([section 5.2.1](#) and [section 5.2.2](#)); these chaperones present a binding surface that comprises many charged residues, and electrostatic interactions contribute to the complex formation. In the case of the (unfolded) clients PhoA and MBP binding to SecB, TF and JDPs, the frustrated sites do not appear to be particularly involved in binding ([Figure 7](#)); arguably, this can be expected, as the clients are totally unfolded, and thus their primary structure is important, but not the structure of the folded state.

The different types of interactions (hydrophobic, polar, electrostatic) may also contribute at different points along the complex-formation process: for the Skp chaperone, for example, initial binding is most likely driven by electrostatic interactions and the client is then encapsulated additionally by the hydrophobic interactions [Qu et al. (2007), Qu et al. (2009)]. Similarly, rapid initial binding of the Im7 substrate by the Spy chaperone is thought to be achieved through the electrostatic (not hydrophobic) interactions, and the complex is further strengthened by the hydrophobic interactions in the binding site Koldewey et al. (2016). Hereby, locally frustrated, and thus inherently unstable/unfolded parts of the partly folded client are particular hotspots of interaction with the chaperone He and Hiller (2019). The self-folding of the client on the chaperone surface triggers client release, and thereby the chaperone-client interaction is terminated without the need for any particular trigger event Koldewey et al. (2016).

ATP-driven chaperone machineries exploit this combination of interaction types in a more active manner: Hsp60 chaperonins “capture” their clients through hydrophobic interactions close to the chamber entry, and then, upon ATP hydrolysis and allosteric closure of the chamber including rotation of some helices, the

client protein finds itself in a much more hydrophilic environment (Yebeles et al., 2011; Lopez et al., 2015).

Client-specificity of a chaperone, and chaperone-specificity of a given client protein are important in the complex environment of the cell, where the right cellular localization of a protein is also related to the way it is transported to e.g., cellular compartments. The question about specificity comprises two questions: does a given client protein interact in similar ways with different chaperones? And does a given chaperone hold different client proteins in a similar manner? The examples discussed here show that there are clearly similarities when a client binds to different chaperones: for example, OMPs are similarly behaved (namely unfolded) in Skp and SurA; PhoA and MBP engage with similar stretches along their sequence when binding to SecB, TF and Hsp40; the Tim23 client wraps its hydrophobic transmembrane part around the clefts of TIM9.10 and TIM8.13 similarly. However, in detail the interactions differ, which opens the possibility for introducing specificity. We have discussed the case of Tim23-binding to TIM8.13 vs. TIM9.10, where only one of the two (TIM8.13) is able to engage in hydrophilic contacts with a 100-residue long stretch of the client, while the other one is better in binding hydrophobic stretches; accordingly, these two chaperones are able to diversify their clientome Sućec et al. (2020). The Hsp40 (JDP) chaperone system is another important case, and we discussed the example where a given Hsp40 uses a Ser/Thr rich part when interacting with some clients, but not with others (section 5.6). Along these lines, we described the NMR study that suggested that different JDPs use either only one or two of the CTDs to engage with a given client, thus also allowing for specificity.

In the context of specificity, it must be mentioned that the cellular localization is also of central importance. Chaperones are often positioned at strategic points, for example on the ribosome Craig and Marszalek (2017) or near the exit of translocation pores [e.g. in the two mitochondrial membranes Shiota et al. (2015), Craig (2018)]. Interactions of the chaperones with these machineries (ribosome, translocases) keeps them right at the location where they are required.

Furthermore, chaperones collaborate, and a given client protein is often handed over from one chaperone to the next. Such interaction networks can either involve functionally redundant chaperones or a step-wise substrate transfer within chaperone cascades. For example, bacterial OMPs are handled by different periplasmic chaperones. A holistic view on the periplasmic chaperone network was obtained from a mathematical model that integrated available experimental information from *in vivo* and *in vitro* studies. From these simulations, the authors concluded that functional robustness does not necessarily rely on the concept of specific pathways [Costello et al. (2016), Chum et al. (2019)]. On the other hand, within the GET pathway, tail-anchored membrane proteins (TA-MBPs) are transported to the ER membrane via a step-wise substrate transfer from highly promiscuous Hsp70 (Ssa1 in yeast) to the selective Get3 that traps TA-MBPs for membrane insertion. Such a cascade, engaging more specialized chaperones with increasing affinity allows for efficient, selective, and unidirectional targeting of nascent TAs, while protecting them

from reaction with other cytoplasmic chaperones [reviewed in Shan (2019)]. In a similar way, a network of chaperones including different Hsp40s, Sti, Hsp70 and Hsp90 is important for safeguarding mitochondrial precursor proteins across the cytoplasm Bykov et al. (2020).

A common theme found in many chaperone systems is their oligomeric nature, found for example in small TIMs (hexamer), Skp (trimer), Hsp40 (dimer), Spy (dimer) or Hsp60 (14- or 16-mer). The dimerization represents several advantages for a chaperone. Most importantly, it enables the chaperone to present a larger binding surface which is often also used as a cavity-like architecture. Through the multiple interaction sites with the client, the oligomeric chaperone strongly enhances its affinity to the client by avidity. When the subunits are allosterically coupled to each other, such an oligomeric machinery may perform even concerted large scale movements; positive intra-ring allostery and negative inter-ring allostery in chaperonins are an example of the functional complexity that can be achieved from relatively small building units Yebeles et al. (2011). The oligomeric nature often implies that subunits can go in and out in a dynamic manner. In small TIM chaperones Weinhäupl et al. (2021) and Skp Mas et al. (2020), the oligomeric chaperones are in equilibrium with monomeric subunits. Together with a protease system that clears exclusively the monomers Baker et al. (2012) and replenishment by newly synthesized subunits, the concentration of chaperones can thus be adjusted very efficiently. This may present a simple way to adjust the chaperone level to the state of the cell. The dynamic monomer-oligomer equilibrium also presents a direct way of regulating the chaperone activity. In the Skp system, the presence of client proteins shifts the equilibrium from the monomeric to the oligomeric state; thus, the presence of clients generates a higher effective chaperone concentration, and therefore the chaperone activity is very rapidly adjusted to the needs. Interestingly, in some chaperone systems, the inverse process occurs: the chaperone exists in a high-oligomer “resting state”, and the presence of client leads to a deoligomerization of the chaperone into smaller subunits and activation of its chaperone activity [Haslbeck et al., 1999; Jehle et al., 2010].

A central property of all the chaperones discussed here is their dynamic nature. These complexes are often held together by a multitude of individually weak interactions, which, due to their large number, can result in a strong overall affinity. Dynamics is important in chaperone-client complexes for several reasons. First, the multi-conformation dynamical ensemble results in a more favorable entropic contribution for binding than a single conformation would do. In contrast to complexes of folded proteins, which are rich in highly specific interactions (such as salt bridges) which make a large enthalpic contribution to binding, chaperone-client complexes often do not have such interactions. Thus, while the free energy of binding in rather rigid complexes of folded proteins is dominated by a strong enthalpic component (and an entropic penalty), this is not the case for chaperone-client complexes, and the entropic component shall be favorable (or at least less disfavored) than for complexes of folded proteins. Second, as discussed above, dynamics in chaperone complexes can lead to



high overall affinity (particularly important for holdases of highly aggregation-prone client proteins), while avoiding high energy barriers for release (important for efficient release).

The details of the dynamics of a given client bound to different chaperones, or of different clients bound to a given chaperone certainly differ. Such differences may provide a further layer by which the cell can differentiate client proteins. Technically, different dynamic behaviors of diverse clients means that their experimental characterization has to be adapted. For example, the NMR signature of a protein sliding in a chaperone's binding pocket depends on the time scale of this motion; if it falls into the millisecond range, NMR line broadening is induced, which challenges the extraction of information about the conformation and inter-molecular contacts. Due to the broad range of dynamics in these complexes, their characterization generally requires multiple techniques, and it will certainly continue to be a playground of integrated structural biology, where advanced computational methods, such as explicit ensembles derived from MD simulations are confronted with a multitude of experimental observables.

It must be stressed that we are only starting to decipher the chaperone function at the atomic level, and that it can be expected that there is much more diversity than what appears from the examples described here. The proteins that have been selected in these studies may well provide a biased view: the clients are often model proteins, or artificially denatured by urea or mutations. The complexes have also rather divergent life times and affinities: Skp or TIM chaperones capture their clients for many hours, while Spy-Im7 or Hsp90-Tau complexes have life times in the millisecond range. Whether this impressive factor of  $10^5$ – $10^6$  difference is related to the fact that the former rely much more on hydrophobic interactions, or whether the difference comes from the interaction surface area or the architecture of the chaperone is not clear and needs further investigation.

Although we have just scratched the surface of the molecular and structural features of these complexes, the features that these

examples have revealed will provide an important foundation as the community will explore more complex targets. Moving towards such more complicated and larger clients, and to higher-order complexes is certainly on the “to do list” for the field. It is becoming increasingly clear, for example, that many JDP clients are mature folded proteins, and that the JDPs remodel large multiprotein complexes, acting on the protein-protein interactions within these complexes. Studying such complexes will likely shed light onto new mechanisms. Another field of central importance is the one of membrane-insertion machineries. The hydrophobic membrane environment corresponds to physico-chemical properties very different from the aqueous solution in which most chaperones are studied currently. Continued technical improvements will be important for tackling these complex membrane-integrated/membrane-associated machineries.

## AUTHOR CONTRIBUTIONS

IS, BB, and PS wrote this manuscript. IS prepared the figures.

## FUNDING

This study was supported by the European Research Council (StG-2012-311318-ProtDyn2Function) and the Agence Nationale de la Recherche (ANR-18-CE92-0032-MitoMemProtImp).

## ACKNOWLEDGMENTS

We thank Juan C. Fontecilla-Camps for insightful discussions related to ATP-driven machineries, and Elif Karagöz for providing the structural model of the Hsp90-Tau complex.

## REFERENCES

- Arié, J.-P., Sassoon, N., and Betton, J.-M. (2001). Chaperone Function of FkpA, a Heat Shock Prolyl Isomerase, in the Periplasm of *Escherichia coli*. *Mol. Microbiol.* 39, 199–210. doi:10.1046/j.1365-2958.2001.02250.x
- Arranz, R., Martín-Benito, J., and Valpuesta, J. M. (2018). Structure and Function of the Cochaperone Prefoldin. *Adv. Exp. Med. Biol.* 1106, 119–131. doi:10.1007/978-3-030-00737-9\_9
- Bagshaw, C. R., and Trentham, D. R. (1973). The Reversibility of Adenosine Triphosphate Cleavage by Myosin. *Biochem. J.* 133, 323–328. doi:10.1042/bj1330323
- Baker, M. J., Mooga, V. P., Guiard, B., Langer, T., Ryan, M. T., and Stojanovski, D. (2012). Impaired Folding of the Mitochondrial Small TIM Chaperones Induces Clearance by the I-AAA Protease. *J. Mol. Biol.* 424, 227–239. doi:10.1016/j.jmb.2012.09.019
- Balchin, D., Hayer-Hartl, M., and Hartl, F. U. (2016). *In Vivo* aspects of Protein Folding and Quality Control. *Science* 353, aac4354. doi:10.1126/science.aac4354
- Barriot, R., Latour, J., Castanié-Cornet, M.-P., Fichant, G., and Genevaux, P. (2020). J-domain Proteins in Bacteria and Their Viruses. *J. Mol. Biol.* 432, 3771–3789. doi:10.1016/j.jmb.2020.04.014
- Bauer, M. F., Hofmann, S., Neupert, W., and Brunner, M. (2000). Protein Translocation into Mitochondria: the Role of TIM Complexes. *Trends Cel Biol.* 10, 25–31. doi:10.1016/s0962-8924(99)01684-0
- Becker, T., Song, J., and Pfanner, N. (2019). Versatility of Preprotein Transfer from the Cytosol to Mitochondria. *Trends Cel Biol.* 29, 534–548. doi:10.1016/j.tcb.2019.03.007
- Behrens, S., Maier, R., De Cock, H., Schmid, F. X., and Gross, C. A. (2001). The SurA Periplasmic PPIase Lacking its Parvulin Domains Functions *In Vivo* and Has Chaperone Activity. *EMBO J.* 20, 285–294. doi:10.1093/emboj/20.1.285
- Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002). *Biochemistry*. 5th edn. New York, NY: W. H. Freeman.
- Beverly, K. N., Sawaya, M. R., Schmid, E., and Koehler, C. M. (2008). The Tim8-Tim13 Complex Has Multiple Substrate Binding Sites and Binds Cooperatively to Tim23. *J. Mol. Biol.* 382, 1144–1156. doi:10.1016/j.jmb.2008.07.069
- Biebl, M. M., and Buchner, J. (2019). Structure, Function, and Regulation of the Hsp90 Machinery. *Cold Spring Harb. Perspect. Biol.* 11, a034017. doi:10.1101/cshperspect.a034017
- Bigotti, M. G., and Clarke, A. R. (2005). Cooperativity in the Thermosome. *J. Mol. Biol.* 348, 13–26. doi:10.1016/j.jmb.2005.01.066
- Bitto, E., and McKay, D. B. (2004). Binding of Phage-Display-Selected Peptides to the Periplasmic Chaperone Protein SurA Mimics Binding of Unfolded Outer Membrane Proteins. *FEBS Lett.* 568, 94–98. doi:10.1016/j.febslet.2004.05.014
- Bitto, E., and McKay, D. B. (2002). Crystallographic Structure of SurA, a Molecular Chaperone that Facilitates Folding of Outer Membrane Porins. *Structure* 10, 1489–1498. doi:10.1016/s0969-2126(02)00877-8



- Bitto, E., and McKay, D. B. (2003). The Periplasmic Molecular Chaperone Protein SurA Binds a Peptide Motif that Is Characteristic of Integral Outer Membrane Proteins. *J. Biol. Chem.* 278, 49316–49322. doi:10.1074/jbc.m308853200
- Burmam, B. M., Wang, C., and Hiller, S. (2013). Conformation and Dynamics of the Periplasmic Membrane-Protein-Chaperone Complexes OmpX-Skp and tOmpA-Skp. *Nat. Struct. Mol. Biol.* 20, 1265–1272. doi:10.1038/nsmb.2677
- Bykov, Y. S., Rapaport, D., Herrmann, J. M., and Schuldiner, M. (2020). Cytosolic Events in the Biogenesis of Mitochondrial Proteins. *Trends Biochem. Sci.* 45, 650–667. doi:10.1016/j.tibs.2020.04.001
- Calabrese, A. N., Schiffrin, B., Watson, M., Karamanos, T. K., Walko, M., Humes, J. R., et al. (2020). Inter-domain Dynamics in the Chaperone SurA and Multi-Site Binding to its Outer Membrane Protein Clients. *Nat. Commun.* 11, 2155–2216. doi:10.1038/s41467-020-15702-1
- Chamachi, N., Hartmann, A., Ma, M. Q., Krainer, G., and Schlierf, M. (2021). *Chaperones Skp and SurA Dynamically Expand Unfolded Outer Membrane Protein X and Synergistically Disassemble Oligomeric Aggregates*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Chatzi, K. E., Sardis, M. F., Karamanou, S., and Economou, A. (2013). Breaking on through to the Other Side: Protein export through the Bacterial Sec System. *Biochem. J.* 449, 25–37. doi:10.1042/bj20121227
- Chen, R., and Henning, U. (1996). Aperioplasmic Protein (Skp) of *Escherichia coli* Selectively Binds a Class of Outer Membrane Proteins. *Mol. Microbiol.* 19, 1287–1294. doi:10.1111/j.1365-2958.1996.tb02473.x
- Christiansen, A., Wang, Q., Cheung, M. S., and Wittung-Stafshede, P. (2013). Effects of Macromolecular Crowding Agents on Protein Folding *In Vitro* and *In Silico*. *Biophys. Rev.* 5, 137–145. doi:10.1007/s12551-013-0108-0
- Chum, A. P., Shoemaker, S. R., Fleming, P. J., and Fleming, K. G. (2019). Plasticity and Transient Binding Are Key Ingredients of the Periplasmic Chaperone Network. *Protein Sci.* 28, 1340–1349. doi:10.1002/pro.3641
- Clausen, T., Kaiser, M., Huber, R., and Ehrmann, M. (2011). HTRA Proteases: Regulated Proteolysis in Protein Quality Control. *Nat. Rev. Mol. Cell Biol.* 12, 152–162. doi:10.1038/nrm3065
- Clausen, T., Southern, C., and Ehrmann, M. (2002). The HtrA Family of Proteases. *Mol. Cell* 10, 443–455. doi:10.1016/s1097-2765(02)00658-5
- Costello, S. M., Plummer, A. M., Fleming, P. J., and Fleming, K. G. (2016). Dynamic Periplasmic Chaperone Reservoir Facilitates Biogenesis of Outer Membrane Proteins. *Proc. Natl. Acad. Sci. USA* 113, E4794–E4800. doi:10.1073/pnas.1601002113
- Craig, E. A. (2018). Hsp70 at the Membrane: Driving Protein Translocation. *BMC Biol.* 16, 11. doi:10.1186/s12915-017-0474-3
- Craig, E. A., and Marszalek, J. (2017). How Do J-Proteins Get Hsp70 to Do So Many Different Things? *Trends Biochem. Sci.* 42, 355–368. doi:10.1016/j.tibs.2017.02.007
- Cumberworth, A., Lamour, G., Babu, M. M., and Gspomer, J. (2013). Promiscuity as a Functional Trait: Intrinsically Disordered Regions as central Players of Interactomes. *Biochem. J.* 454, 361–369. doi:10.1042/bj20130545
- Curran, S. P., Leuenberger, D., Oppliger, W., and Koehler, C. M. (2002). The Tim9p-Tim10p Complex Binds to the Transmembrane Domains of the ADP/ATP Carrier. *EMBO J.* 21, 942–953. doi:10.1093/emboj/21.5.942
- Davis, A. J., Alder, N. N., Jensen, R. E., and Johnson, A. E. (2007). The Tim9p/10p and Tim8p/13p Complexes Bind to Specific Sites on Tim23p during Mitochondrial Protein Import. *MBoC* 18, 475–486. doi:10.1091/mbc.e06-06-0546
- Dekker, C., Kruijff, B. d., and Gros, P. (2003). Crystal Structure of SecB from *Escherichia coli*. *J. Struct. Biol.* 144, 313–319. doi:10.1016/j.jsb.2003.09.012
- Ellis, R. J. (1990). The Molecular Chaperone Concept. *Semin. Cell Biol.* 1, 1–9.
- Feldman, D. E., and Frydman, J. (2000). Protein Folding *In Vivo*: the Importance of Molecular Chaperones. *Curr. Opin. Struct. Biol.* 10, 26–33. doi:10.1016/s0959-440x(99)00044-5
- Ferreiro, D. U., Komives, E. A., and Wolynes, P. G. (2014). Frustration in Biomolecules. *Quart. Rev. Biophys.* 47, 285–363. doi:10.1017/s0033583514000092
- Finka, A., and Goloubinoff, P. (2013). Proteomic Data from Human Cell Cultures Refine Mechanisms of Chaperone-Mediated Protein Homeostasis. *Cell Stress and Chaperones* 18, 591–605. doi:10.1007/s12192-013-0413-3
- Fotin, A., Cheng, Y., Grigorieff, N., Walz, T., Harrison, S. C., and Kirchhausen, T. (2004). Structure of an Auxilin-Bound Clathrin Coat and its Implications for the Mechanism of Uncoating. *Nature* 432, 649–653. doi:10.1038/nature03078
- Fürst, M., Zhou, Y., Merfort, J., and Müller, M. (2018). Involvement of PpiD in Sec-Dependent Protein Translocation. *Biochim. Biophys. Acta (Bba) - Mol. Cell Res.* 1865, 273–280. doi:10.1016/j.bbamcr.2017.10.012
- Ge, X., Lyu, Z.-X., Liu, Y., Wang, R., Zhao, X. S., Fu, X., et al. (2014). Identification of FkpA as a Key Quality Control Factor for the Biogenesis of Outer Membrane Proteins under Heat Shock Conditions. *J. Bacteriol.* 196, 672–680. doi:10.1128/jb.01069-13
- Geibert, N., Chacinska, A., Wagner, K., Guaiard, B., Koehler, C. M., Rehling, P., et al. (2008). Assembly of the Three Small Tim Proteins Precedes Docking to the Mitochondrial Carrier Translocase. *EMBO Rep.* 9, 548–554. doi:10.1038/embor.2008.49
- Gentle, I., Gabriel, K., Beech, P., Waller, R., and Lithgow, T. (2004). The Omp85 Family of Proteins Is Essential for Outer Membrane Biogenesis in Mitochondria and Bacteria. *J. Cell Biol.* 164, 19–24. doi:10.1083/jcb.200310092
- Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014). Energy Landscapes of Functional Proteins Are Inherently Risky. *Nat. Chem. Biol.* 10, 884–891. doi:10.1038/nchembio.1670
- Goemans, C., Denoncin, K., and Collet, J.-F. (2014). Folding Mechanisms of Periplasmic Proteins. *Biochim. Biophys. Acta (Bba) - Mol. Cell Res.* 1843, 1517–1528. doi:10.1016/j.bbamcr.2013.10.014
- Habib, S. J., Waizenegger, T., Lech, M., Neupert, W., and Rapaport, D. (2005). Assembly of the TOB Complex of Mitochondria. *J. Biol. Chem.* 280, 6434–6440. doi:10.1074/jbc.m411510200
- Hardy, S., and Randall, L. (1991). A Kinetic Partitioning Model of Selective Binding of Nonnative Proteins by the Bacterial Chaperone SecB. *Science* 251, 439–443. doi:10.1126/science.1989077
- Hartl, F.-U., Lecker, S., Schiebel, E., Hendrick, J. P., and Wickner, W. (1990). The Binding cascade of SecB to SecA to SecYE Mediates Preprotein Targeting to the *E. coli* Plasma Membrane. *Cell* 63, 269–279. doi:10.1016/0092-8674(90)90160-g
- Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular Chaperones in Protein Folding and Proteostasis. *Nature* 475, 324–332. doi:10.1038/nature10317
- Hartl, F. U., Martin, J., and Neupert, W. (1992). Protein Folding in the Cell: the Role of Molecular Chaperones Hsp70 and Hsp60. *Annu. Rev. Biophys. Biomol. Struct.* 21, 293–322. doi:10.1146/annurev.bb.21.060192.001453
- Haslbeck, M., Walke, S., Stromer, T., Ehrnsperger, M., White, H. E., Chen, S., et al. (1999). Hsp26: a Temperature-Regulated Chaperone. *EMBO J.* 18, 6744–6751. doi:10.1093/emboj/18.23.6744
- He, C., Pan, S., Li, G., and Zhao, X. S. (2020). Molecular Mechanism of SurA's Chaperoning Function to Outer Membrane Proteins Revealed by Purification-After-Crosslinking Single-Molecule FRET. *Sci. China Chem.* 63, 1142–1152. doi:10.1007/s11426-020-9758-2
- He, L., Sharpe, T., Mazur, A., and Hiller, S. (2016). A Molecular Mechanism of Chaperone-Client Recognition. *Sci. Adv.* 2, e1601625–12. doi:10.1126/sciadv.1601625
- He, L., and Hiller, S. (2018). Common Patterns in Chaperone Interactions with a Native Client Protein. *Angew. Chem. Int. Ed.* 57, 5921–5924. doi:10.1002/anie.201713064
- He, L., and Hiller, S. (2019). Frustrated Interfaces Facilitate Dynamic Interactions between Native Client Proteins and Holdase Chaperones. *ChemBioChem* 20, 2803–2806. doi:10.1002/cbic.201900215
- Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J., and Behrens, S. (2005). The Periplasmic Chaperone SurA Exploits Two Features Characteristic of Integral Outer Membrane Proteins for Selective Substrate Recognition. *J. Biol. Chem.* 280, 23540–23548. doi:10.1074/jbc.m413742200
- Hiller, S., and Burmann, B. M. (2018). Chaperone-client Complexes: A Dynamic Liaison. *J. Magn. Reson.* 289, 142–155. doi:10.1016/j.jmr.2017.12.008
- Hoffmann, A., Bukau, B., and Kramer, G. (2010). Structure and Function of the Molecular Chaperone Trigger Factor. *Biochim. Biophys. Acta (Bba) - Mol. Cell Res.* 1803, 650–661. doi:10.1016/j.bbamcr.2010.01.017
- Höhr, A. I. C., Straub, S. P., Warscheid, B., Becker, T., and Wiedemann, N. (2015). Assembly of  $\beta$ -barrel Proteins in the Mitochondrial Outer Membrane. *Biochim. Biophys. Acta (Bba) - Mol. Cell Res.* 1853, 74–88. doi:10.1016/j.bbamcr.2014.10.006
- Holdbrook, D. A., Burmann, B. M., Huber, R. G., Petoukhov, M. V., Svergun, D. I., Hiller, S., et al. (2017). A Spring-Loaded Mechanism Governs the Clamp-like Dynamics of the Skp Chaperone. *Structure* 25, 1079–1088. doi:10.1016/j.str.2017.05.018
- Hoppins, S. C., and Nargang, F. E. (2004). The Tim8-Tim13 Complex of *Neurospora Crassa* Functions in the Assembly of Proteins into Both Mitochondrial Membranes. *J. Biol. Chem.* 279, 12396–12405. doi:10.1074/jbc.m313037200
- Horowitz, A. (1998). Structural Aspects of GroEL Function. *Curr. Opin. Struct. Biol.* 8, 93–100. doi:10.1016/s0959-440x(98)80015-8

- Horowitz, S., Salmon, L., Koldewey, P., Ahlstrom, L. S., Martin, R., Quan, S., et al. (2016). Visualizing Chaperone-Assisted Protein Folding. *Nat. Struct. Mol. Biol.* 23, 691–697. doi:10.1038/nsmb.3237
- Huang, C., Rossi, P., Saio, T., and Kalodimos, C. G. (2016). Structural Basis for the Antifolding Activity of a Molecular Chaperone. *Nature* 537, 202–206. doi:10.1038/nature18965
- Hwang, W., and Karplus, M. (2019). Structural Basis for Power Stroke vs. Brownian Ratchet Mechanisms of Motor Proteins. *Proc. Natl. Acad. Sci. USA* 116, 19777–19785. doi:10.1073/pnas.1818589116
- Jakob, U., and Buchner, J. (1994). Assisting Spontaneity: the Role of Hsp90 and Small Hsps as Molecular Chaperones. *Trends Biochemical Sciences* 19, 205–211. doi:10.1016/0968-0004(94)90023-x
- Jarchow, S., Lück, C., Görg, A., and Skerra, A. (2008). Identification of Potential Substrate Proteins for the Periplasmic *Escherichia coli* Chaperone Skp. *Proteomics* 8, 4987–4994. doi:10.1002/pmic.200800288
- Jehle, S., Rajagopal, P., Bardiaux, B., Markovic, S., Kühne, R., Stout, J. R., et al. (2010). Solid-state NMR and SAXS Studies Provide a Structural Basis for the Activation of  $\alpha$ B-crystallin Oligomers. *Nat. Struct. Mol. Biol.* 17, 1037–1042. doi:10.1038/nsmb.1891
- Jia, M., Wu, B., Yang, Z., Chen, C., Zhao, M., Hou, X., et al. (2020). Conformational Dynamics of the Periplasmic Chaperone SurA. *Biochemistry* 59, 3235–3246. doi:10.1021/acs.biochem.0c00507
- Jiang, J., Zhang, X., Chen, Y., Wu, Y., Zhou, Z. H., Chang, Z., et al. (2008). Activation of DegP Chaperone-Protease via Formation of Large Cage-like Oligomers upon Binding to Substrate Proteins. *Proc. Natl. Acad. Sci.* 105, 11939–11944. doi:10.1073/pnas.0805464105
- Jiang, Y., Rossi, P., and Kalodimos, C. G. (2019). Structural Basis for Client Recognition and Activity of Hsp40 Chaperones. *Science* 365, 1313–1319. doi:10.1126/science.aax1280
- Jin, M., Han, W., Liu, C., Zang, Y., Li, J., Wang, F., et al. (2019). An Ensemble of Cryo-EM Structures of TRiC Reveal its Conformational Landscape and Subunit Specificity. *Proc. Natl. Acad. Sci. USA* 116, 19513–19522. doi:10.1073/pnas.1903976116
- Jores, T., Lawatscheck, J., Beke, V., Franz-Wachtel, M., Yunoki, K., Fitzgerald, J. C., et al. (2018). Cytosolic Hsp70 and Hsp40 Chaperones Enable the Biogenesis of Mitochondrial  $\beta$ -barrel Proteins. *J. Cel Biol.* 217, 3091–3108. doi:10.1083/jcb.201712029
- Jores, T., Klinger, A., Groß, L. E., Kawano, S., Flinner, N., Duchardt-Ferner, E., et al. (2016). Characterization of the Targeting Signal in Mitochondrial  $\beta$ -barrel Proteins. *Nat. Commun.* 7, 12036. doi:10.1038/ncomms12036
- Jorgensen, P. L., Håkansson, K. O., and Karlsh, S. J. D. (2003). Structure and Mechanism of Na,K-ATPase: Functional Sites and Their Interactions. *Annu. Rev. Physiol.* 65, 817–849. doi:10.1146/annurev.physiol.65.092101.142558
- Kafri, G., and Horowitz, A. (2003). Transient Kinetic Analysis of ATP-Induced Allosteric Transitions in the Eukaryotic Chaperonin Containing TCP-1. *J. Mol. Biol.* 326, 981–987. doi:10.1016/s0022-2836(03)00046-9
- Kakkar, V., Kuiper, E. F., Pandey, A., Braakman, L., and Kampinga, H. H. (2016a). Versatile Members of the Dnaj Family Show Hsp70 Dependent Anti-aggregation Activity on Ring1 Mutant Parkin C289g. *Sci. Rep.* 6, 34830–34912. doi:10.1038/srep34830
- Kakkar, V., Månsson, C., de Mattos, E. P., Bergink, S., van der Zwaag, M., van Waarde, M. A. W. H., et al. (2016b). The S/t-Rich Motif in the Dnajb6 Chaperone Delays Polyglutamine Aggregation and the Onset of Disease in a Mouse Model. *Mol. Cel* 62, 272–283. doi:10.1016/j.molcel.2016.03.017
- Kampinga, H. H., and Craig, E. A. (2010). The HSP70 Chaperone Machinery: J Proteins as Drivers of Functional Specificity. *Nat. Rev. Mol. Cel Biol.* 11, 579–592. doi:10.1038/nrm2941
- Karagöz, G. E., Duarte, A. M. S., Akoury, E., Ippel, H., Biernat, J., Morán Luengo, T., et al. (2014). Hsp90- $\tau$  Complex Reveals Molecular Basis for Specificity in Chaperone Action. *Cell* 156, 963–974. doi:10.1016/j.cell.2014.01.037
- Kelley, W., and Georgopoulos, C. (1992). Chaperones and Protein Folding. *Curr. Opin. Cel Biol.* 4, 984–991. doi:10.1016/0955-0674(92)90130-5
- Kim, D. H., Xu, Z.-Y., Na, Y. J., Yoo, Y.-J., Lee, J., Sohn, E.-J., et al. (2011). Small Heat Shock Protein Hsp17.8 Functions as an AKR2A Cofactor in the Targeting of Chloroplast Outer Membrane Proteins in Arabidopsis. *Plant Physiol.* 157, 132–146. doi:10.1104/pp.111.178681
- Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M., and Ulrich Hartl, F. (2013). Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annu. Rev. Biochem.* 82, 323–355. doi:10.1146/annurev-biochem-060208-092442
- Knoblauch, N. T. M., Rüdiger, S., Schönfeld, H.-J., Driessen, A. J. M., Schneider-Mergener, J., and Bukau, B. (1999). Substrate Specificity of the SecB Chaperone. *J. Biol. Chem.* 274, 34219–34225. doi:10.1074/jbc.274.48.34219
- Koehler, C. M., Jarosch, E., Tokatlidis, K., Schmid, K., Schweyen, R. J., and Schatz, G. (1998a). Import of Mitochondrial Carriers Mediated by Essential Proteins of the Intermembrane Space. *Science* 279, 369–373. doi:10.1126/science.279.5349.369
- Koehler, C. M., Merchant, S., Oppliger, W., Schmid, K., Jarosch, E., Dolfini, L., et al. (1998b). Tim9p, an Essential Partner Subunit of Tim10p for the Import of Mitochondrial Carrier Proteins. *EMBO J.* 17, 6477–6486. doi:10.1093/emboj/17.22.6477
- Koehler, C. M., Merchant, S., and Schatz, G. (1999). How Membrane Proteins Travel across the Mitochondrial Intermembrane Space. *Trends Biochem. Sci.* 24, 428–432. doi:10.1016/s0968-0004(99)01462-0
- Koldewey, P., Horowitz, S., and Bardwell, J. C. A. (2017). Chaperone-client Interactions: Non-specificity Engenders Multifunctionality. *J. Biol. Chem.* 292, 12010–12017. doi:10.1074/jbc.r117.796862
- Koldewey, P., Stull, F., Horowitz, S., Martin, R., and Bardwell, J. C. A. (2016). Forces Driving Chaperone Action. *Cell* 166, 369–379. doi:10.1016/j.cell.2016.05.054
- Komarudin, A. G., and Driessen, A. J. M. (2019). SecA-Mediated Protein Translocation through the SecYEG Channel. *Microbiol. Spectr.* 7 (4). doi:10.1128/microbiolspec.PSIB-0028-2019
- Korndörfer, I. P., Dommel, M. K., and Skerra, A. (2004). Structure of the Periplasmic Chaperone Skp Suggests Functional Similarity with Cytosolic Chaperones Despite Differing Architecture. *Nat. Struct. Mol. Biol.* 11, 1015–1020. doi:10.1038/nsmb828
- Kozjak, V., Wiedemann, N., Milenkovic, D., Lohaus, C., Meyer, H. E., Guiard, B., et al. (2003). An Essential Role of Sam50 in the Protein Sorting and Assembly Machinery of the Mitochondrial Outer Membrane. *J. Biol. Chem.* 278, 48520–48523. doi:10.1074/jbc.c300442200
- Krojer, T., Garrido-Franco, M., Huber, R., Ehrmann, M., and Clausen, T. (2002). Crystal Structure of DegP (HtrA) Reveals a New Protease-Chaperone Machine. *Nature* 416, 455–459. doi:10.1038/416455a
- Krojer, T., Pangel, K., Kurt, J., Sawa, J., Sting, C., Mechtler, K., et al. (2008a). Interplay of PDZ and Protease Domain of DegP Ensures Efficient Elimination of Misfolded Proteins. *Proc. Natl. Acad. Sci.* 105, 7702–7707. doi:10.1073/pnas.0803392105
- Krojer, T., Sawa, J., Schäfer, E., Saibil, H. R., Ehrmann, M., and Clausen, T. (2008b). Structural Basis for the Regulated Protease and Chaperone Function of DegP. *Nature* 453, 885–890. doi:10.1038/nature07004
- Kuehn, M., Ogg, D., Kihlberg, J., Slonim, L., Flemmer, K., Bergfors, T., et al. (1993). Structural Basis of Pilus Subunit Recognition by the PapD Chaperone. *Science* 262, 1234–1241. doi:10.1126/science.7901913
- Lecker, S., Lill, R., Ziegelhoffer, T., Georgopoulos, C., Bassford, P. J., Kumamoto, C. A., et al. (1989). Three Pure Chaperone Proteins of *Escherichia coli*--SecB, Trigger Factor and GroEL--form Soluble Complexes with Precursor Proteins *In Vitro*. *EMBO J.* 8, 2703–2709. doi:10.1002/j.1460-2075.1989.tb08411.x
- Lee, J., Lee, H., Kim, J., Lee, S., Kim, D. H., Kim, S., et al. (2011). Both the Hydrophobicity and a Positively Charged Region Flanking the C-Terminal Region of the Transmembrane Domain of Signal-Anchored Proteins Play Critical Roles in Determining Their Targeting Specificity to the Endoplasmic Reticulum or Endosymbiotic Organelles in Arabidopsis Cells. *Plant Cell* 23, 1588–1607. doi:10.1105/tpc.110.082230
- Lithgow, T., and Schneider, A. (2010). Evolution of Macromolecular Import Pathways in Mitochondria, Hydrogenosomes and Mitosomes. *Phil. Trans. R. Soc. B* 365, 799–817. doi:10.1098/rstb.2009.0167
- Lopez, T., Dalton, K., and Frydman, J. (2015). The Mechanism and Function of Group II Chaperonins. *J. Mol. Biol.* 427, 2919–2930. doi:10.1016/j.jmb.2015.04.013
- Lycklama a Nijeholt, J. A., and Driessen, A. J. M. (2012). The Bacterial Sec-Translocase: Structure and Mechanism. *Phil. Trans. R. Soc. B* 367, 1016–1028. doi:10.1098/rstb.2011.0201
- Maduke, M., and Roise, D. (1996). Structure and Function of Mitochondrial Presequences. *Membr. Protein Transport* 3, 49–79. doi:10.1016/s1874-592x(96)80004-7
- Malet, H., Canellas, F., Sawa, J., Yan, J., Thalassinou, K., Ehrmann, M., et al. (2012). Newly Folded Substrates inside the Molecular Cage of the HtrA Chaperone DegQ. *Nat. Struct. Mol. Biol.* 19, 152–157. doi:10.1038/nsmb.2210

- Martinez-Hackert, E., and Hendrickson, W. A. (2009). Promiscuous Substrate Recognition in Folding and Assembly Activities of the Trigger Factor Chaperone. *Cell* 138, 923–934. doi:10.1016/j.cell.2009.07.044
- Marx, D. C., Leblanc, M. J., Plummer, A. M., Krueger, S., and Fleming, K. G. (2020a). Domain Interactions Determine the Conformational Ensemble of the Periplasmic Chaperone SurA. *Protein Sci.* 29, 2043–2053. doi:10.1002/pro.3924
- Marx, D. C., Plummer, A. M., Faustino, A. M., Devlin, T., Roskopf, M. A., Leblanc, M. J., et al. (2020b). SurA Is a Cryptically Grooved Chaperone that Expands Unfolded Outer Membrane Proteins. *Proc. Natl. Acad. Sci. USA* 117, 28026–28035. doi:10.1073/pnas.2008175117
- Mas, G., Burmann, B. M., Sharpe, T., Claudi, B., Bumann, D., and Hiller, S. (2020). Regulation of Chaperone Function by Coupled Folding and Oligomerization. *Sci. Adv.* 6, eabc5822. doi:10.1126/sciadv.abc5822
- Matern, Y., Barion, B., and Behrens-Kneip, S. (2010). PpiD Is a Player in the Network of Periplasmic Chaperones in *Escherichia coli*. *BMC Microbiol.* 10, 251. doi:10.1186/1471-2180-10-251
- Mayer, M. P. (2021). The Hsp70-Chaperone Machines in Bacteria. *Front. Mol. Biosci.* 8, 694012. doi:10.3389/fmolb.2021.694012
- Mayer, M. P., and Bukau, B. (2005). Hsp70 Chaperones: Cellular Functions and Molecular Mechanism. *CMLS, Cel. Mol. Life Sci.* 62, 670–684. doi:10.1007/s00018-004-4464-6
- Mayer, M. P., and Gierasch, L. M. (2019). Recent Advances in the Structural and Mechanistic Aspects of Hsp70 Molecular Chaperones. *J. Biol. Chem.* 294 (6), 2085–2097. doi:10.1074/jbc.REV118.002810
- McMorran, L. M., Brockwell, D. J., and Radford, S. E. (2015). Mechanistic Studies of the Biogenesis and Folding of Outer Membrane Proteins *In Vitro* and *In Vivo*: What Have We Learned to Date? *Arch. Biochem. Biophys.* 564, 265–280. doi:10.1016/j.abb.2014.02.011
- Medeiros, R., Baglietto-Vargas, D., and Laferla, F. M. (2011). The Role of Tau in Alzheimer's Disease and Related Disorders. *CNS Neurosci. Ther.* 17, 514–524. doi:10.1111/j.1755-5949.2010.00177.x
- Merdanovic, M., Clausen, T., Kaiser, M., Huber, R., and Ehrmann, M. (2011). Protein Quality Control in the Bacterial Periplasm. *Annu. Rev. Microbiol.* 65, 149–168. doi:10.1146/annurev-micro-090110-102925
- Meyer, A. S., Gillespie, J. R., Walther, D., Millet, I. S., Doniach, S., and Frydman, J. (2003). Closing the Folding Chamber of the Eukaryotic Chaperonin Requires the Transition State of ATP Hydrolysis. *Cell* 113, 369–381. doi:10.1016/s0092-8674(03)00307-6
- Mitra, R., Gadkari, V. V., Meinen, B. A., van Mierlo, C. P. M., Ruotolo, B. T., and Bardwell, J. C. A. (2021). Mechanism of the Small ATP-independent Chaperone Spy Is Substrate Specific. *Nat. Commun.* 12, 851–913. doi:10.1038/s41467-021-21120-8
- Morgenstern, M., Stiller, S. B., Lübbert, P., Peikert, C. D., Dannenmaier, S., Drepper, F., et al. (2017). Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. *Cel Rep.* 19, 2836–2852. doi:10.1016/j.celrep.2017.06.014
- Nakagawa, A., Moriya, K., Arita, M., Yamamoto, Y., Kitamura, K., Ishiguro, N., et al. (2014). Dissection of the ATP-dependent Conformational Change Cycle of a Group II Chaperonin. *J. Mol. Biol.* 426, 447–459. doi:10.1016/j.jmb.2013.09.034
- Neupert, W., Hartl, F.-U., Craig, E. A., and Pfanner, N. (1990). How Do Polypeptides Cross the Mitochondrial Membranes? *Cell* 63, 447–450. doi:10.1016/0092-8674(90)90437-j
- Nunes, J. M., Mayer-Hartl, M., Hartl, F. U., and Müller, D. J. (2015). Action of the Hsp70 Chaperone System Observed with Single Proteins. *Nat. Commun.* 6, 6307. doi:10.1038/ncomms7307
- Oswald, J., Njenga, R., Natriashvili, A., Sarmah, P., and Koch, H. G. (2021). The Dynamic SecYEG Translocon. *Front. Mol. Biosci.* 8, 1–30. doi:10.3389/fmolb.2021.664241
- Pan, S., Yang, C., and Zhao, X. S. (2020). Affinity of Skp to OmpC Revealed by Single-Molecule Detection. *Sci. Rep.* 10, 14871. doi:10.1038/s41598-020-71608-4
- Paschen, S. A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., et al. (2003). Evolutionary Conservation of Biogenesis of  $\beta$ -barrel Membrane Proteins. *Nature* 426, 862–866. doi:10.1038/nature02208
- Priya, S., Sharma, S. K., Sood, V., Mattoo, R. U. H., Finka, A., Azem, A., et al. (2013). Groel and Cct Are Catalytic Unfoldases Mediating Out-Of-Cage Polypeptide Refolding without Atp. *Proc. Natl. Acad. Sci.* 110, 7199–7204. doi:10.1073/pnas.1219867110
- Qu, J., Behrens-Kneip, S., Holst, O., and Kleinschmidt, J. H. (2009). Binding Regions of Outer Membrane Protein A in Complexes with the Periplasmic Chaperone Skp. A Site-Directed Fluorescence Study. *Biochemistry* 48, 4926–4936. doi:10.1021/bi9004039
- Qu, J., Mayer, C., Behrens, S., Holst, O., and Kleinschmidt, J. H. (2007). The Trimeric Periplasmic Chaperone Skp of *Escherichia coli* Forms 1:1 Complexes with Outer Membrane Proteins via Hydrophobic and Electrostatic Interactions. *J. Mol. Biol.* 374, 91–105. doi:10.1016/j.jmb.2007.09.020
- Quan, S., Wang, L., Petrotchenko, E. V., Makepeace, K. A., Horowitz, S., Yang, J., et al. (2014). Super Spy Variants Implicate Flexibility in Chaperone Action. *Elife* 3, e01584. doi:10.7554/eLife.01584
- Quan, S., Koldewey, P., Tapley, T., Kirsch, N., Ruane, K. M., Pfizenmaier, J., et al. (2011). Genetic Selection Designed to Stabilize Proteins Uncovers a Chaperone Called Spy. *Nat. Struct. Mol. Biol.* 18, 262–269. doi:10.1038/nsmb.2016
- Randall, L. L., Topping, T. B., Suci, D., and Hardy, S. J. S. (1998). Calorimetric Analyses of the Interaction between SecB and its Ligands. *Protein Sci.* 7, 1195–1200. doi:10.1002/pro.5560070514
- Randall, L. (1992). Peptide Binding by Chaperone SecB: Implications for Recognition of Nonnative Structure. *Science* 257, 241–245. doi:10.1126/science.1631545
- Rehling, P., Model, K., Brandner, K., Kovermann, P., Sickmann, A., Meyer, H. E., et al. (2003). Protein Insertion into the Mitochondrial Inner Membrane by a Twin-Pore Translocase. *Science* 299, 1747–1751. doi:10.1126/science.1080945
- Reissmann, S., Parnot, C., Booth, C. R., Chiu, W., and Frydman, J. (2007). Essential Function of the Built-In Lid in the Allosteric Regulation of Eukaryotic and Archaeal Chaperonins. *Nat. Struct. Mol. Biol.* 14, 432–440. doi:10.1038/nsmb1236
- Rocchio, S., Duman, R., El Omari, K., Mykhaylyk, V., Orr, C., Yan, Z., et al. (2019). Identifying Dynamic, Partially Occupied Residues Using Anomalous Scattering. *Acta Cryst. Sect D Struct. Biol.* 75, 1084–1095. doi:10.1107/s2059798319014475
- Roesch, K., Hynds, P. J., Varga, R., Tranebjærge, L., and Koehler, C. M. (2004). The Calcium-Binding Aspartate/glutamate Carriers, Citrin and Aralar1, Are New Substrates for the DDP1/TIMM8a-TIMM13 Complex. *Hum. Mol. Genet.* 13, 2101–2111. doi:10.1093/hmg/ddh217
- Rollauer, S. E., Soorshjani, M. A., Noinaj, N., and Buchanan, S. K. (2015). Outer Membrane Protein Biogenesis in Gram-Negative Bacteria. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370. doi:10.1098/rstb.2015.0023
- Rosenzweig, R., Nillegoda, N. B., Mayer, M. P., and Bukau, B. (2019). The Hsp70 Chaperone Network. *Nat. Rev. Mol. Cel Biol.* 20, 665–680. doi:10.1038/s41580-019-0133-3
- Rüdiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate Specificity of the DnaK Chaperone Determined by Screening Cellulose-Bound Peptide Libraries. *EMBO J.* 16, 1501–1507. doi:10.1093/emboj/16.7.1501
- Sahi, C., and Craig, E. A. (2007). Network of General and Specialty J Protein Chaperones of the Yeast Cytosol. *PNAS* 104, 7163–7168. doi:10.1073/pnas.0702357104
- Saio, T., Guan, X., Rossi, P., Economou, A., and Kalodimos, C. G. (2014). Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone. *Science* 344, 1250494. doi:10.1126/science.1250494
- Sala, A., Bordes, P., and Genevieux, P. (2014). Multitasking SecB Chaperones in Bacteria. *Front. Microbiol.* 5, 666–712. doi:10.3389/fmicb.2014.00666
- Salmon, L., Ahlstrom, L. S., Horowitz, S., Dickson, A., Brooks, C. L., and Bardwell, J. C. A. (2016). Capturing a Dynamic Chaperone-Substrate Interaction Using NMR-Informed Molecular Modeling. *J. Am. Chem. Soc.* 138, 9826–9839. doi:10.1021/jacs.6b02382
- Sandlin, C. W., Zaccari, N. R., and Fleming, K. G. (2015). Skp Trimer Formation Is Insensitive to Salts in the Physiological Range. *Biochemistry* 54, 7059–7062. doi:10.1021/acs.biochem.5b00806
- Sauer, F. G., Knight, S. D., Waksman, G. J., and Hultgren, S. J. (2000). PapD-like Chaperones and Pilus Biogenesis. *Semin. Cel Dev. Biol.* 11, 27–34. doi:10.1006/scdb.1999.0348
- Sawa, J., Malet, H., Krojer, T., Canellas, F., Ehrmann, M., and Clausen, T. (2011). Molecular Adaptation of the DegQ Protease to Exert Protein Quality Control in the Bacterial Cell Envelope. *J. Biol. Chem.* 286, 30680–30690. doi:10.1074/jbc.m111.243832
- Schiffirin, B., Brockwell, D. J., and Radford, S. E. (2017). Outer Membrane Protein Folding from an Energy Landscape Perspective. *BMC Biol.* 15, 123–216. doi:10.1186/s12915-017-0464-5
- Schiffirin, B., Calabrese, A. N., Devine, P. W. A., Harris, S. A., Ashcroft, A. E., Brockwell, D. J., et al. (2016). Skp Is a Multivalent Chaperone of Outer-



- Membrane Proteins. *Nat. Struct. Mol. Biol.* 23, 786–793. doi:10.1038/nsmb.3266
- Shan, S.-o. (2019). Guiding Tail-Anchored Membrane Proteins to the Endoplasmic Reticulum in a Chaperone cascade. *J. Biol. Chem.* 294, 16577–16586. doi:10.1074/jbc.rev119.006197
- Shiota, T., Imai, K., Qiu, J., Hewitt, V. L., Tan, K., Shen, H.-h., et al. (2015). Molecular Architecture of the Active Mitochondrial Protein Gate. *Science* 349, 1544–1548. doi:10.1126/science.aac6428
- Silver, P. A., and Way, J. C. (1993). Eukaryotic DnaJ Homologs and the Specificity of Hsp70 Activity. *Cell* 74, 5–6. doi:10.1016/0092-8674(93)90287-z
- Skjærven, L., Cuellar, J., Martínez, A., and Valpuesta, J. M. (2015). Dynamics, Flexibility, and Allostery in Molecular Chaperonins. *FEBS Lett.* 589, 2522–2532. doi:10.1016/j.febslet.2015.06.019
- Sklar, J. G., Wu, T., Kahne, D., and Silhavy, T. J. (2007). Defining the Roles of the Periplasmic Chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev.* 21, 2473–2484. doi:10.1101/gad.1581007
- Soltes, G. R., Schwalm, J., Ricci, D. P., and Silhavy, T. J. (2016). The Activity of *Escherichia coli* Chaperone SurA Is Regulated by Conformational Changes Involving a Parvulin Domain. *J. Bacteriol.* 198, 921–929. doi:10.1128/jb.00889-15
- Spieß, C., Beil, A., and Ehrmann, M. (1999). A Temperature-dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. *Cell* 97, 339–347. doi:10.1016/s0092-8674(00)80743-6
- Spieß, C., Miller, E. J., McClellan, A. J., and Frydman, J. (2006). Identification of the TRiC/CCT Substrate Binding Sites Uncovers the Function of Subunit Diversity in Eukaryotic Chaperonins. *Mol. Cell* 24, 25–37. doi:10.1016/j.molcel.2006.09.003
- Stenberg, G., and Fersht, A. R. (1997). Folding of Barnase in the Presence of the Molecular Chaperone SecB. *J. Mol. Biol.* 274, 268–275. doi:10.1006/jmbi.1997.1398
- Stull, F., Koldewey, P., Humes, J. R., Radford, S. E., Bardwell, J. C. A., and Biology, C. (2016). Substrate Protein Folds while it Is Bound to the ATP-independent Chaperone Spy. *Nat. Struct. Mol. Biol.* 23, 53–58. doi:10.1038/nsmb.3133
- Subrini, O., and Betton, J.-M. (2009). Assemblies of DegP Underlie its Dual Chaperone and Protease Function. *FEMS Microbiol. Lett.* 296, 143–148. doi:10.1111/j.1574-6968.2009.01658.x
- Sućec, I., Wang, Y., Dakhlou, O., Weinhäupl, K., Jores, T., Costa, D., et al. (2020). Structural Basis of Client Specificity in Mitochondrial Membrane-Protein Chaperones. *Sci. Adv.* 6, eabd0263. doi:10.1126/sciadv.abd0263
- Suo, Y., Hardy, S. J. S., and Randall, L. L. (2015). The Basis of Asymmetry in the SecA:SecB Complex. *J. Mol. Biol.* 427, 887–900. doi:10.1016/j.jmb.2014.12.008
- Szolańska, E., and Chroboczek, J. (2011). Faithful Chaperones. *Cell. Mol. Life Sci.* 68, 3307–3322. doi:10.1007/s00018-011-0740-4
- Taipale, M., Jarosz, D. F., and Lindquist, S. (2010). Hsp90 at the Hub of Protein Homeostasis: Emerging Mechanistic Insights. *Nat. Rev. Mol. Cell Biol.* 11, 515–528. doi:10.1038/nrm2918
- Thoma, J., Burmann, B. M., Hiller, S., and Müller, D. J. (2015). Impact of Holdase Chaperones Skp and SurA on the Folding of  $\beta$ -barrel Outer-Membrane Proteins. *Nat. Struct. Mol. Biol.* 22, 795–802. doi:10.1038/nsmb.3087
- Tomasek, D., and Kahne, D. (2021). The Assembly of  $\beta$ -barrel Outer Membrane Proteins. *Curr. Opin. Microbiol.* 60, 16–23. doi:10.1016/j.mib.2021.01.009
- Tomba, P., and Fuxreiter, M. (2008). Fuzzy Complexes: Polymorphism and Structural Disorder in Protein-Protein Interactions. *Trends Biochem. Sci.* 33, 2–8. doi:10.1016/j.tibs.2007.10.003
- Tsirigotaki, A., De Geyter, J., Šoštarić, N., Economou, A., and Karamanou, S. (2017). Protein export through the Bacterial Sec Pathway. *Nat. Rev. Microbiol.* 15, 21–36. doi:10.1038/nrmicro.2016.161
- Vainberg, I. E., Lewis, S. A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H. L., et al. (1998). Prefoldin, a Chaperone that Delivers Unfolded Proteins to Cytosolic Chaperonin. *Cell* 93, 863–873. doi:10.1016/s0092-8674(00)81446-4
- Vergnolle, M. A. S., Baud, C., Golovanov, A. P., Alcock, F., Luciano, P., Lian, L.-Y., et al. (2005). Distinct Domains of Small Tims Involved in Subunit Interaction and Substrate Recognition. *J. Mol. Biol.* 351, 839–849. doi:10.1016/j.jmb.2005.06.010
- Vickery, L. E., and Cupp-Vickery, J. R. (2007). Molecular Chaperones HscA/ssq1 and HscB/jac1 and Their Roles in Iron-Sulfur Protein Maturation. *Crit. Rev. Biochem. Mol. Biol.* 42, 95–111. doi:10.1080/10409230701322298
- von Heijne, G. (2002). “Targeting Sequences,” in *Protein Targeting, Transport, and Translocation*. Editors R. E. Dalbey and G. von Heijne (London: Academic Press), 35–46. doi:10.1016/B978-012200731-6.50005-7
- Walton, T. A., Sandoval, C. M., Fowler, C. A., Pardi, A., and Sousa, M. C. (2009). The Cavity-Chaperone Skp Protects its Substrate from Aggregation but Allows Independent Folding of Substrate Domains. *Pnas* 106, 1772–1777. doi:10.1073/pnas.0809275106
- Walton, T. A., and Sousa, M. C. (2004). Crystal Structure of Skp, a Prefoldin-like Chaperone that Protects Soluble and Membrane Proteins from Aggregation. *Mol. Cell* 15, 367–374. doi:10.1016/j.molcel.2004.07.023
- Wang, J. (2018). Misreading Chaperone-Substrate Complexes from Random Noise. *Nat. Struct. Mol. Biol.* 25, 989–990. doi:10.1038/s41594-018-0144-3
- Wang, X., Peterson, J. H., and Bernstein, H. D. (2021). Bacterial Outer Membrane Proteins Are Targeted to the Bam Complex by Two Parallel Mechanisms. *mBio* 12. doi:10.1128/mbio.00597-21
- Wang, Y., Wang, R., Jin, F., Liu, Y., Yu, J., Fu, X., et al. (2016). A Supercomplex Spanning the Inner and Outer Membranes Mediates the Biogenesis of  $\beta$ -Barrel Outer Membrane Proteins in Bacteria. *J. Biol. Chem.* 291, 16720–16729. doi:10.1074/jbc.m115.710715
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). Swiss-model: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* 46, W296–W303. doi:10.1093/nar/gky427
- Webb, C. T., Gorman, M. A., Lazarou, M., Ryan, M. T., and Gulbis, J. M. (2006). Crystal Structure of the Mitochondrial Chaperone TIM910 Reveals a Six-Bladed  $\alpha$ -Propeller. *Mol. Cell* 21, 123–133. doi:10.1016/j.molcel.2005.11.010
- Weickert, S., Wawrzyniuk, M., John, L. H., Rüdiger, S. G. D., and Drescher, M. (2020). The Mechanism of Hsp90-Induced Oligomerization of Tau. *Sci. Adv.* 6, eaax6999. doi:10.1126/sciadv.aax6999
- Weinhäupl, K., Lindau, C., Hessel, A., Wang, Y., Schütze, C., Jores, T., et al. (2018). Structural Basis of Membrane Protein Chaperoning through the Mitochondrial Intermembrane Space. *Cell* 175, 1365–1379. doi:10.1016/j.cell.2018.10.039
- Weinhäupl, K., Wang, Y., Hessel, A., Brennich, M., Lindorff-Larsen, K., and Schanda, P. (2021). Architecture and Assembly Dynamics of the Essential Mitochondrial Chaperone Complex TIM9-10-12. *Structure* 29, 1–9. doi:10.1016/j.str.2021.04.009
- Wiedemann, N., Pfanner, N., and Rehling, P. (2006). Import of Precursor Proteins into Isolated Yeast Mitochondria. *Methods Mol. Biol.* 313, 373–383. doi:10.1385/1-59259-958-3:373
- Wiedemann, N., Pfanner, N., and Ryan, M. T. (2001). The Three Modules of ADP/ATP Carrier Cooperate in Receptor Recruitment and Translocation into Mitochondria. *EMBO J.* 20, 951–960. doi:10.1093/emboj/20.5.951
- Wu, S., Ge, X., Lv, Z., Zhi, Z., Chang, Z., and Zhao, X. S. (2011). Interaction between Bacterial Outer Membrane Proteins and Periplasmic Quality Control Factors: A Kinetic Partitioning Mechanism. *Biochem. J.* 438, 505–511. doi:10.1042/bj20110264
- Xu, X., Wang, S., Hu, Y.-X., and McKay, D. B. (2007). The Periplasmic Bacterial Molecular Chaperone SurA Adapts its Structure to Bind Peptides in Different Conformations to Assert a Sequence Preference for Aromatic Residues. *J. Mol. Biol.* 373, 367–381. doi:10.1016/j.jmb.2007.07.069
- Yébenes, H., Mesa, P., Muñoz, I. G., Montoya, G., and Valpuesta, J. M. (2011). Chaperonins: Two Rings for Folding. *Trends Biochem. Sci.* 36, 424–432. doi:10.1016/j.tibs.2011.05.003
- Zhang, Y., and Zuiderweg, E. R. P. (2004). The 70-kDa Heat Shock Protein Chaperone Nucleotide-Binding Domain in Solution Unveiled as a Molecular Machine that Can Reorient its Functional Subdomains. *Proc. Natl. Acad. Sci.* 101, 10272–10277. doi:10.1073/pnas.0401313101

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Sućec, Bersch and Schanda. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

RESEARCH ARTICLE

Open Access



# The mitochondrial carrier pathway transports non-canonical substrates with an odd number of transmembrane segments

Heike Rampelt<sup>1,2\*</sup> , Iva Sucec<sup>3</sup>, Beate Bersch<sup>3</sup>, Patrick Horten<sup>1,2,4</sup>, Inge Perschil<sup>1</sup>, Jean-Claude Martinou<sup>5</sup>, Martin van der Laan<sup>6</sup>, Nils Wiedemann<sup>1,2,7</sup>, Paul Schanda<sup>3</sup> and Nikolaus Pfanner<sup>1,2,7\*</sup>

## Abstract

**Background:** The mitochondrial pyruvate carrier (MPC) plays a central role in energy metabolism by transporting pyruvate across the inner mitochondrial membrane. Its heterodimeric composition and homology to SWEET and semiSWEET transporters set the MPC apart from the canonical mitochondrial carrier family (named MCF or SLC25). The import of the canonical carriers is mediated by the carrier translocase of the inner membrane (TIM22) pathway and is dependent on their structure, which features an even number of transmembrane segments and both termini in the intermembrane space. The import pathway of MPC proteins has not been elucidated. The odd number of transmembrane segments and positioning of the N-terminus in the matrix argues against an import via the TIM22 carrier pathway but favors an import via the flexible presequence pathway.

**Results:** Here, we systematically analyzed the import pathways of Mpc2 and Mpc3 and report that, contrary to an expected import via the flexible presequence pathway, yeast MPC proteins with an odd number of transmembrane segments and matrix-exposed N-terminus are imported by the carrier pathway, using the receptor Tom70, small TIM chaperones, and the TIM22 complex. The TIM9-10 complex chaperones MPC proteins through the mitochondrial intermembrane space using conserved hydrophobic motifs that are also required for the interaction with canonical carrier proteins.

**Conclusions:** The carrier pathway can import paired and non-paired transmembrane helices and translocate N-termini to either side of the mitochondrial inner membrane, revealing an unexpected versatility of the mitochondrial import pathway for non-cleavable inner membrane proteins.

**Keywords:** Mitochondrial pyruvate carrier, MPC, Mitochondrial protein biogenesis, Protein import, TIM22 complex, Tim9, Tim10, TIM23 complex

## Background

Despite its crucial role in physiology, the molecular identity of the mitochondrial pyruvate carrier (MPC) was uncovered only in recent years [1, 2]. MPC enables transport of pyruvate into the mitochondrial matrix for oxidative metabolism via pyruvate dehydrogenase and the tricarboxylic acid cycle. Due to this central position in energy metabolism, the MPC plays a crucial role in metabolic switches between glycolytic and respiratory growth and

affects cancer stemness [3–5]. The functional unit of the MPC is an inner membrane-integrated heterodimer consisting of MPC1 and MPC2 in mammals and of Mpc1 with either Mpc2 or Mpc3 in yeast [1, 2, 6, 7].

The inner mitochondrial membrane harbors a multitude of carrier proteins that belong to the mitochondrial carrier family (termed MCF or SLC25 for solute carrier family 25) and transport nucleotides, amino acids, and other metabolites across the inner membrane. These canonical, well-studied carrier proteins are characterized by three structural modules, each consisting of two transmembrane helices with a connecting matrix loop, and expose both termini of the polypeptide chain to the

\* Correspondence: [heike.rampelt@biochemie.uni-freiburg.de](mailto:heike.rampelt@biochemie.uni-freiburg.de); [nikolaus.pfanner@biochemie.uni-freiburg.de](mailto:nikolaus.pfanner@biochemie.uni-freiburg.de)

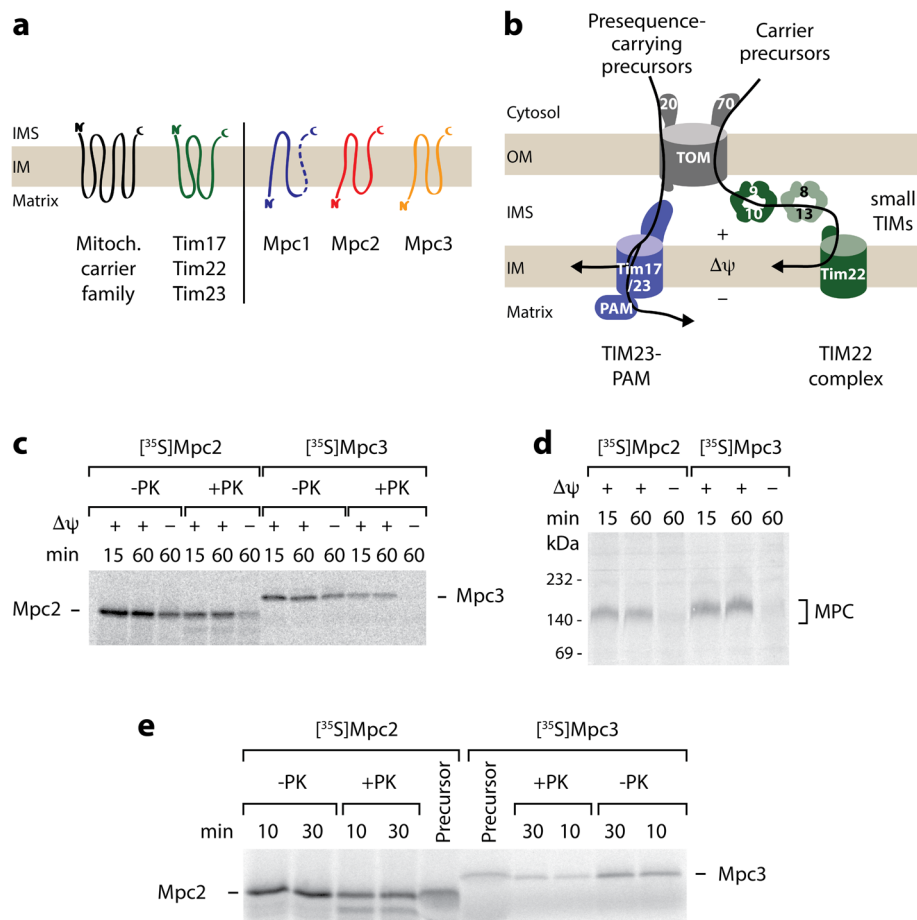
<sup>1</sup>Institute of Biochemistry and Molecular Biology, ZBMZ, Faculty of Medicine, University of Freiburg, 79104 Freiburg, Germany

Full list of author information is available at the end of the article



intermembrane space (Fig. 1a) [8–10]. MPC proteins do not belong to the established mitochondrial carrier family but are related to the SWEET (sugars will eventually be exported transporter) and semiSWEET sugar transporters that function as two triple-helix bundles [11, 12]. In contrast to the canonical carriers with six transmembrane helices, Mpc2 and Mpc3 were shown to contain three transmembrane helices with the N-terminus facing the matrix, based on the accessibility to protease or to thiol

labeling (Fig. 1a) [6, 7]. The N-terminus of Mpc1 faces the matrix; its exact number of transmembrane segments has not been defined as biochemical approaches suggested the presence of two transmembrane segments, whereas a recent homology analysis indicated that Mpc1 displays a similar topology as Mpc2 and Mpc3 [6, 7]. The active MPC complexes are heterodimers; Mpc1-Mpc3 promotes pyruvate transport during respiratory growth, whereas Mpc1-Mpc2 functions during fermentable growth [6, 7, 13].



**Fig. 1** Import of MPC precursors into the mitochondria. **a** Membrane topology of substrates of the carrier translocase TIM22 in the inner mitochondrial membrane (IM). Left, all TIM22 substrates known so far possess a uniform topology with an even number of transmembrane segments and both termini facing the intermembrane space (IMS): canonical mitochondrial carriers (black) and translocase components (green). Right, the mitochondrial pyruvate carrier subunits Mpc2 and Mpc3 possess an odd number of transmembrane segments and expose the N-terminus to the matrix. The N-terminus of Mpc1 is also located in the matrix, Mpc1 likely possesses three transmembrane segments like Mpc2/3. **b** Overview of the presequence (TIM23) pathway and the carrier (TIM22) pathway to the IM. Precursors with N-terminal presequence are recognized by the receptor Tom20, translocated through the TOM complex, and transferred to TIM23 for sorting to the IM or matrix. Precursors of the mitochondrial carrier family are recognized by the receptor Tom70, translocated through TOM, and handed over to small TIM chaperones in the IMS (TIM9-10, TIM8-13); the precursors are inserted into the IM by the TIM22 complex.  $\Delta\psi$ , membrane potential across the IM; PAM, presequence translocase-associated motor. **c** Mpc2 and Mpc3 precursors radiolabeled with [<sup>35</sup>S] methionine were imported at 25 °C into isolated yeast wild-type mitochondria for the indicated periods. Non-imported precursors were degraded with proteinase K (PK) where indicated; the mitochondria were analyzed by SDS-PAGE and autoradiography. **d** Mpc2 and Mpc3 import and assembly into a native complex is  $\Delta\psi$ -dependent. Radiolabeled Mpc2 and Mpc3 precursors were imported as in **c**; mitochondria were PK-treated and analyzed by BN-PAGE and autoradiography. **e** Mpc2 and Mpc3 are not proteolytically processed upon import into mitochondria. Mpc2 and Mpc3 were imported into mitochondria as in **c**. The reactions were analyzed by SDS-PAGE and autoradiography. For comparison, 20% of reticulocyte lysate (precursor) used in the import reactions were included

The import pathway of canonical carrier precursors from the cytosol to the carrier translocase of the inner mitochondrial membrane (TIM22) has been well established (Fig. 1b) [14–16]. After recognition of internal, non-cleavable signals by the receptor Tom70 of the translocase of the outer membrane (TOM) [17–19], carrier precursors pass through the TOM channel into the intermembrane space. There, the hydrophobic precursors are bound by small TIM chaperones and are transferred to the TIM22 complex for membrane potential ( $\Delta\psi$ )-dependent insertion into the inner membrane (Fig. 1b) [15, 20–24]. In contrast to the highly versatile presequence translocase of the inner membrane (TIM23) that handles a large variety of precursor proteins, including cleavable and non-cleavable matrix and inner membrane proteins with differing topologies, the carrier translocase TIM22 is thought to have a narrow, well-defined substrate repertoire (Fig. 1a, b). The only known physiological substrates of the TIM22 pathway are the mitochondrial carriers with 6 transmembrane segments (> 30 members in fungi and > 50 members in mammals [9]) and the translocase components Tim17, Tim22, and Tim23 with 4 transmembrane segments, all sharing the same topology with both termini facing the intermembrane space (Fig. 1a) [14, 15]. Precursors imported by the carrier pathway are bound and transported by the TOM complex in a modular fashion with pairs of transmembrane helices being translocated [14, 18, 19, 23, 25]. Binding to the small TIM chaperones also takes place in a modular fashion [26]. Mutational studies with truncated carrier precursors indicated that the cooperation of the 2-helix modules is crucial for import and assembly via the carrier pathway [19, 27–29]. Truncated carrier precursors with 4 or less transmembrane segments were even mistargeted via the TIM23 complex into the matrix or remained in the intermembrane space [28, 29]. A remarkable exception in the carrier family is Ugo1 that contains an odd number of transmembrane segments (3). Indeed, Ugo1 is not imported by the TIM22 pathway but is an integral component of the mitochondrial outer membrane [30–33]. The findings reported so far thus strongly support the model of strict substrate selectivity of the TIM22 pathway.

The biogenesis pathway of MPC proteins from their synthesis in the cytosol to their mature destination in the inner membrane has not been elucidated. The odd number of transmembrane segments and positioning of the N-terminus in the matrix argues against an import via the TIM22 carrier pathway but favors an import via the flexible presequence pathway. Here, we systematically analyzed the import pathways of Mpc2 and Mpc3 and unexpectedly observed a clear dependence on the carrier import pathway, including the receptor Tom70,

TIM chaperones, and TIM22 complex, but not on the presequence pathway. These findings substantially expand the substrate spectrum and translocation flexibility of the mitochondrial carrier pathway.

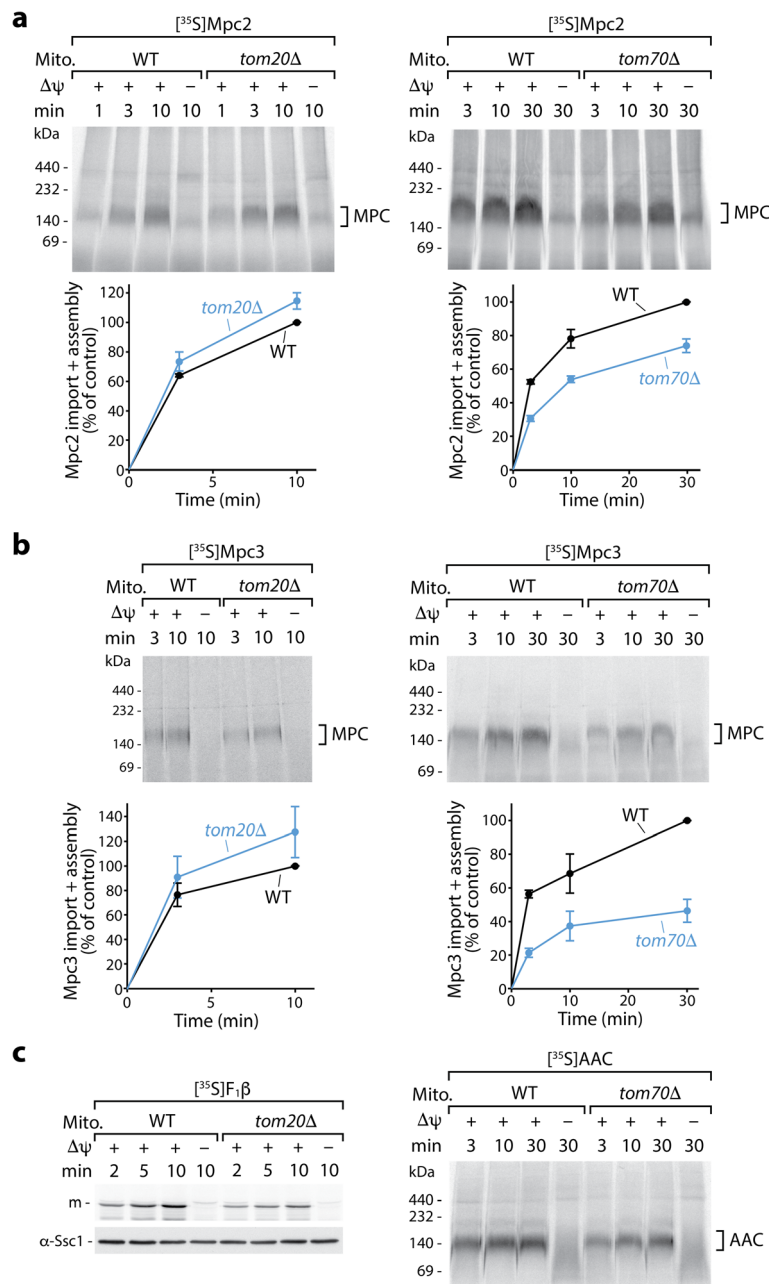
## Results

### Targeting and $\Delta\psi$ -dependent import of MPC precursors into mitochondria

We synthesized and radiolabeled the precursors of Mpc2 and Mpc3 in a cell-free system and imported them into isolated yeast wild-type mitochondria. The precursors were transported to a protease-protected location (Fig. 1c) and assembled into a complex migrating at ~ 150 kDa in blue native gel electrophoresis (Fig. 1d) like the mature assembled MPC dimers detected by Western blotting (Additional file 1: Figure S1a-d) [1, 6]. The relatively slow migration of the ~ 30 kDa MPC dimers on blue native electrophoresis is likely due to considerable amounts of lipids and detergent bound to MPC, similar to observations with other small membrane proteins [7, 26, 34, 35]. In the absence of a membrane potential  $\Delta\psi$ , the transport to a protease-protected location was impaired and the assembly into the ~ 150 kDa complex was blocked (Fig. 1c, d), demonstrating that  $\Delta\psi$  across the inner membrane was required for the biogenesis of the MPC proteins in line with the inner membrane localization of mature MPC. The strong  $\Delta\psi$  dependence of the formation of the 150 kDa MPC complex upon importing radiolabeled precursors provided an efficient assay for studying import and assembly of Mpc2 and Mpc3 *in organello*. The imported proteins (+ $\Delta\psi$ ) showed the identical SDS gel mobility as the non-imported precursors ( $-\Delta\psi$ ) and the precursors synthesized in the cell-free system (Fig. 1c, e), indicating that neither Mpc2 nor Mpc3 carried a cleavable presequence, in agreement with a systematic proteomic study that did not detect a cleavable presequence in Mpc3 (termed Fmp43 before the assignment as MPC subunit) [36].

Precursor proteins imported via the presequence pathway are typically recognized by the receptor Tom20, whereas canonical carrier precursors are recognized by Tom70 [17, 35, 37–42]. Import and assembly of Mpc2 and Mpc3 into *tom20* $\Delta$  mitochondria were not inhibited, but even slightly better than that into wild-type mitochondria, whereas import of the presequence pathway substrate  $F_1$ -ATPase subunit  $\beta$  ( $F_1\beta$ ) was inhibited in the mutant mitochondria as expected (Fig. 2a–c, Additional file 2: Figure S2a–c). When Mpc2 or Mpc3 were imported into the mitochondria lacking Tom70, however, we observed a reduction of import and assembly similar to that observed for the ADP/ATP carrier (AAC) (Fig. 2a–c, Additional file 2: Figure S2d–g). The individual TOM receptors do not





**Fig. 2** Import of Mpc2 and Mpc3 precursors occurs via the receptor Tom70, not Tom20. Radiolabeled Mpc2 (**a**) and Mpc3 (**b**) were imported at 25 °C into mitochondria from wild-type (WT), *tom20Δ*, or *tom70Δ* yeast strains and analyzed as described in Fig. 1d. **a, b** (lower panels) Quantification of import and assembly efficiency; the efficiency into WT mitochondria upon the longest import period was set to 100% (control);  $n = 3$  except Mpc2 import into *tom70Δ*:  $n = 4$ ; error bars: SEM. As controls, the matrix-targeted precursor of F<sub>1</sub>β was imported into *tom20Δ* mitochondria (**c**, left panel, with α-Ssc1 immunodecoration as a loading control), and the carrier protein AAC was imported into *tom70Δ* mitochondria (**c**, right panel). In all experiments, non-imported precursors were degraded with proteinase K. m, mature form

exclusively recognize one defined substrate class but possess an overlapping specificity [37, 43, 44]. In particular, precursors with N-terminal presequences recognized by Tom20 can contain additional internal targeting signals that interact with Tom70, and thus, these precursors employ both receptors [43–45]. The

selective dependence of Mpc2 and Mpc3 on Tom70 and not on Tom20 (Fig. 2a, b, Additional file 2: Figure S2 g), however, does not fit to the typical receptor dependence of preproteins with N-terminal targeting signals but to that of the mitochondrial carrier family MCF/SLC25.



### MPC precursors are imported via the TIM22 complex and not the TIM23 complex

To directly determine whether the TIM22 complex or the TIM23 complex is responsible for membrane insertion of Mpc2 and Mpc3, we imported the precursor proteins into mitochondria which were isolated from yeast mutants that specifically affect one of the translocases.

To date, all imported mitochondrial proteins that expose the N-terminus to the matrix have been found to be transported by the TIM23 pathway [14, 15]. Since the MPC proteins also expose their N-termini to the matrix, we analyzed the dependence on the TIM23 machinery. The yeast mutants *tim17-5* and *tim17-4* selectively impair TIM23-mediated matrix import or lateral sorting of cleavable preproteins into the inner membrane, respectively, without disturbing the inner membrane potential and the canonical carrier import [46, 47]. Import and assembly of Mpc2 and Mpc3, however, were not inhibited in *tim17-5* mitochondria after an in vitro heat shock at 37 °C (Fig. 3a, Additional file 3: Figure S3a; the corresponding wild-type mitochondria were subjected to the same heat shock conditions), whereas import of the TIM23-dependent matrix protein F<sub>1</sub>β was considerably impaired (Fig. 3b). Unexpectedly, heat-shocked *tim17-4* mitochondria, which were impaired in the inner membrane sorting of the TIM23 model substrate b<sub>2</sub>(220)-DHFR [46, 47], efficiently imported and assembled Mpc2 and Mpc3 in a Δψ-dependent manner (Fig. 3c, d; Additional file 3: Figure S3b), indicating that the MPC proteins are not imported by the presequence pathway.

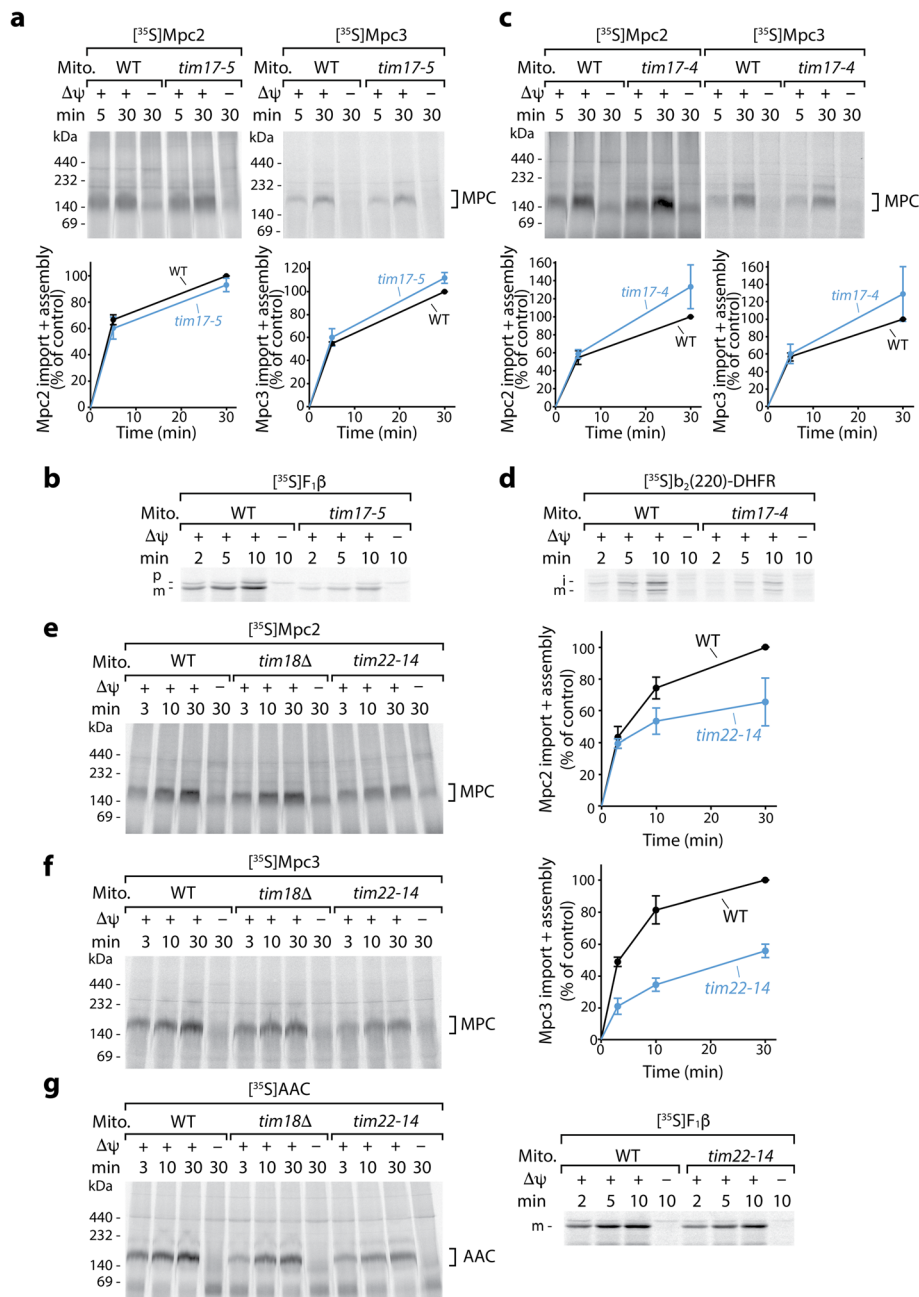
The lack of the non-essential subunit Tim18 of the TIM22 complex only mildly affected the import and assembly of Mpc2, Mpc3, and AAC (Fig. 3e–g, Additional file 3: Figure S3c) and thus did not provide an answer on the translocase dependence. Therefore, we used the yeast temperature-sensitive mutant *tim22-14* of the essential translocase subunit Tim22 at a permissive temperature [48] (Additional file 3: Figure S3d). The mutant mitochondria are disturbed in the assembly of the carrier translocase TIM22 [48]. Despite mildly reduced levels of the TIM22 substrate Tim23 (Additional file 3: Figure S3d), neither the presequence import pathway (Fig. 3g, right panel) nor the inner membrane potential is impaired [48]. Import and assembly of Mpc2 and Mpc3, however, were partially reduced in *tim22-14* mitochondria, like import and assembly of the canonical substrate AAC (Fig. 3e–g, Additional file 3: Figure S3c, e, f), supporting the view that the MPC proteins use the carrier import pathway.

Taken together, we conclude that the two MPC proteins are imported via the TIM22 pathway into the inner membrane despite their non-canonical carrier topology and their odd number of transmembrane segments.

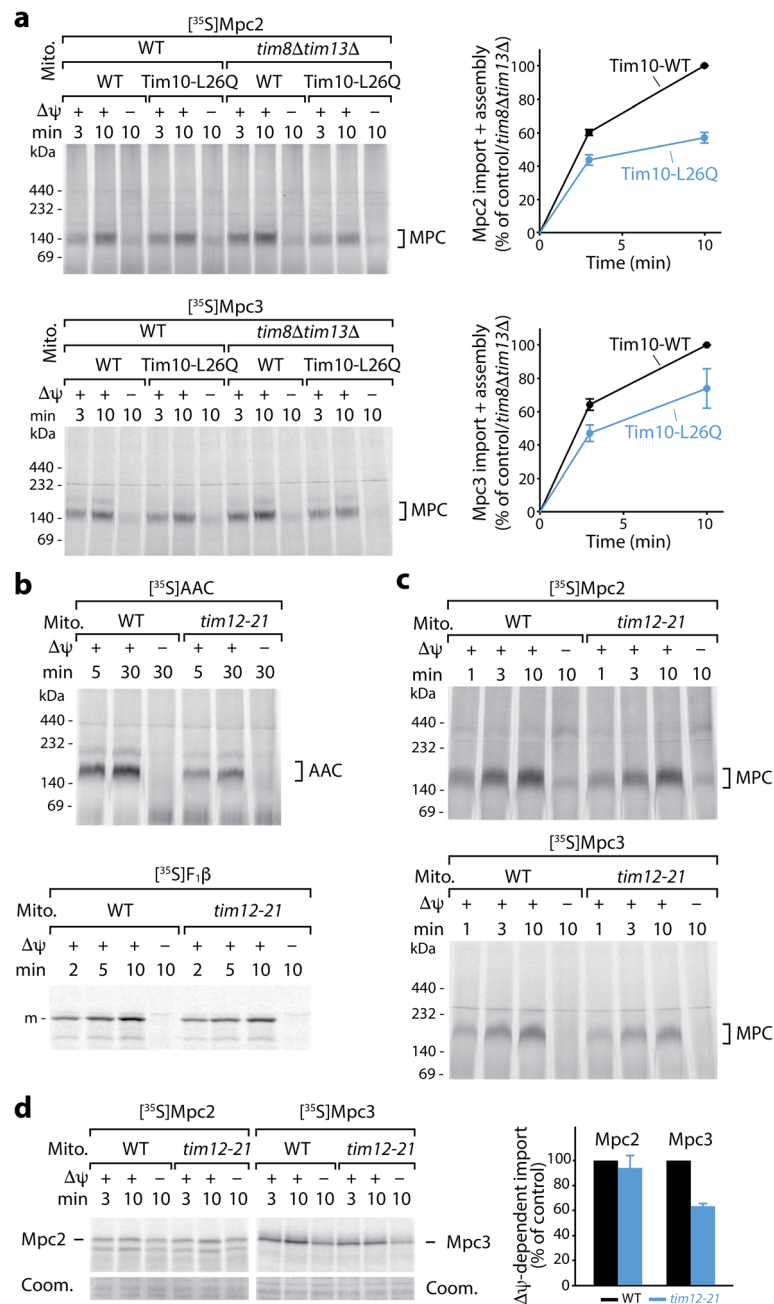
### Import of MPC precursors involves small TIM chaperones of the intermembrane space

Canonical carrier proteins with their six hydrophobic transmembrane segments rely on chaperoning by the small TIM proteins during their transit through the aqueous intermembrane space, providing a strict difference to the presequence import pathway where precursors are directly transferred from the TOM complex to the TIM23 complex [20–23, 26, 45, 46, 49–53]. Carrier precursors are preferentially bound by the essential Tim9-Tim10 complex (TIM9-10), while the alternative Tim8-Tim13 complex (TIM8-13) provides some redundancy and, together with TIM9-10, promotes the import of β-barrel precursors to the outer membrane [26, 54]. The association of carrier precursors with the TIM22 complex is accomplished via a membrane-bound module of TIM22 comprising Tim9, Tim10, and Tim12 [20, 22, 24, 55].

The model of MPC import via the canonical carrier import pathway implies that MPC precursors should depend on small TIM chaperones for crossing the intermembrane space. We thus asked if any of the TIM chaperones participated in the import of Mpc2 and Mpc3. We used a yeast mutant of the TIM9-10 complex containing an amino acid replacement in the chaperone motif of the essential Tim10 protein, resulting in a temperature-sensitive growth defect. Tim10-L26Q mutant mitochondria are delayed in the import of canonical carrier proteins and the four-transmembrane substrate Tim23 under permissive conditions, whereas Δψ and the presequence pathway are not affected [26]. Import and assembly of Mpc2 and Mpc3 into the Tim10-L26Q mitochondria at permissive temperature were reduced both in the presence and in the absence of TIM8-13, similarly to the biogenesis of AAC (Fig. 4a, Additional file 4: Figure S4a–c). The lack of TIM8-13 alone did not impede Mpc2/3 import (Fig. 4a). The steady-state levels of Mpc1 and Mpc3 were reduced in the Tim10-L26Q mutant strains, but not in the *tim8Δ-tim13Δ* strain, similarly to the levels of the canonical carrier protein Yhm2 (Additional file 4: Figure S4d). The increased levels of Mpc2 in the Tim10-L26Q mutant strains are likely due to the decreased levels of Mpc1 as the lack of Mpc1 leads to a strong induction of Mpc2 levels (Additional file 1: Figure S1a–f) [1]. A preferential dependence on the essential TIM9-10 chaperone and a backup function of the non-essential TIM8-13 chaperone are consistent with the import behavior of carrier pathway substrates like AAC and Tim23 and distinguish Mpc2/3 from the import characteristics of β-barrel precursors that typically use both TIM9-10 and TIM8-13 [26]. To address a possible requirement for inner membrane-bound Tim12, we tested the import of the MPC precursors into mitochondria from the temperature-sensitive *tim12-21* mutant, employing the



**Fig. 3** Mpc2 and Mpc3 are imported by TIM22 and are independent of TIM23. **a** Wild-type (WT) and *tim17-5* mitochondria, which display a specific defect in TIM23-mediated matrix import [46, 47], were heat-shocked for 10 min at 37 °C prior to import of radiolabeled Mpc2 or Mpc3 at 25 °C. Samples were analyzed by BN-PAGE and autoradiography. Quantification of import and assembly efficiency; the efficiency into WT mitochondria after 30 min was set to 100% (control),  $n = 3$ ; error bars: SEM. **b** As a control, the matrix protein F<sub>1</sub>β was imported into heat-shocked WT and *tim17-5* mitochondria. Samples were analyzed by SDS-PAGE and autoradiography. p, precursor; m, mature form. **c** Mpc2 and Mpc3 were imported at 25 °C into heat-shocked WT mitochondria and *tim17-4* mitochondria that display a defect in TIM23-mediated sorting into the inner membrane [46, 47]. Samples were analyzed and quantitated as in **a**;  $n = 3$ ; error bars: SEM. **d** As a control, the IM sorting substrate b<sub>2</sub>(220)-DHFR was imported into heat-shocked WT and *tim17-4* mitochondria. Samples were analyzed by SDS-PAGE and autoradiography. i, intermediate form; m, mature form. **e** Mpc2 was imported at 25 °C into mitochondria from WT and TIM22-specific yeast mutant strains, *tim18Δ* or *tim22-14*, and analyzed by BN-PAGE and autoradiography. Quantification of import and assembly efficiency as in **a**;  $n = 3$ ; error bars: SEM. **f** Mpc3 was imported at 25 °C into mitochondria from WT, *tim18Δ* and *tim22-14* strains as in **e**. Quantification of import and assembly efficiency as in **a**;  $n = 3$ ; error bars: SEM. **g** The model carrier substrate AAC was imported at 25 °C into *tim18Δ* and *tim22-14* mitochondria (left panel) and analyzed as the Mpc2/ Mpc3 import reactions. As a control, the matrix-targeted precursor of F<sub>1</sub>β was imported into these mitochondria (right panel) and analyzed by SDS-PAGE and autoradiography. m, mature form. In all experiments, non-imported precursors were degraded with PK



**Fig. 4** Mpc2 and Mpc3 import depends on small TIM chaperones. **a** Radiolabeled Mpc2 and Mpc3 were imported at 25 °C into wild-type (WT) mitochondria, mitochondria with the mutant form Tim10-L26Q, mitochondria lacking Tim8 and Tim13, or mitochondria affected in Tim10, Tim8, and Tim13 [26]. Samples were analyzed by BN-PAGE and autoradiography. Quantification of import and assembly efficiency; the efficiency into Tim10-WT/*tim8Δtim13Δ* mitochondria after 10 min was set to 100% (control); *n* = 3 for Mpc2 import, *n* = 4 for Mpc3 import; error bars: SEM. **b** AAC (upper panel) and F<sub>1</sub>β (lower panel) were imported at 30 °C into wild-type or *tim12-21* mutant mitochondria, followed by BN-PAGE (AAC) or SDS-PAGE (F<sub>1</sub>β) analysis and autoradiography. m, mature form. **c** Mpc2 (upper panel) and Mpc3 (lower panel) were imported at 30 °C into wild-type or *tim12-21* mutant mitochondria and analyzed by BN-PAGE and autoradiography. **d** Mpc2 or Mpc3 were imported at 30 °C into wild-type or *tim12-21* mutant mitochondria. Mitoplasts were generated by hypo-osmotic swelling, treated with proteinase K, and analyzed by SDS-PAGE and autoradiography (upper panel) or Coomassie Blue R-250 staining (Coom.) as a loading control (lower panel). Quantification (right panel) of membrane potential (Δψ)-dependent import yield after 10 min relative to the WT control; *n* = 3; error bars: SEM. In all experiments, non-imported precursors were degraded with proteinase K

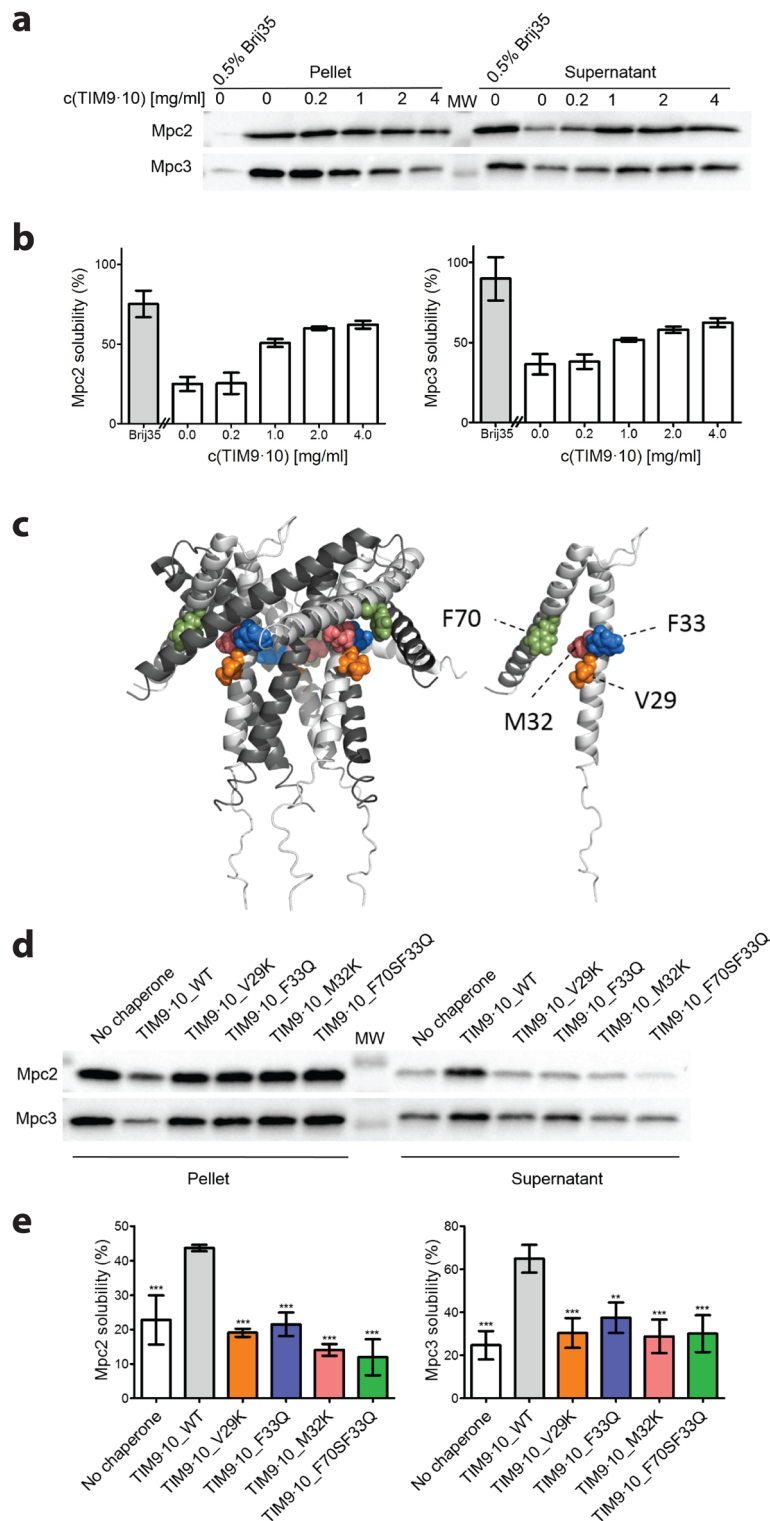
elevated temperature of 30 °C. The *tim12-21* mutant mitochondria were impaired in the carrier pathway (AAC), but not in the presequence pathway ( $F_1\beta$ ) (Fig. 4b) [55]. Import and assembly of Mpc2 were not significantly diminished in the mutant mitochondria, whereas Mpc3 was partially affected (Fig. 4c, d) and the import of AAC was more strongly reduced (Fig. 4b). These results suggest that the biogenesis of Mpc2/3 involves small TIM proteins, in particular the major soluble TIM chaperone, the TIM9-10 complex.

To directly determine if the MPC precursors depend on the chaperone function of small TIMs, we synthesized cysteine-free forms of the precursors in a cell-free translation system [56] and performed an aggregation assay. The majority of the hydrophobic Mpc2 and Mpc3 precursors aggregated in the cell-free system in the absence of detergent (Fig. 5a, b). Weinhäupl et al. [26] showed that the TIM9-10 chaperone prevented the aggregation of a canonical carrier precursor in vitro. We thus added recombinantly produced TIM9-10 and observed a significant improvement of the solubility of Mpc2 and Mpc3 in a concentration-dependent manner (Fig. 5a, b). Importantly, the positive effect of TIM9-10 on the solubility of MPC precursors was abrogated with Tim10 point mutants in which hydrophobic residues in the binding cleft were replaced by hydrophilic ones (Fig. 5c–e). These mutant forms also disrupt the interaction with carrier precursors [26], suggesting that MPC precursors bind to the same hydrophobic motif of the chaperone as carriers. In addition, we studied the influence of TIM9-10 on Mpc1, whose topology has not been fully clarified but based on a recent homology analysis is likely similar to Mpc2/3, including the lack of a cleavable presequence [6, 7, 36]. We observed a similar prevention of aggregation and dependence on specific Tim10 residues for Mpc1 as for Mpc2 and Mpc3 (Additional file 5: Figure S5a, b). The levels of Mpc1 are considerably reduced in *tim22-14* mitochondria and partially reduced in *tim12-21* and Tim10-L26Q mitochondria (Additional file 3: Figure S3d, Additional file 4: S4d), suggesting that the biogenesis of Mpc1 occurs via the carrier import pathway. Since Mpc1 levels are stable in the absence of Mpc2 and/or Mpc3 (Additional file 1: Figure S1e, f), the observed decrease in *tim22-14*, *tim12-21*, and Tim10-L26Q mitochondria likely reflects a defect in Mpc1 biogenesis rather than an indirect destabilization. In line with our *in organello* import results, the TIM8-13 complex only mildly improved the solubility of MPC precursors (Additional file 5: Figure S5c). We conclude that the TIM9-10 complex chaperones all MPC precursors. Interaction of TIM9-10 with the MPC proteins is mediated by the same conserved Tim10 motifs that are responsible for the chaperone activity toward established substrates [26].

## Discussion

The mitochondrial pyruvate carrier differs substantially from the well-characterized family of mitochondrial carriers, by both its topology and its heterodimeric composition. In particular, all three MPC proteins have their N-termini in the matrix, and for Mpc2 and Mpc3, the presence of three transmembrane helices has been established [6, 7]. Proteins with this topology have been expected to be imported by the TIM23 pathway [14, 15]. In contrast, our results demonstrate that MPC subunits are imported into the inner mitochondrial membrane by the carrier pathway, using all of its characteristic components. They are recognized on the mitochondrial surface by the receptor Tom70, are chaperoned through the intermembrane space by the TIM9-10 complex, and are inserted into the inner membrane by the carrier translocase TIM22. This surprising finding strongly changes the view of the substrate selection of this major transport pathway to the mitochondrial inner membrane.

All studies available so far supported the model that the carrier pathway can only handle pairs of transmembrane helices with their termini in the intermembrane space [18, 19, 23, 25, 26]. Different precursor forms such as truncated carrier precursors or the three-helix Ugo1 are either imported by the highly flexible TIM23 presequence pathway (bypassing the small TIMs), remain in the intermembrane space, or are even directed to the mitochondrial outer membrane [28–33]. The basic requirements of proteins imported by the carrier pathway include paired transmembrane helices with a defined topology, positively charged matrix-exposed segments and the ability to interact with the small TIM chaperones [14, 15, 19, 25–29, 57]. The MPC proteins display a fundamentally different topology but are able to interact with the TIM chaperones, and their matrix-exposed N-termini and loops (between transmembrane helices 2 and 3) are positively charged [1, 2, 7]. The two C-terminal transmembrane helices of Mpc2 and Mpc3 may be handled by the TIM22 machinery similarly to a paired helix of a canonical carrier. The N-terminus of MPCs was suggested to form an amphipathic helix whose function is unknown [7]. As observed for mitochondrial matrix and inner membrane proteins, the matrix-exposed positively charged amino acid residues are likely involved in the translocation of preprotein segments across the inner membrane by responding to  $\Delta\psi$  (negative on the matrix side) [24, 25, 57–59]. For the interaction with TIM chaperones, the same residues in the hydrophobic substrate-binding cleft of the TIM9-10 complex are required for the interaction with both types of substrates, MPC precursors and canonical carriers [26], providing strong evidence that the MPCs are bona fide substrates of the carrier import pathway.



**Fig. 5** (See legend on next page.)



(See figure on previous page.)

**Fig. 5** Interaction of Mpc2 and Mpc3 with the TIM9-10 chaperone in vitro. **a** Cell-free reaction mixtures producing Mpc2 (upper panel) or Mpc3 (lower panel) were supplemented with detergent (Brij35) or different concentrations of recombinantly produced TIM9-10 complex. Immunoblot of the soluble (supernatant) and insoluble (pellet) fractions of the reaction mixtures. **b** Mpc2 and Mpc3 solubility quantification. In the presence of detergent (absence of TIM9-10), both Mpc2 and Mpc3 were largely found in the soluble fraction. In the absence of detergent and chaperone, the majority of Mpc2 and Mpc3 was found in the insoluble fraction. Increasing the concentration of TIM9-10 complex in the cell-free reaction mixture resulted in increased solubility of Mpc2 and Mpc3;  $n = 4-5$  for Mpc2;  $n = 3$  for Mpc3; error bars indicate standard deviation. **c** Structural view of the TIM9-10 complex [26, 68]. In the chaperone complex (left), Tim9 monomers are shown in dark gray and Tim10 in light gray. Altered amino acids of the mutant variants in the TIM9-10 complex [26] are shown as colored spheres. Tim10 monomer (right) and altered amino acids in the hydrophobic cleft of TIM9-10. **d** Immunoblot of the soluble and insoluble fractions of the cell-free reaction mixtures producing Mpc2 or Mpc3 in the absence of TIM chaperones or in the presence of wild-type TIM9-10 (TIM9-10\_WT) or mutant variants of Tim10 in the TIM9-10 complex (TIM9-10\_V29K, TIM9-10\_F33Q, TIM9-10\_M32K, TIM9-10\_F70SF33Q). **e** Solubility quantification shows solubility of Mpc2 and Mpc3 in the presence of TIM9-10 mutant variants comparable to the reaction condition without added chaperone complex.  $n = 3$ ; error bars indicate standard deviation; \*\*\* and \*\* indicate the significant difference with  $P < 0.001$  and  $P < 0.005$ , respectively, in comparison with the reaction with the WT chaperone

## Conclusions

We conclude that the mitochondrial carrier pathway possesses a much higher flexibility than anticipated and can transport transmembrane helices in a paired or non-paired fashion and direct the precursor N-termini into the intermembrane space (canonical carriers, Tim17/22/23) or matrix (MPC proteins). Due to their high conservation, we expect that human MPC subunits [1, 2] are similarly imported into mitochondria via the carrier translocase pathway. These findings represent a striking example that the search for non-canonical substrates can change even long-established views of an essential protein translocation pathway.

## Material and methods

### Yeast strains and growth

The *Saccharomyces cerevisiae* strains used in this study are summarized in Table 1. The strains *tom20Δ*, *tom70Δ*, *tim18Δ*, *tim22-14*, *tim12-21*, *tim17-4*, *tim17-5*, Tim10-L26Q, *tim8Δ tim13Δ*, Tim10-L26Q *tim8Δ tim13Δ*, *mpc2Δ mpc3Δ*, and *mpc1Δ mpc2Δ mpc3Δ* and their corresponding wild types were described [6, 26, 46, 48, 55, 60–62]. Deletion strains *mpc1Δ*, *mpc2Δ*, and *mpc3Δ* and the corresponding BY4741 wild-type strain were obtained from Euroscarf. Cells for mitochondrial import experiments were grown on YPG media (1% [w/v] yeast extract, 2% [w/v] peptone, 3% [v/v] glycerol) or on YPLac media (1% [w/v] yeast extract, 2% [w/v] peptone, 3% [v/v] glycerol, 0.05% [w/v] CaCl<sub>2</sub>, 0.06% [w/v] MgCl<sub>2</sub>, 0.1% [w/v] KH<sub>2</sub>PO<sub>4</sub>, 0.1% [w/v] NH<sub>4</sub>Cl, 0.05% [w/v] NaCl, 0.05% [w/v] glucose, 2% [v/v] lactate). For the analysis of mitochondrial protein and complex levels in MPC deletion strains, cells were grown on YPG media or on YPD media (1% [w/v] yeast extract, 2% [w/v] peptone, 2% [w/v] glucose). The growth temperature was 30 °C except for the following strains: Tim10-L26Q, *tim8Δ tim13Δ*, Tim10-L26Q *tim8Δ tim13Δ*, and the corresponding wild-type strain were grown at 21 °C; *tim12-21*, *tim17-4*, *tom20Δ*, and the corresponding wild-type

strains were grown at 24 °C; and *tim17-5* and the corresponding wild-type strain were grown at 23 °C.

### Isolation of mitochondria

Mitochondria were isolated by fractionation [64]. After pre-treatment with DTT buffer (100 mM Tris-H<sub>2</sub>SO<sub>4</sub> pH 9.4, 10 mM DTT) and digestion of the cell wall with zymolyase in zymolyase buffer (20 mM potassium phosphate buffer pH 7.4, 1.2 M sorbitol), the cells were lysed on ice in homogenization buffer (10 mM Tris-HCl pH 7.4, 0.6 M sorbitol, 1 mM EDTA, 0.2% bovine serum albumin, 1 mM phenylmethylsulfonyl fluoride (PMSF)) with a glass Teflon homogenizer. After two centrifugation steps at 2000×g to remove the cell debris and nuclei, crude mitochondria were isolated from the supernatant by centrifugation at 17,000×g. Mitochondria were resuspended in SEM buffer (250 mM sucrose, 1 mM EDTA, 10 mM MOPS-KOH pH 7.2) and stored at – 80 °C.

### In organello import

In vitro synthesis of [<sup>35</sup>S]methionine-labeled precursor proteins was performed with the mMessage mMachine SP6 transcription kit (Ambion, Cat.# 1340) and the Flexi rabbit reticulocyte translation kit (Promega, Cat. # L4540), or with the TNT SP6 coupled reticulocyte transcription/translation kit (Promega, Cat. # L2080). The following plasmids were used as templates: pGEM4Z-AAC (*Neurospora crassa*), pGEM-F1β (*S. cerevisiae*), pGEM4Z-b<sub>2</sub>(220)-DHFR, pGEM4Z-Mpc1, pGEM4Z-Mpc2, and pGEM4Z-Mpc3. The radiolabeled precursors were imported into the isolated mitochondria at 25 °C in import buffer (10 mM MOPS-KOH pH 7.2, 3% [w/v] bovine serum albumin, 250 mM sucrose, 80 mM KCl, 5 mM MgCl<sub>2</sub>, 2 mM KH<sub>2</sub>PO<sub>4</sub>, 5 mM methionine) with 2–4 mM NADH and an ATP-regenerating system including 2–4 mM ATP, 5–10 mM creatine phosphate, and 0.1–0.2 mg/ml creatine kinase. Import reactions into *tim12-21* and the control wild-type mitochondria were performed at 30 °C. *tim17-4* mitochondria and *tim17-5*

**Table 1** *S. cerevisiae* strains used in this study

Strain (lab ID no.)	Genotype	Reference
RL285-16C (SHY WT) (4928)	<i>MATa his3Δ1 ura3Δ0</i>	[6]
<i>mpc1Δ</i> (SHY9) (4929)	<i>MATa his3Δ1 ura3Δ0 mpc1::kanMX</i>	[6]
<i>mpc2Δmpc3Δ</i> (SHY14) (4932)	<i>MATa his3Δ1 ura3Δ0 mpc2::natMX mpc3::hphMX</i>	[6]
<i>mpc1Δmpc2Δmpc3Δ</i> (SHY15) (4933)	<i>MATa his3Δ1 ura3Δ0 mpc1::kanMX mpc2::natMX mpc3::hphMX</i>	[6]
YPH499 (WT) (1501)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1</i>	[63]
<i>tom20Δ</i> (1273)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tom20::URA3 pYEP-TOM22</i>	[62]
<i>tom70Δ</i> (1183)	<i>ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tom70::HIS3</i>	[60, 61]
<i>tim18Δ</i> (1383)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tim18::kanMX6</i>	[48]
<i>tim22-14</i> (1370) (YPH499 22-M4)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tim22-M4</i> (amino acid alterations in Tim22-14: I11M, K16R, E21K, G31R, N37D, F63L, A85T, T86A, K120R, C141S, Y153H, M193 T, K194Q)	[48], this study
<i>tim12-21</i> (2462) (YPH-BG-12-1)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tim12::ADE2 pFL39-TIM12-1ts</i> (amino acid alterations in Tim12-21: S7G, V14D, A22E, D64A)	[55]
<i>tim17-4</i> (1758) (YPH-BG17-9d)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 BG17-9d (tim17-4)</i> (amino acid alteration in Tim17-4: C10R)	[46, 47], this study
<i>tim17-5</i> (1759) (YPH-BG17-21-7)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 BG17-21-7 (tim17-5)</i> (amino acid alterations in Tim17-5: P42L, R109G, S115P)	[46, 47], this study
WT for Tim10 mutants (5118)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tim10::ADE2 pFL39-TIM10</i>	[26]
Tim10-L26Q (5210)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tim10::ADE2 pFL39-TIM10-L26Q</i>	[26]
<i>tim8Δtim13Δ</i> (5084)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tim8::natNT2 tim13::hphNT1 tim10::ADE2 pFL39-TIM10</i>	[26]
Tim10-L26Q <i>tim8Δtim13Δ</i> (5206)	<i>MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1 tim8::natNT2 tim13::hphNT1 tim10::ADE2 pFL39-TIM10-L26Q</i>	[26]
BY4741 (WT) (1354)	<i>MATa ura3Δ0 his3Δ1 leu2Δ0 met15Δ0</i>	Euroscarf
<i>mpc1Δ</i> (4774)	<i>MATa ura3Δ0 his3Δ1 leu2Δ0 met15Δ0 mpc1::kanMX4</i>	Euroscarf
<i>mpc2Δ</i> (4775)	<i>MATa ura3Δ0 his3Δ1 leu2Δ0 met15Δ0 mpc2::kanMX4</i>	Euroscarf
<i>mpc3Δ</i> (4776)	<i>MATa ura3Δ0 his3Δ1 leu2Δ0 met15Δ0 mpc3::kanMX4</i>	Euroscarf

mitochondria and the corresponding wild-type mitochondria were heat-shocked for 10 min at 37 °C in import buffer prior to the addition of NADH, the ATP-regenerating system, and the radiolabeled precursor proteins (in reticulocyte lysate), followed by the import reaction at 25 °C. Reactions included a control sample where the membrane potential was dissipated with AVO mix (8 μM antimycin A, 1 μM valinomycin, 20 μM oligomycin) before the addition of precursor. The import reactions were stopped by the addition of AVO and transfer on ice. Non-imported precursor was removed by a 15-min incubation with 50 μg/ml proteinase K on ice, unless indicated otherwise. After the inactivation of proteinase K with 2 mM PMSF, the mitochondria were reisolated and washed in SEM buffer. To generate mitoplasts after the import reaction, the mitochondria were resuspended in hypotonic EM buffer (1 mM EDTA, 10 mM MOPS-KOH pH 7.2). The mitoplasts were treated with 50 μg/ml proteinase K and subsequently treated with PMSF and re-isolated as described above. Quantification of import/assembly efficiency was performed with

Fiji ImageJ software. Replicates used for quantification were independent import and assembly assays of incubation of isolated yeast mitochondria (wild-type and mutant mitochondria) with radiolabeled precursor proteins, followed by independent gel separation and analysis. The individual data values from independent replicates are listed in Additional file 6: Table S1 and Additional file 7: Table S2.

#### Gel electrophoresis and Western blotting

Import reactions were analyzed by SDS-PAGE or blue native gel electrophoresis (BN-PAGE) and autoradiography. For BN-PAGE analysis [65], mitochondria were solubilized in solubilization buffer (20 mM Tris-HCl pH 7.4, 50 mM NaCl, 0.1 mM EDTA, 10% [v/v] glycerol, 1% [w/v] digitonin, 1 mM PMSF) or in low-ionic strength buffer (50 mM imidazole-HCl pH 7.0, 500 mM 6-aminohexanoic acid, 1 mM EDTA, 3% [w/v] digitonin, 1 mM PMSF) [66] for 15 min on ice. Analysis of protein levels and native protein complexes was performed by SDS-PAGE or BN-PAGE,

respectively, followed by Western blot analysis. The following rabbit antisera were used (source: Pfanner Lab, non-commercial antisera specifically prepared for the lab):  $\alpha$ -Mpc1 (GR5021-1, 1:100),  $\alpha$ -Mpc2 (GR5024-4, affinity purified, 1:100),  $\alpha$ -Mpc3 (GR5025-5, affinity purified, 1:100),  $\alpha$ -Tim22 (GR5113-4, 1:250),  $\alpha$ -Tim54 (GR2012-3, 1:200),  $\alpha$ -Tim18 (GR5114-3, 1:250),  $\alpha$ -Tim12 (GR905-1, 1:500),  $\alpha$ -Yhm2 (GR3053-5, 1:500),  $\alpha$ -Ssc1 (GR1830-7, 1:250),  $\alpha$ -Tom70 (GR657-5, 1:500),  $\alpha$ -Tom40 (168-12/5, 1:500),  $\alpha$ -Tom20 (GR3225-7, 1:5000),  $\alpha$ -Tim23 (133-6, 1:500),  $\alpha$ -Tim17 (GR1844-4, 1:500),  $\alpha$ -Cor1 (GR371-6, 1:500),  $\alpha$ -Tim13 (GR2044-5, 1:500),  $\alpha$ -Tim10 (GR2041-7, 1:250), and  $\alpha$ -Atp4 (GR1958-4, 1:500).  $\alpha$ -rabbit IgG-peroxidase was obtained from Sigma-Aldrich (A6154, 1:5000–1:10,000).

#### Cell-free expression of MPC proteins in the absence or presence of TIM chaperones

Genes coding for *S. cerevisiae* Mpc1(C87A), Mpc2(C86A, C111S), and Mpc3(C87A) were cloned by GeneCust in customized pIVEX2.3d cell-free expression plasmids between NdeI and XhoI cloning sites. Cysteine-free variants were used since previous studies with the chaperoning assay [26] indicated that the presence of Cys residues can lead to enhanced aggregation, likely due to disulfide formation. The plasmid codes for the TEV-protease-cleavable N-terminal His<sub>6</sub>-tag, and it includes the stop codon before the C-terminal His<sub>6</sub>-tag of the original plasmid. The produced MPC proteins contain a cleavable His<sub>6</sub>-tag at the N-terminus.

MPC proteins were produced in 50  $\mu$ l cell-free reaction mixtures [67] for 2.5 h at 28 °C. The final composition of the cell-free reaction buffer was 0.08 mM rUTP, 0.08 mM rGTP, 0.08 mM rCTP, 0.55 mM HEPES, 0.12 mM ATP, 6.8  $\mu$ M folinic acid, 0.064 mM cyclic AMP, 0.34 mM DTT, 2.75 mM NH<sub>4</sub>OAc, 80 mM phosphocreatine, 0.208 M potassium glutamate, 10.48 mM magnesium acetate, 1 mM of amino acid mix, 1.25  $\mu$ g creatine kinase, 0.25  $\mu$ g T7 polymerase, 20  $\mu$ l S30 *E. coli* extract, 0.5  $\mu$ g plasmid DNA, and 0.175 mg/ml tRNAs. The reaction condition with the detergent contained additionally 0.5% of Brij35. To test the specificity of the binding of MPC proteins by TIM chaperones, the solubility of MPC proteins was monitored at increasing concentration of either TIM8-13 or TIM9-10 complexes. The concentrations of the chaperones in the reaction mixtures were 0, 0.2, 1, 2, and 4 mg/ml. To test the effect of selected Tim10 mutant variants in the TIM9-10 chaperone complex on the binding and subsequently the solubility of MPC proteins, 4 mg/ml of the TIM9-10<sub>WT</sub>, TIM9-10<sub>V29K</sub>, TIM9-10<sub>F33Q</sub>, TIM9-10<sub>M32K</sub>, and TIM9-10<sub>F70SF33Q</sub> were used. Chaperone complexes of TIM8-13, TIM9-10, and mutant variants of TIM9-10 for

cell-free experiments were expressed and purified as described previously [26].

The cell-free reaction was stopped after 2.5 h, and the soluble fraction was separated from the insoluble pellet by centrifugation at 16.800 $\times$ g. The amount of His-tagged MPC proteins in the soluble fraction and the pellet were quantified from the membranes, after the immunodecoration with anti-His antibody (Sigma-Aldrich monoclonal  $\alpha$ -polyHistidine-peroxidase antibody; product no: A7058), as relative band intensities using BioRad ImageLab program/software. The solubility of the proteins was calculated from at least three experiments for each condition, as a percentage of protein in the supernatant in relation to the total amount of protein in the pellet and supernatant. Significance of the difference in solubility between wild-type TIM9-10 and the mutant variants was analyzed with GraphPad Prism 5 using one-way ANOVA and Tukey's multiple comparison test. The individual data values from independent replicates are listed in Additional file 6: Table S1 and in Additional file 7: Table S2.

#### Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12915-019-0733-6>.

**Additional file 1: Figure S1.** Assembly and level of MPC subunits. (PDF)

**Additional file 2: Figure S2.** Characterization of mitochondria lacking Tom20 or Tom70. (PDF)

**Additional file 3: Figure S3.** Characterization of mitochondria affected in TIM23 or TIM22 translocases. (PDF)

**Additional file 4: Figure S4.** Characterization of mitochondria affected in small TIM chaperones. (PDF)

**Additional file 5: Figure S5.** Interaction of Mpc1, Mpc2 and Mpc3 with TIM chaperones in vitro. (PDF)

**Additional file 6: Table S1.** Individual data values for quantifications in main figures. (XSLX)

**Additional file 7: Table S2.** Individual data values for quantifications in supplementary figures. (XSLX)

#### Acknowledgements

We thank Thomas Becker, Tom Bender, Katrin Brandner, Agnieszka Chacinska, Bernard Guiard, Sébastien Herzig, Caroline Lindau, Christoph U. Mårtensson, Chantal Priesnitz, and Lena-Sophie Wenz for materials and discussion. Work included in this study has also been performed in partial fulfillment of the requirements for the doctoral thesis of PH at the University of Freiburg. This work used the cell-free protein expression facility at the Grenoble INSTRUCT-ERIC 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. We thank Lionel Imbert for the excellent support.

#### Authors' contributions

HR was responsible for conception and design, acquisition of the data, analysis and interpretation of the data, and drafting of the article. IS and BB were responsible for acquisition of the data, analysis and interpretation of the data, and drafting of the article. PH and IP were responsible for acquisition of the data, analysis and interpretation of the data. JCM was responsible for analysis and interpretation of the data. MvdL, NW, PS, and NP were responsible for conception and design, analysis and interpretation of



the data, and drafting of the article. All authors read and approved the final manuscript.

#### Funding

This study was supported by the Deutsche Forschungsgemeinschaft grants PF 202/8-1, PF 202/9-1 (NP), Collaborative Research Centers SFB 746 (MvdL, NP), SFB 894 (MvdL), SFB 1381 Project-ID 403222702 (NW), Germany's Excellence Strategy/Initiative CIBSS-EXC-2189 (Project ID 390939984), EXC 294 BIOS, GSC-4 Spemann Graduate School (NP, NW, MvdL), the European Research Council Starting Grant No. 311318 (PS), the European Research Council Consolidator Grant No. 648235 (NW), the Agence Nationale de la Recherche and the Deutsche Forschungsgemeinschaft binational project MitoMemProtImp (ANR-18-CE92-0032) (WI 4506/1-1) (PS, NW), a Peter und Traudl Engelhorn Stiftung Postdoctoral Fellowship (HR), and the Müller-Fahrenberg-Stiftung (HR). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. The article processing charge was funded by the German Research Foundation (DFG) and the University of Freiburg in the funding programme Open Access Publishing.

#### Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary data (Additional files 1, 2, 3, 4, 5, 6, and 7).

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that no competing interests exist.

#### Author details

<sup>1</sup>Institute of Biochemistry and Molecular Biology, ZBMZ, Faculty of Medicine, University of Freiburg, 79104 Freiburg, Germany. <sup>2</sup>CIBSS Centre for Integrative Biological Signalling Studies, University of Freiburg, 79104 Freiburg, Germany. <sup>3</sup>Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, CEA, CNRS, 38000 Grenoble, France. <sup>4</sup>Faculty of Biology, University of Freiburg, 79104 Freiburg, Germany. <sup>5</sup>Department of Cell Biology, University of Geneva, Genève 4, Switzerland. <sup>6</sup>Medical Biochemistry and Molecular Biology, Center for Molecular Signaling, PZMS, Saarland University, 66421 Homburg, Germany. <sup>7</sup>BIOS Centre for Biological Signalling Studies, University of Freiburg, 79104 Freiburg, Germany.

Received: 1 July 2019 Accepted: 17 December 2019

Published online: 06 January 2020

#### References

- Bricker DK, Taylor EB, Schell JC, Orsak T, Boutron A, Chen Y-C, et al. A mitochondrial pyruvate carrier required for pyruvate uptake in yeast, *Drosophila*, and humans. *Science*. 2012;337:96–100.
- Herzig S, Raemy E, Montessuit S, Veuthey J-L, Zamboni N, Westermann B, et al. Identification and functional expression of the mitochondrial pyruvate carrier. *Science*. 2012;337:93–6.
- Schell JC, Olson KA, Jiang L, Hawkins AJ, Van Vranken JG, Xie J, et al. A role for the mitochondrial pyruvate carrier as a repressor of the Warburg effect and colon cancer cell growth. *Mol Cell*. 2014;56:400–13.
- Vanderperre B, Bender T, Kunji ERS, Martinou J-C. Mitochondrial pyruvate import and its effects on homeostasis. *Curr Opin Cell Biol*. 2015;33:35–41.
- Rauchhorst AJ, Taylor EB. Mitochondrial pyruvate carrier function and cancer metabolism. *Curr Opin Genet Dev*. 2016;38:102–9.
- Bender T, Pena G, Martinou J-C. Regulation of mitochondrial pyruvate uptake by alternative pyruvate carrier complexes. *EMBO J*. 2015;34:911–24.
- Tavoulari S, Thangaratnarajah C, Mavridou V, Harbour ME, Martinou J-C, Kunji ERS. The yeast mitochondrial pyruvate carrier is a hetero-dimer in its functional state. *EMBO J*. 2019;38:e100785.
- Pebay-Peyroula E, Dahout-Gonzalez C, Kahn R, Trézéguet V, Lauquin GJ-M, Brandolin G. Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature*. 2003;426:39–44.
- Palmieri F, Monné M. Discoveries, metabolic roles and diseases of mitochondrial carriers: a review. *Biochim Biophys Acta*. 1863;2016:2362–78.
- Ruprecht JJ, King MS, Zögg T, Aleksandrova AA, Pardon E, Crichton PG, et al. The molecular mechanism of transport by the mitochondrial ADP/ATP carrier. *Cell*. 2019;176:435–447.e15.
- Xu Y, Tao Y, Cheung LS, Fan C, Chen L-Q, Xu S, et al. Structures of bacterial homologues of SWEET transporters in two distinct conformations. *Nature*. 2014;515:448–52.
- Feng L, Frommer WB. Structure and function of SemiSWEET and SWEET sugar transporters. *Trends Biochem Sci*. 2015;40:480–6.
- Timón-Gómez A, Proft M, Pascual-Ahuir A. Differential regulation of mitochondrial pyruvate carrier genes modulates respiratory capacity and stress tolerance in yeast. *PLoS One*. 2013;8:e79405.
- Neupert W. A perspective on transport of proteins into mitochondria: a myriad of open questions. *J Mol Biol*. 2015;427:1135–58.
- Wiedemann N, Pfanner N. Mitochondrial machineries for protein import and assembly. *Annu Rev Biochem*. 2017;86:685–714.
- Hansen KG, Herrmann JM. Transport of proteins into mitochondria. *Protein J*. 2019;77:1496–13.
- Brix J, Dietmeier K, Pfanner N. Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. *J Biol Chem*. 1997;272:20730–5.
- Endres M, Neupert W, Brunner M. Transport of the ADP/ATP carrier of mitochondria from the TOM complex to the TIM22.54 complex. *EMBO J*. 1999;18:3214–21.
- Wiedemann N, Pfanner N, Ryan MT. The three modules of ADP/ATP carrier cooperate in receptor recruitment and translocation into mitochondria. *EMBO J*. 2001;20:951–60.
- Koehler CM, Jarosch E, Tokatlidis K, Schmid K, Schweyen RJ, Schatz G. Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. *Science*. 1998;279:369–73.
- Koehler CM, Merchant S, Oppliger W, Schmid K, Jarosch E, Dolfini L, et al. Tim9p, an essential partner subunit of Tim10p for the import of mitochondrial carrier proteins. *EMBO J*. 1998;17:6477–86.
- Sirrenberg C, Endres M, Fölsch H, Stuart RA, Neupert W, Brunner M. Carrier protein import into mitochondria mediated by the intermembrane proteins Tim10/Mrs11 and Tim12/Mrs5. *Nature*. 1998;391:912–5.
- Curran SP, Leuenberger D, Schmidt E, Koehler CM. The role of the Tim8p–Tim13p complex in a conserved import pathway for mitochondrial polytopic inner membrane proteins. *J Cell Biol*. 2002;158:1017–27.
- Rehling P, Model K, Brandner K, Kovermann P, Sickmann A, Meyer HE, et al. Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. *Science*. 2003;299:1747–51.
- Davis AJ, Ryan KR, Jensen RE. Tim23p contains separate and distinct signals for targeting to mitochondria and insertion into the inner membrane. *Mol Biol Cell*. 1998;9:2577–93.
- Weinhäupl K, Lindau C, Hessel A, Wang Y, Schütze C, Jores T, et al. Structural basis of membrane protein chaperoning through the mitochondrial intermembrane space. *Cell*. 2018;175:1365–1379.e25.
- Brandner K, Rehling P, Truscott KN. The carboxyl-terminal third of the dicarboxylate carrier is crucial for productive association with the inner membrane twin-pore translocase. *J Biol Chem*. 2005;280:6215–21.
- Vergnolle MAS, Sawney H, Junne T, Dolfini L, Tokatlidis K. A cryptic matrix targeting signal of the yeast ADP/ATP carrier normally inserted by the TIM22 complex is recognized by the TIM23 machinery. *Biochem J*. 2005;385:173–80.
- Yamano K, Ishikawa D, Esaki M, Endo T. The phosphate carrier has an ability to be sorted to either the TIM22 pathway or the TIM23 pathway for its import into yeast mitochondria. *J Biol Chem*. 2005;280:10011–7.
- Coonrod EM, Karren MA, Shaw JM. Ugo1p is a multipass transmembrane protein with a single carrier domain required for mitochondrial fusion. *Traffic*. 2007;8:500–11.
- Hoppins SC, Horner J, Song C, McCaffery JM, Nunnari J. Mitochondrial outer and inner membrane fusion requires a modified carrier protein. *J Cell Biol*. 2009;184:569–81.
- Becker T, Wenz L-S, Kruger V, Lehmann W, Müller JM, Goroncy L, et al. The mitochondrial import protein Mim1 promotes biogenesis of multispanning outer membrane proteins. *J Cell Biol*. 2011;194:387–95.
- Papic D, Krumpe K, Dukanovic J, Dimmer KS, Rapaport D. Multispanning mitochondrial outer membrane protein Ugo1 follows a unique Mim1-dependent import pathway. *J Cell Biol*. 2011;194:397–405.

34. Crichton PG, Harding M, Ruprecht JJ, Lee Y, Kunji ERS. Lipid, detergent, and Coomassie Blue G-250 affect the migration of small membrane proteins in blue native gels: mitochondrial carriers migrate as monomers not dimers. *J Biol Chem*. 2013;288:22163–73.
35. Ellenrieder L, Dieterle MP, Doan KN, Mårtensson CU, Floerchinger A, Campo ML, et al. Dual role of mitochondrial porin in metabolite transport across the outer membrane and protein transfer to the inner membrane. *Mol Cell*. 2019;73:1056–1065.e7.
36. Vögtle FN, Wortelkamp S, Zahedi RP, Becker D, Leidhold C, Gevaert K, et al. Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. *Cell*. 2009;139:428–39.
37. Brix J, Rüdiger S, Bukau B, Schneider-Mergener J, Pfanner N. Distribution of binding sequences for the mitochondrial import receptors Tom20, Tom22, and Tom70 in a presequence-carrying preprotein and a non-cleavable preprotein. *J Biol Chem*. 1999;274:16522–30.
38. Saitoh T, Igura M, Obita T, Ose T, Kojima R, Maenaka K, et al. Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states. *EMBO J*. 2007;26:4777–87.
39. Yamano K, Yatsukawa Y-I, Esaki M, Hobbs AEA, Jensen RE, Endo T. Tom20 and Tom22 share the common signal recognition pathway in mitochondrial protein import. *J Biol Chem*. 2008;283:3799–807.
40. Yamamoto H, Itoh N, Kawano S, Yatsukawa YI, Momose T, Makio T, et al. Dual role of the receptor Tom20 in specificity and efficiency of protein import into mitochondria. *Proc Natl Acad Sci U S A*. 2011;108:91–6.
41. Schmidt O, Harbauer AB, Rao S, Eylich B, Zahedi RP, Stojanovski D, et al. Regulation of mitochondrial protein import by cytosolic kinases. *Cell*. 2011;144:227–39.
42. Becker T, Song J, Pfanner N. Versatility of preprotein transfer from the cytosol to mitochondria. *Trends Cell Biol*. 2019;29:534–48.
43. Yamamoto H, Fukui K, Takahashi H, Kitamura S, Shiota T, Terao K, et al. Roles of Tom70 in import of presequence-containing mitochondrial proteins. *J Biol Chem*. 2009;284:31635–46.
44. Backes S, Hess S, Boos F, Woellhaf MW, Gödel S, Jung M, et al. Tom70 enhances mitochondrial preprotein import efficiency by binding to internal targeting sequences. *J Cell Biol*. 2018;217:1369–82.
45. Bohnert M, Rehling P, Guiard B, Herrmann JM, Pfanner N, van der Laan M. Cooperation of stop-transfer and conservative sorting mechanisms in mitochondrial protein transport. *Curr Biol*. 2010;20:1227–32.
46. Chacinska A, Lind M, Frazier AE, Dudek J, Meisinger C, Geissler A, et al. Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. *Cell*. 2005;120:817–29.
47. Chacinska A, van der Laan M, Mehnert CS, Guiard B, Mick DU, Hutu DP, et al. Distinct forms of mitochondrial TOM-TIM supercomplexes define signal-dependent states of preprotein sorting. *Mol Cell Biol*. 2010;30:307–18.
48. Wagner K, Gebert N, Guiard B, Brandner K, Truscott KN, Wiedemann N, et al. The assembly pathway of the mitochondrial carrier translocase involves four preprotein translocases. *Mol Cell Biol*. 2008;28:4251–60.
49. Tamura Y, Harada Y, Shiota T, Yamano K, Watanabe K, Yokota M, et al. Tim23-Tim50 pair coordinates functions of translocators and motor proteins in mitochondrial protein import. *J Cell Biol*. 2009;184:129–41.
50. Reinhold R, Krüger V, Meinecke M, Schulz C, Schmidt B, Grunau SD, et al. The channel-forming Sym1 protein is transported by the TIM23 complex in a presequence-independent manner. *Mol Cell Biol*. 2012;32:5009–21.
51. Ieva R, Schrempf SG, Opaliński Ł, Wollweber F, Höß P, Heißwolf AK, et al. Mgr2 functions as lateral gatekeeper for preprotein sorting in the mitochondrial inner membrane. *Mol Cell*. 2014;56:641–52.
52. Waegemann K, Popov-Celeketić D, Neupert W, Azem A, Mokranjac D. Cooperation of TOM and TIM23 complexes during translocation of proteins into mitochondria. *J Mol Biol*. 2015;427:1075–84.
53. Moulin C, Caumont-Sarcos A, Ieva R. Mitochondrial presequence import: multiple regulatory knobs fine-tune mitochondrial biogenesis and homeostasis. *Biochim Biophys Acta Mol Cell Res*. 1866;2019:930–44.
54. Wiedemann N, Truscott KN, Pfannschmidt S, Guiard B, Meisinger C, Pfanner N. Biogenesis of the protein import channel Tom40 of the mitochondrial outer membrane: intermembrane space components are involved in an early stage of the assembly pathway. *J Biol Chem*. 2004;279:18188–94.
55. Gebert N, Chacinska A, Wagner K, Guiard B, Koehler CM, Rehling P, et al. Assembly of the three small Tim proteins precedes docking to the mitochondrial carrier translocase. *EMBO Rep*. 2008;9:548–54.
56. Foshag D, Henrich E, Hiller E, Schäfer M, Kerger C, Burger-Kentischer A, et al. The *E. coli* S30 lysate proteome: a prototype for cell-free protein production. *New Biotechnol*. 2018;40:245–60.
57. Káldi K, Bauer MF, Sirrenberg C, Neupert W, Brunner M. Biogenesis of Tim23 and Tim17, integral components of the TIM machinery for matrix-targeted preproteins. *EMBO J*. 1998;17:1569–76.
58. Martin J, Mahlke K, Pfanner N. Role of an energized inner membrane in mitochondrial protein import:  $\Delta\Psi$  drives the movement of presequences. *J Biol Chem*. 1991;266:18051–7.
59. Bohnert M, Zerbes RM, Davies KM, Mühleip AW, Rampelt H, Horvath SE, et al. Central role of Mic10 in the mitochondrial contact site and cristae organizing system. *Cell Metab*. 2015;21:747–55.
60. Moczko M, Ehmann B, Gärtner F, Hönlinger A, Schäfer E, Pfanner N. Deletion of the receptor MOM19 strongly impairs import of cleavable preproteins into *Saccharomyces cerevisiae* mitochondria. *J Biol Chem*. 1994;269:9045–51.
61. Bömer U, Pfanner N, Dietmeier K. Identification of a third yeast mitochondrial Tom protein with tetratricopeptide repeats. *FEBS Lett*. 1996;382:153–8.
62. Krimmer T, Rapaport D, Ryan MT, Meisinger C, Kassenbrock CK, Blachly-Dyson E, et al. Biogenesis of porin of the outer mitochondrial membrane involves an import pathway via receptors and the general import pore of the TOM complex. *J Cell Biol*. 2001;152:289–300.
63. Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*. 1989;122:19–27.
64. Meisinger C, Pfanner N, Truscott KN. Isolation of yeast mitochondria. *Methods Mol Biol*. 2006;313:33–9.
65. Stojanovski D, Pfanner N, Wiedemann N. Import of proteins into mitochondria. *Methods Cell Biol*. 2007;80:783–806.
66. Wittig I, Braun H-P, Schägger H. Blue native PAGE. *Nat Protoc*. 2006;1:418–28.
67. Schneider B, Junge F, Shirokov VA, Durst F, Schwarz D, Dötsch V, et al. Membrane protein expression in cell-free systems. *Methods Mol Biol*. 2010;601:165–86.
68. Webb CT, Gorman MA, Lazarou M, Ryan MT, Gulbis JM. Crystal structure of the mitochondrial chaperone TIM9•10 reveals a six-bladed  $\alpha$ -propeller. *Mol Cell*. 2006;21:123–33.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**







- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**


Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)



# Functional control of a 0.5 MDa TET aminopeptidase by a flexible loop revealed by MAS NMR

Diego F. Gauto<sup>1,8</sup>, Pavel Macek<sup>1,9</sup>, Duccio Malinverni <sup>2</sup>, Hugo Fraga<sup>1,3,4</sup>, Matteo Paloni <sup>5</sup>, Iva Sučec <sup>1</sup>, Audrey Hessel<sup>1</sup>, Juan Pablo Bustamante<sup>6</sup>, Alessandro Barducci<sup>3</sup>  & Paul Schanda <sup>1,7</sup> 

Large oligomeric enzymes control a myriad of cellular processes, from protein synthesis and degradation to metabolism. The 0.5 MDa large TET2 aminopeptidase, a prototypical protease important for cellular homeostasis, degrades peptides within a ca. 60 Å wide tetrahedral chamber with four lateral openings. The mechanisms of substrate trafficking and processing remain debated. Here, we integrate magic-angle spinning (MAS) NMR, mutagenesis, co-evolution analysis and molecular dynamics simulations and reveal that a loop in the catalytic chamber is a key element for enzymatic function. The loop is able to stabilize ligands in the active site and may additionally have a direct role in activating the catalytic water molecule whereby a conserved histidine plays a key role. Our data provide a strong case for the functional importance of highly dynamic - and often overlooked - parts of an enzyme, and the potential of MAS NMR to investigate their dynamics at atomic resolution.

<sup>1</sup>Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale (IBS), 71, Avenue des Martyrs, F-38044 Grenoble, France. <sup>2</sup>Department of Structural Biology and Center for Data Driven Discovery, St Jude Children's Research Hospital, Memphis, TN, USA. <sup>3</sup>Departamento de Biomedicina, Faculdade de Medicina da Universidade do Porto, Porto, Portugal. <sup>4</sup>i3S, Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal. <sup>5</sup>CBS (Centre de Biologie Structurale), Univ Montpellier, CNRS, INSERM, Montpellier, France. <sup>6</sup>Instituto de Bioingeniería y Bioinformática, IBB (CONICET-UNER), Oro Verde, Entre Ríos, Argentina. <sup>7</sup>Institute of Science and Technology Austria, Am Campus 1, A-3400 Klosterneuburg, Austria. <sup>8</sup>Present address: ICSN, CNRS UPR2301, Univ. Paris-Saclay, Gif-sur-Yvette, France. <sup>9</sup>Present address: Celonic AG, Eulerstrasse 55, 4051 Basel, Switzerland. email: [barducci@cbs.cnrs.fr](mailto:barducci@cbs.cnrs.fr); [paul.schanda@ista.ac.at](mailto:paul.schanda@ista.ac.at)

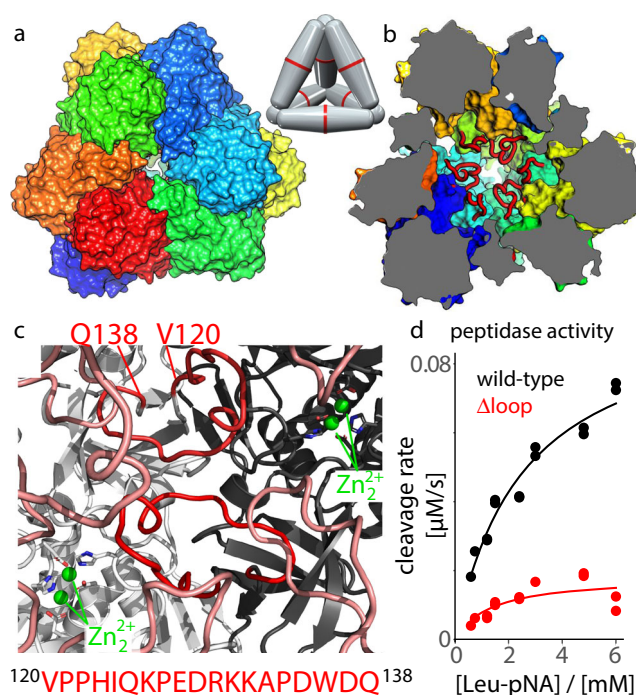
Cells use large protein assemblies to perform many essential biological processes. The cellular protein quality control machinery comprises a collection of such large protein assemblies, including chaperones, unfoldases, proteases and peptidases. Collectively, these proteins eliminate damaged or misfolded proteins either by refolding them to a functional state or by proteolysis. Many proteases and peptidases form large oligomeric assemblies, often in the molecular weight range of hundreds of kilodaltons. The self-compartmentalization of these machineries allows for specificity, as only unfolded proteins or small fragments can access the protease reaction centers. The proteasome, a prominent example, cleaves proteins to peptides of ca. 7–15 residues length<sup>1</sup>. These peptide fragments are then further digested to amino acids by aminopeptidases<sup>2</sup>, such as tetrahedral aminopeptidases, present in all forms of life. While structures of many proteases, peptidases, and chaperones are available, the precise mechanisms of their action, including substrate entry, fixation, and product release, often remain difficult to decipher. Motions and allosteric regulation are intimately linked to enzymatic function, as shown for machines of the protein quality-control system<sup>3–5</sup>. An increasing number of cases reveals that enzymatic turnover can directly depend on the inter-conversions of states, such as conformations in which the active site is open or closed<sup>6–10</sup> or where larger domains reorganize e.g., for binding additional accessory proteins<sup>11</sup>. Characterizing the link between enzyme structure, dynamics, and function at the atomic scale remains, however, experimentally challenging.

We study here the 468 kDa large tetrahedral aminopeptidase TET2 from the hyperthermophilic archaeon *Pyrococcus horikoshii*, a member of the metallo-peptidase family M42. Archaeal TET aminopeptidases and homologous structures in other organisms<sup>12,13</sup> assemble to dodecameric tetrahedral structures, encapsulating twelve Zn<sub>2</sub> active sites within a large hollow lumen with a diameter of ca. 60 Å<sup>14–17</sup> (Fig. 1). Pores on each of the four faces of the tetrahedron, each ca. 18 Å wide, allow the passage of unfolded or short  $\alpha$ -helical, or  $\beta$ -hairpin peptides while preventing folded proteins from entry to the catalytic chamber. The processing occurs in a sequential manner from the N-terminus of the peptide<sup>18</sup>. The fastest hydrolysis is observed for peptides of up to ca. 12 amino acids length, and the longest peptides processed by TET2 are ca. 35 residues long<sup>14</sup>. The active center contains two metal ions, labeled M1 and M2 in the nomenclature of Schechter and Berger<sup>19</sup>, separated by circa 3.5 Å. Besides the catalytic role of the two metals, they are important for assembly to dodecamers<sup>20,21</sup>. In particular, site M1 is implicated in the stabilization of the oligomeric interface. The two metal sites are generally occupied by zinc. Site M1, but not M2, can be exchanged to Co<sup>2+</sup>. The M2 site is considered as the catalytic metal that hosts the catalytically active water molecule, and sits in the specificity pocket that hosts the side chain of the substrate. A third metal site, M3, has been found recently, adjacent to the active site, and shown to broaden the substrate specificity<sup>20</sup>. The binuclear active site is chelated by histidines (H323 at site M1; H68 at M2) and carboxylates (E213, E212, D235, D182; the latter bridges both zinc ions), and a water molecule<sup>15</sup>.

The proposed catalytic mechanism is common to binuclear metallo-aminopeptidases, such as Leucyl-aminopeptidase (LAP) and AAP<sup>12,22,23</sup>. In this mechanism, peptide bond cleavage proceeds via the activation (deprotonation) of a water molecule that bridges the zinc ions. The water molecule donates a proton to a conserved glutamate (E212 in TET2), and the remaining hydroxyl group then attacks the peptide bond of the substrate<sup>15,22,24</sup>. The tetrahedral intermediate state, a diol, where one oxygen stems from the attacking OH and one from the carbonyl group, is stabilized by interaction with the zinc ions. Insight into this state comes from structures obtained with inhibitors; in particular, the

transition-state substrate analog amastatin binds with affinity in the nanomolar range<sup>22</sup> with its scissile bond positioned at the binuclear metal center (Supplementary Figure 1). The proton that was donated by the water molecule and resides on Glu212 is then transferred to what becomes the new N-terminus of the peptide, thus completing the bond cleavage. The cleaved amino acids may be released from the TET particle either through small pores located on the tetrahedral faces close to the central entry pore<sup>15</sup>, or through small pores at the apices of the tetrahedron<sup>18</sup>, or the large entry pores. The kinetics of the reaction proceeds with a rate constant,  $k_{cat}$ , of the order of up to 50 s<sup>-1</sup> at ambient temperature and several thousand s<sup>-1</sup> at the physiological temperatures of the hyperthermophilic TET enzymes<sup>20</sup>. TET can also process substrates when it is in its dimeric form (which is also found in vivo) but the catalytic efficiency is strongly reduced compared to the dodecameric assembly, in particular towards large substrates. Thus, the compartmentalization to dodecameric hollow assemblies has an important role for activity<sup>18</sup>.

Different TET isoforms have different substrate preferences and may assemble into hetero-dodecameric assemblies with improved efficiency for peptide processing: mixed dodecameric assemblies comprising subunits of TET1, TET2 and TET3 may be tuned to degrade entire peptides that may remain within the catalytic chamber throughout the sequential degradation<sup>16,18,25,26</sup>. Homo-dodecameric TET2 displays highest activity for cleavage of



**Fig. 1 Structure and activity of the TET2 aminopeptidase.** Assembly of the dodecameric tetrahedron seen from outside (a), with one of the four entry pores in the center, and in a cut-open view (b), where the loops are shown in red. Different subunits are shown with different colors. The structure is based on PDB ID 1YOR, and the loops (unresolved in the crystal structure) have been modeled with Swiss-Model. The schematic model shows the arrangements of the six dimers within the dodecamer. **c** View into the catalytic chamber and onto two adjacent subunits (light grey, black) with their respective loops (red), and the catalytic zinc ions in the active sites, as well as loops from adjacent subunits. **d** Enzyme kinetics data, obtained as the initial rate of absorbance signal following the cleavage of leucine-p-nitroanilide (Leu-pNA). Data points at identical substrate concentration indicate duplicate measurements. Source data are provided as a Source Data file.



hydrophobic residues, with a preference for leucine as the N-terminal amino acid, showing a high activity up to 100 °C over a broad pH range<sup>27</sup>. Eukaryotic homologs of TET2 are involved in blood pressure regulation in humans<sup>28,29</sup> and hemoglobin degradation in the human malaria parasite *P. falciparum*<sup>30</sup>.

Although static high-resolution structures are available, important mechanistic aspects of the function of TET peptidases remain debated, including substrate entry and trafficking within the lumen of the TET particle, as well as the release of amino acids<sup>18</sup>. To understand such mechanistic details, characterizing the dynamics and interactions within the protein and with substrates is of great interest. Solution-state NMR is ideally suited to study protein dynamics and thus link structures to function. However, for proteins of the size of TET2, solution-NMR suffers from rapid signal loss for most sites. Methyl-specific labeling is often the only way to obtain site-specific information for proteins beyond ca. 100 kDa<sup>31–34</sup>. This approach is by definition limited to methyl-bearing residues.

Unlike solution-state NMR, magic-angle spinning (MAS) NMR does not face inherent protein size limitations and allows to see, in principle, each atom. MAS NMR is a powerful technique for studying dynamics at atomic resolution, and has been applied to sedimented, crystalline, membrane and amyloid proteins<sup>35–47</sup>. Most of the previous MAS NMR dynamics studies focused on proteins below 20–30 kDa, as the resonance overlap often encountered in larger proteins complicates analyses. We have recently achieved the resonance assignment of ca. 90% of the backbone atoms, and about 70 % of the side-chain heavy atoms, as well as of methyl groups of Ile, Leu and Val residues<sup>48,49</sup> and Phe ring C-H moieties<sup>50</sup> in TET2. With 353 residues per subunit, TET2 is among the largest proteins for which such near-complete assignment has been achieved. This assignment, along with distance restraints, has allowed us to develop an approach that uses medium-resolution cryo-EM data along with (primarily solid-state) NMR data to solve the structure of TET2<sup>48</sup>.

We use here quantitative MAS NMR experiments, co-evolution analyses and molecular dynamics (MD) simulations to probe at the atomic level the dynamic contacts formed between the active sites and a functionally important loop. Enzyme kinetics experiments and mutants allow linking these findings to the function of this enzyme. Our study provides direct insight into the functional control of an enzyme through a region which is not even visible in high-resolution crystal structures, and demonstrates the maturity of MAS NMR for studying the structure-function link of even very large proteins.

## Results

**A highly dynamic loop in the catalytic chamber controls enzyme activity.** The catalytic chamber of TET2 comprises twelve long loops, one from each subunit. Interestingly, in 3D structures of TET2 obtained by crystallography<sup>15</sup>, these loop regions have not been modeled (residues 120–132 are missing); similarly, in our recent cryo-EM data<sup>48</sup> this region had very weak electron density. In crystal structures of the homologous TET1 and TET3, the loop has been modeled partially (5 and 9 residues are missing in PDB 2wyr and 2wzn, respectively); the modeled residues have B-factors well above average (see Supplementary Figure 1, which also lists the loop residues that have not been modeled). In MD simulations we confirm that these loop regions in TET1 and TET3 are dynamic, even though they have been partially modeled (discussed further below). In a homologous peptidase that forms the same dodecameric assembly, such as those from *Streptococcus pneumoniae* (PDB 3kl9) a 15-residue long stretch is missing, similarly to TET2. All these observations point to large mobility of these loops. When modelled into the TET2 structure, the loops

fill close to 30% of the catalytic chamber volume (Fig. 1a–c). Thus, one may assume that they represent a significant steric penalty for substrate trafficking in the catalytic chamber, raising the question of their possible functional role.

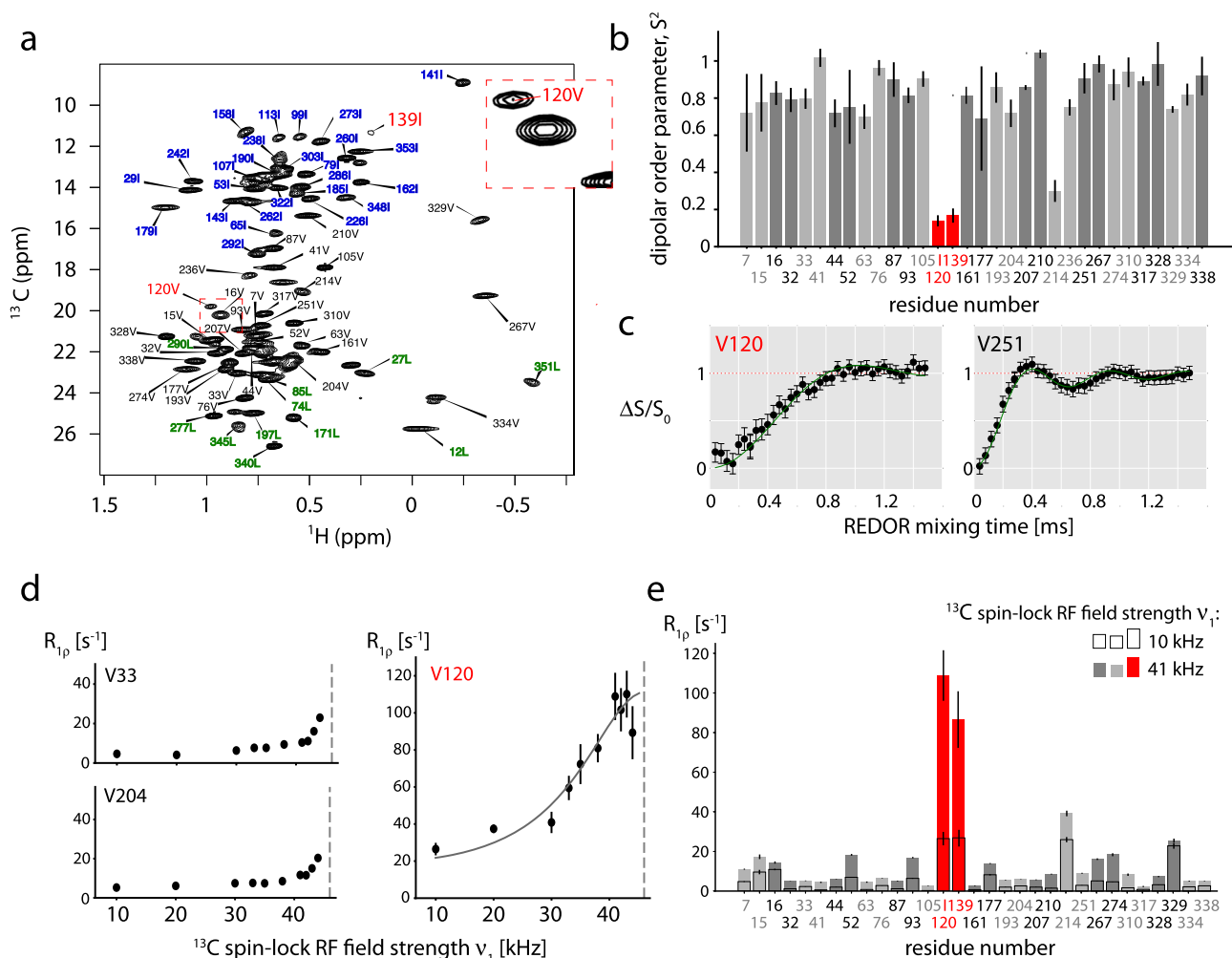
To probe whether the loops influence catalytic activity, we measured the peptidase activity of a TET2 mutant in which the loop has been shortened to a two-residue  $\beta$ -turn ( $\Delta(120-138)$ , henceforth called  $\Delta$ loop mutant). Cleavage of the peptide bond of a small substrate, leucine-p-nitroanilide (H-Leu-pNA), is detected by the absorbance of the reaction product, pNA. The  $\Delta$ loop variant showed a dramatically reduced enzymatic activity compared to the wild-type protein (Fig. 1d and Supplementary Table 1). To ensure that this loop shortening does not significantly impact the protein structure which would lead to the observed drop in activity, we have collected MAS NMR correlation experiments (2D hNH and 3D hCONH). Based on the observation of very similar chemical shifts of the  $\Delta$  loop and WT variants (discussed below, Fig. 3) we can rule out structural distortions in this mutant, and it must be the loop itself which plays an important role for catalytic activity of WT TET2.

We used MAS NMR to probe the conformational behavior of the loop in more detail. In dipolar-coupling based MAS NMR experiments, which are inherently most sensitive for rigid parts, the backbone of residues S119 to K132 and W136 to Q138 could not be assigned<sup>48</sup>. The absence of these peaks in such dipolar-coupling transfers may point to large-amplitude motions that would render the dipolar-coupling based transfer inefficient. In the presence of very fast motion (tens of nanoseconds at most), scalar-coupling based transfers shall lead to efficient transfer. We collected such scalar-coupling based hNH correlation experiments, but we did not observe additional peaks (Supplementary Figure 2).

The <sup>15</sup>N backbone amide site of the assigned and resolved residue D135 shows rapid <sup>15</sup>N  $R_{1\rho}$  spin relaxation ( $\geq 12$  s<sup>-1</sup>; Supplementary Figure 3a), pointing to motions in the nanosecond-to-millisecond range. We have performed additional <sup>15</sup>N  $R_{1\rho}$  Bloch-McConnell relaxation-dispersion experiments<sup>51</sup>. The relaxation-dispersion profile of D135 is non-flat, which shows that this residue undergoes microsecond dynamics (Supplementary Figure 3c). Collectively, the absence of observable signals for most of the backbone sites in the loop, and the direct evidence from spin relaxation of D135 indicates that the loop undergoes  $\mu$ s motion.

In an additional experiment that aims to detect the loop signals we exploited a specific and very sensitive isotope-labeling method of phenylalanines which we developed recently<sup>50</sup>. In this isotope-labeling scheme, a single <sup>1</sup>H-<sup>13</sup>C pair in the para-position of the Phe ring is introduced in an otherwise fully deuterated sample. We have shown previously that it allows the sensitive detection of the side chain of Phe. We have prepared such a sample in which we positioned a Phe at position 123 of the loop; for reasons outlined in detail further below this position was chosen. We, thus, expected to see an additional correlation peak in the <sup>1</sup>H-<sup>13</sup>C spectrum of the phenylalanine (para-CH) labeled sample of H123F TET2. However, the experiment only features the ten native Phe sites, but not the F123 (Supplementary Figure 4). This observation, although indirect, is another hint to the presence of large-scale  $\mu$ s-ms dynamics of the loop.

To gain further insight into the loop dynamics, we used methyl-directed MAS NMR of a specifically Ile/Leu/Val <sup>13</sup>CHD<sub>2</sub>-labeled and otherwise deuterated sample (Fig. 2a). The well-resolved cross-peaks of Val 120, located toward the beginning of the loop, and of Ile 139, located at the C-terminal junction of the loop to a  $\beta$ -strand, were assigned through a mutagenesis approach<sup>49</sup>. The additional methyl group signal in the loop,  $\delta$ 1 of Ile 124, is not spectrally resolved<sup>49</sup>. V120 and I139 are convenient probes of the loop conformational dynamics, which we quantitatively measured.



**Fig. 2** The loop samples a wide conformational space within the cavity. **a** Methyl spectrum of  $u$ -[ $^2\text{H},^{15}\text{N}$ ],Ile- $\delta$  1-[ $^{13}\text{C}\text{HD}_2$ ],Val- $\gamma$ 2-[ $^{13}\text{C}\text{HD}_2$ ],Leu- $\delta$  2-[ $^{13}\text{C}\text{HD}_2$ ] labeled TET2. **b** Dipolar-coupling-derived  $^1\text{H}$ - $^{13}\text{C}$  order parameters of all Val- $\gamma$ 2  $\text{CHD}_2$  methyl groups and one Ile- $\delta$  1  $\text{CHD}_2$  (Ile139), derived from a Rotational Echo Dipolar Recoupling (REDOR) experiment at 55.555 kHz MAS frequency. Data of all Val- $\gamma$ 2, Ile- $\delta$ 1, and Leu- $\delta$ 2 methyls are shown in Supplementary Figure 5. Data for the two sites in the loop are highlighted in red. **c** Example REDOR curves. All REDOR curves are shown in Supplementary Figure 6. **d** Example  $^{13}\text{C}$   $R_{1\rho}$  profiles, up to the NERRD regime, where the  $^{13}\text{C}$  spin-lock radio-frequency (RF) field strength approaches the MAS frequency (46 kHz, dashed vertical line). Data for all methyl groups are reported in Supplementary Figure 7. See Supplementary Figure 9 for discussion of the estimated time scale. **e**  $^{13}\text{C}$   $R_{1\rho}$  values at two  $^{13}\text{C}$  spin-lock field strengths (i.e., two points of the NERRD profile), highlighting that only V120 and I139 have a strong NERRD effect. Data in panels (**b**) and (**e**) are presented as best-fit values (in terms of minimal chi-square)  $\pm$  one standard deviation, where the standard deviation has been determined by Monte Carlo error estimation, based on three times the spectral noise level (see Methods). Source data ( $^{13}\text{C}$  relaxation data, REDOR data and analysis scripts) are provided as a Source Data file.

Fig. 2b, c shows the  $^1\text{H}$ - $^{13}\text{C}$  dipolar-coupling tensor data of all Ile- $\delta$ 1, Val- $\gamma$ 2, Leu- $\delta$ 2 methyl groups. The motion-averaged dipolar-coupling tensor reflects the motional amplitude of the methyl group axis, averaged over all time scales up to hundreds of  $\mu\text{s}$ <sup>52</sup>, and can be directly translated to the order parameter, which ranges from 1 for fully rigid to 0 for fully flexible sites. In the case of a valine, it corresponds to motion of a single sidechain torsion angle ( $\chi_1$ ) and of the backbone. While the vast majority of valines are rather rigid, with order parameters ( $S^2$ ) in the range 0.7 to 1, Val120 is highly flexible ( $S^2 = 0.15 \pm 0.02$ ). Ile 139 also displays a similarly low order parameter. It is noteworthy that the two probes may underestimate the full loop mobility: Val 120 is adjacent to a residue that still had been modeled in the crystal structure, although with high B-factor (Ser 119); Ile 139 is located in a 4-residue short helix adjacent to the loop (again with high B-factors). Thus, while the two probes clearly have the largest amplitude among all methyl-bearing residues we observed, with order parameters below ca. 0.2, the actual loop motion may be even larger.

Spin-relaxation rate constants are sensitive to both the time scales and the amplitudes of dynamics. So-called near-rotary-resonance relaxation dispersion (NERRD) experiments are particularly informative of  $\mu\text{s}$ -ms motion<sup>53</sup>. In this type of experiment, the relaxation rate constant in the presence of a spin-lock pulse,  $R_{1\rho}$ , is measured at different spin-lock radio-frequency (RF) field strengths up to the regime where the RF field nutation frequency reaches the MAS frequency  $\nu$ , also termed the  $n = 1$  rotary-resonance condition ( $\nu_{\text{RF},^{13}\text{C}} = \nu_{\text{MAS}}$ )<sup>51,54,55</sup>. The profile of  $R_{1\rho}$  as a function of RF field strength strongly increases as the RF field strength approaches the MAS frequency only if the bond undergoes  $\mu\text{s}$  motion. Unlike in Bloch-McConnell relaxation-dispersion experiments<sup>56</sup> often used in solution-state NMR, NERRD experiments sense  $\mu\text{s}$ -ms dynamics even if the exchanging conformations do not differ in their chemical shifts. Figs. 2d, e show NERRD data for Val sites in TET2. The NERRD curves of almost all methyl sites are flat or show only a modest increase in  $R_{1\rho}$  ( $\leq 10 \text{ s}^{-1}$ ) close to the rotary-resonance condition (see

Supplementary Figure 7). Strong non-flat NERRD profiles are observed for V120 and I139 and unambiguously demonstrate that these sites undergo  $\mu\text{s}$  motions. As the motion experienced by the V120 side chain is presumably complex, involving methyl rotation, side chain motion and loop reorientation, quantitative analysis is challenging, and we limit the analysis to an estimated time scale of ca. 10–1000  $\mu\text{s}$  (see Supplementary Figure 9 for discussion). Motion on this time scale leads to fast transverse relaxation in MAS NMR experiments<sup>52</sup>, which provides an explanation for the broadening beyond detection of most backbone signals and the elevated  $R_{1\rho}$  and relaxation dispersion of D135 (Supplementary Figure 3).

We additionally measured longitudinal relaxation rate constants ( $^{13}\text{C}$   $R_1$ ), which are sensitive to faster motions (nanoseconds)<sup>52,57,58</sup>. Neither V120 nor I139 have particularly fast  $R_1$  decay, demonstrating that their motion does not take place on the ns time scale (Supplementary Figure 8).

We then characterized which parts of the catalytic chamber are in (possibly transient) contact with the loop. We exploited the fact that the chemical shift is a suitable reporter of such contacts: it is sensitive to the local environment around a given atom, averaged over all conformations sampled on time scales up to milliseconds according to their relative population. Therefore, even transient contacts of a given residue with the loop would be imprinted on the chemical shifts of its atoms (as long as the corresponding conformations have a sizeable population level). It shall, thus be possible to detect which residues are in contact with the loop by comparing the chemical shifts of a wild-type (WT) protein and a protein lacking the loop. We have prepared these two samples of TET2 (WT,  $\Delta$  loop) with uniform labeling ( $^2\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ , in  $\text{H}_2\text{O}$  buffer) and probed the backbone  $^1\text{H}^{\text{N}}$ ,  $^{15}\text{N}$ , and  $^{13}\text{C}\alpha$  chemical shifts. Figure 3a–c shows the hCANH-derived chemical-shift differences, between WT and  $\Delta$  loop, which reflect the effects induced by the loop. As expected, the CSP effects are located in the interior of the enzymatic lumen. The large area within the catalytic chamber involved in loop contacts spans residues from the entry pore to the active site. Transient contacts of the loop with all these residues requires large-amplitude motion, in line with the dynamics data reported in Fig. 2.

**Co-evolution and molecular dynamics simulations detect loop contact sites.** Based on the observation that the loop is crucial for function (Fig. 1d) and in contact with many residues, we reasoned that interaction patterns of the loop may be conserved across TET homologs. Thus, we investigated how residues in the loop co-evolved in more than 20,000 different homologous sequences. Analysis of co-evolution (Direct Coupling Analysis, see Methods) highlights the conservation of contacts involving the loop and residues in the catalytic chamber, both intra- and inter-molecularly (Fig. 3d). Co-evolutionary couplings arise from statistical correlation in a multiple-sequence alignment (MSA). Such co-evolution in a loop region is remarkable, as generally the residues in loop regions evolve quickly<sup>59</sup>. Co-evolution is observed between residues of the loop and e.g., V93, located right next to the  $\text{Zn}_2$  center, and P246 in the entry pore. It is noteworthy that the couplings can be explained only assuming a certain degree of flexibility of the loop, because a static loop is unable to fulfill the co-evolved contacts. Our analysis, therefore, implicitly suggests that the motion itself is conserved by evolution and functionally relevant in the family (or in a significant fraction of it).

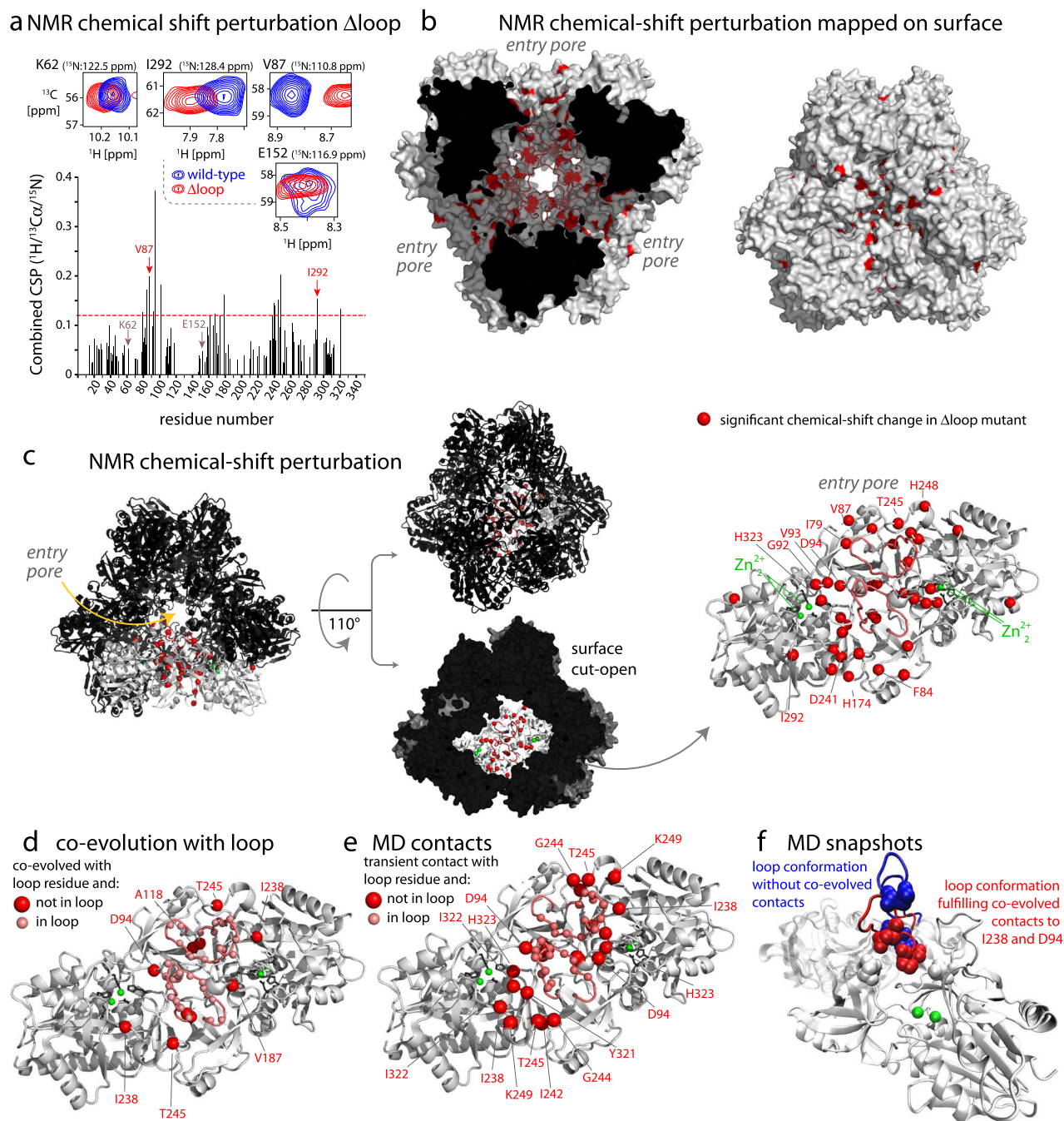
We used one-microsecond-long all-atom molecular dynamics (MD) simulations of the dodecameric TET2 assembly to gain additional insight into the contacts of the loop with other structural parts. These simulations are challenging for several

reasons: with its 468 kDa, TET2 represents a size challenge for all-atom MD, making it difficult to study long time scales; furthermore, as the experimental data revealed, the loop motion occurs on a tens-of-microseconds time scale (Fig. 2 and Supplementary Figure 9). Consequently, in order to obtain convergence from MD simulations, hundreds of microseconds to milliseconds would need to be simulated. Our simulations can, thus, only provide qualitative conclusions, and these are in very good agreement with the experimental observations. Figure 3e highlights the residues within the TET2 cavity which are in transient contact along the MD trajectory; these span the range from the entry pore to the active site, mirroring the NMR CSP data and the co-evolution data. The MD data also allow identifying numerous contacts between the loops of two adjacent monomers within the dimeric building block of TET2, as well as contacts to loops from other subunits (Supplementary Figure 11).

MD provides the possibility to obtain a structural view of the loop conformations. In particular, we aimed to understand which role the evolutionarily-conserved contacts play for the loop conformations. We observed that the loop conformations in which the loop forms contacts to residues D94 and I238 (P121-I238 and P122-D94 and Q125-D94) correspond to states which bring the loop in proximity to the active site (Fig. 3f). This finding suggests that the observed co-evolution may be related to contacts of the loop to substrates. Collectively, three fundamentally different approaches, MAS NMR, MD, and co-evolution analysis, reveal large amplitude motion of the functionally important loop. The MAS NMR relaxation-dispersion data show that this process occurs on a time scale of ca. 10  $\mu\text{s}$  to 1 ms. Interestingly, simulations of TET1, TET2, and TET3 dimeric assemblies also show that this loop is a very dynamic structural element, suggesting that at least within the archaeal TET assemblies the loop flexibility is retained, although the PPH motif (discussed below) is less dynamic in TET1 and TET3 than in TET2 (Supplementary Figure 12).

**Conserved loop residues are important for enzyme-substrate interaction.** To identify the mechanisms of this enzymatic control via a highly flexible loop we investigated the sequence alignment of TET2 homologs, and find a strong conservation of a histidine in this loop, and to lower extent also of a Pro-Pro motif, corresponding to P121, P122 and H123 in TET2 (Fig. 4a). Analysis of the structures of related aminopeptidases and homologs in which the loop has been modeled into the electron density suggests that H123 of a given subunit may be in close vicinity to the active site of the adjacent subunit (Supplementary Figure 1). However, among the crystal structures there is a remarkable variability of the distances between the His and the active site, ranging from ca. 4 Å to over 20 Å. In the cryo-EM structure of TET2, i.e., the sample studied here, the His is over 20 Å away from the active site (Supplementary Figure 13). Histidines often play an important role in enzyme catalysis because the imidazole side chain allows it to form hydrogen bonds and to combine donor and acceptor properties<sup>60</sup>. We envisioned two possible manners how the His in the loop (which is not part of the active site) may play a role in the enzymatic process. On the one hand, the His may stabilize the substrate in the active site. Indeed, hydrogen bonds formed between the substrate and a residue outside the canonical  $\text{Zn}_2$  center play such a stabilizing role in several other aminopeptidases<sup>22</sup>. We speculated that the highly conserved histidine H123 may contribute to stabilizing the substrate in the active site; the Pro-Pro motif preceding H123, is conformationally restricted<sup>61</sup>, and may be important to position the conserved His within the active site. Such a substrate-stabilizing effect shall be reflected in a lowered

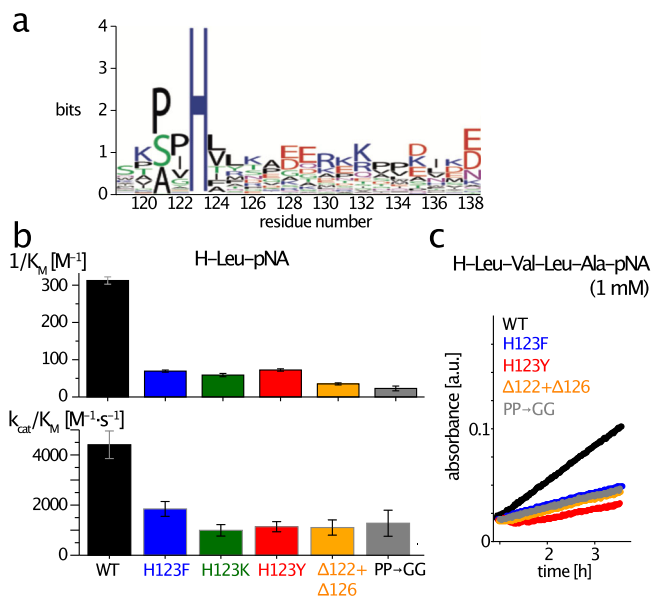




**Fig. 3** The conformational space sampled by the loop detected by MAS NMR, MD and co-evolution analysis. **a** Backbone  $^1\text{H}$ - $^{13}\text{C}\alpha$ - $^{15}\text{N}$  chemical-shift differences of wild-type and  $\Delta$  loop TET2 in 3D hCANH experiments, shown by selected  $^1\text{H}$ - $^{13}\text{C}\alpha$  planes (top) and residue-wise combined CSP values (see methods). **b**, **c** Plot of significant CSP values (3 times standard deviation, indicated by a horizontal line in **a**) on the structure of TET2. In **c**, a dimer (grey) is highlighted, by first showing it in the context of the dodecameric assembly, then rotating and extracting this dimer (right). **d** Co-evolution analysis, showing residues which co-evolved with residues in the loop (residues 120–138), plotted onto the structure of a TET dimeric subunit in the same orientation as in **c**. Red spheres are residues (outside the loop) which co-evolved with residues from the loop (light red spheres). The co-evolution data are listed in Supplementary Table 2. **e** Plot of residues in transient contact with the loop, as observed during the all-atom MD trajectory. Residues were considered in contact when the minimum inter-residue distance between any heavy atoms was below 5 Å. Residues in the loop that are highlighted with spheres have a transient contact with loop from the adjacent subunit. **f** Snapshots of the MD trajectory, in which the evolutionary contacts of loop residues are formed (red) or absent (blue).

Michaelis constant (compared a mutant without these residues). On the other hand the His may play a more active “chemical” role, such as assisting in the nucleophilic attack by activating a water molecule. If the histidine plays a role in the chemistry of the reaction, then one might be able to see the signatures e.g., in pH-dependent effects.

We investigated the role of the conserved residues using functional assays with mutant proteins. Specifically, we mutated H123 to either phenylalanine or tyrosine (ring structure with similar dimensions as His, but without the H-bond donor/acceptor nitrogens) or lysine (to investigate the importance of a positive charge). In an additional mutant,  $\Delta(122,126)$ , we



**Fig. 4 Functional importance of H123, the PP motif and the loop length for enzymatic activity.** **a** Logoplot showing that H123 and the two preceding prolines are highly conserved. **b** Results of a Michaelis-Menten enzyme kinetics assay of TET2 with the short chromogenic substrate H-Leu-pNA. The inverse of the Michaelis constant,  $1/K_M$ , and  $k_{cat}/K_M$  (insert) are shown. The individual points are the best-fit values from individual experiments (replicates). The bar heights and the error bars were obtained from all the replicates. **c** Time traces of enzymatic assays with a longer peptide substrate, H-Leu-Val-Leu-Ala-pNA (1 mM concentration). The limited solubility of the peptide hampered systematic evaluation of  $K_M$  and  $k_{cat}$ , but the results are in good qualitative agreement with those of H-Leu-pNA. Data are presented as best-fit values (in terms of minimal chi-square)  $\pm$  SEM, where the SEM has been determined from a Monte-Carlo based approach described in detail in the Methods section. Source data (kinetic time traces of enzyme kinetics) are provided as a Source Data file.

shortened the loop by one residue on each side of the H123, thus reducing its ability to reach into the active site. To test the importance of the PP motif (residues 121 and 122), we replaced it by a flexible GG stretch.

Figure 4b shows the results of activity assays for the WT and mutant samples processing the chromogenic substrates H-Leu-pNA. The mutants have significantly reduced activity. In particular, the Michaelis constant<sup>62</sup>,  $K_M$ , of the mutants indicates that the stability of the enzyme-substrate complex is reduced by up to one order of magnitude (Fig. 4b and Supplementary Table 1). Experiments with a longer substrate, the tetrapeptide H-Leu-Val-Leu-Ala-pNA, are in good qualitative agreement with the data from the short H-Leu-pNA (Fig. 4c). Taken together, the activity assays show that the stability of the enzyme-substrate complex is reduced through mutations that render the loop either shorter or remove a residue of the conserved Pro-Pro-His motif.

#### Ligand-dependent conformational equilibrium of the loop.

How can a highly dynamic loop without a stably defined position play a crucial role for the activity of an enzyme? We propose that within the wide range of loop conformations there are states which bring important residues, such as H123, close to the active site; in these conformations, contacts to the substrate may increase the stability of the enzyme-substrate complex, as evidenced by the  $K_M$  values. According to this view, the equilibrium of loop conformations is expected to be altered by the presence of substrate bound to the active site.

We experimentally tested this model by measuring the effect of bound ligands on the conformational equilibrium of the loop. The challenge for such experiments comes from the short life time of substrates inside the active site, as they are cleaved within milliseconds<sup>20</sup>. Moreover, the population of TET2 particles that have simultaneously all 12 sites occupied is extremely small. As an experimentally feasible alternative to generate a temporally stable and fully ligand-occupied state of TET2, we prepared samples of TET2 with the inhibitor amastatin (Fig. 5a), a peptide that tightly and non-covalently binds to the active sites of TET2<sup>15</sup>. MAS NMR experiments reveal the chemical-shift perturbations induced by this tightly bound inhibitor (Supplementary Figures 14 and 15a). CSPs are observed in the close vicinity of the binding site, in excellent agreement with the crystal structure<sup>15</sup>. Importantly, a single set of resonances is found in our samples, i.e., the entire population is shifted from the apo state to the amastatin bound one.

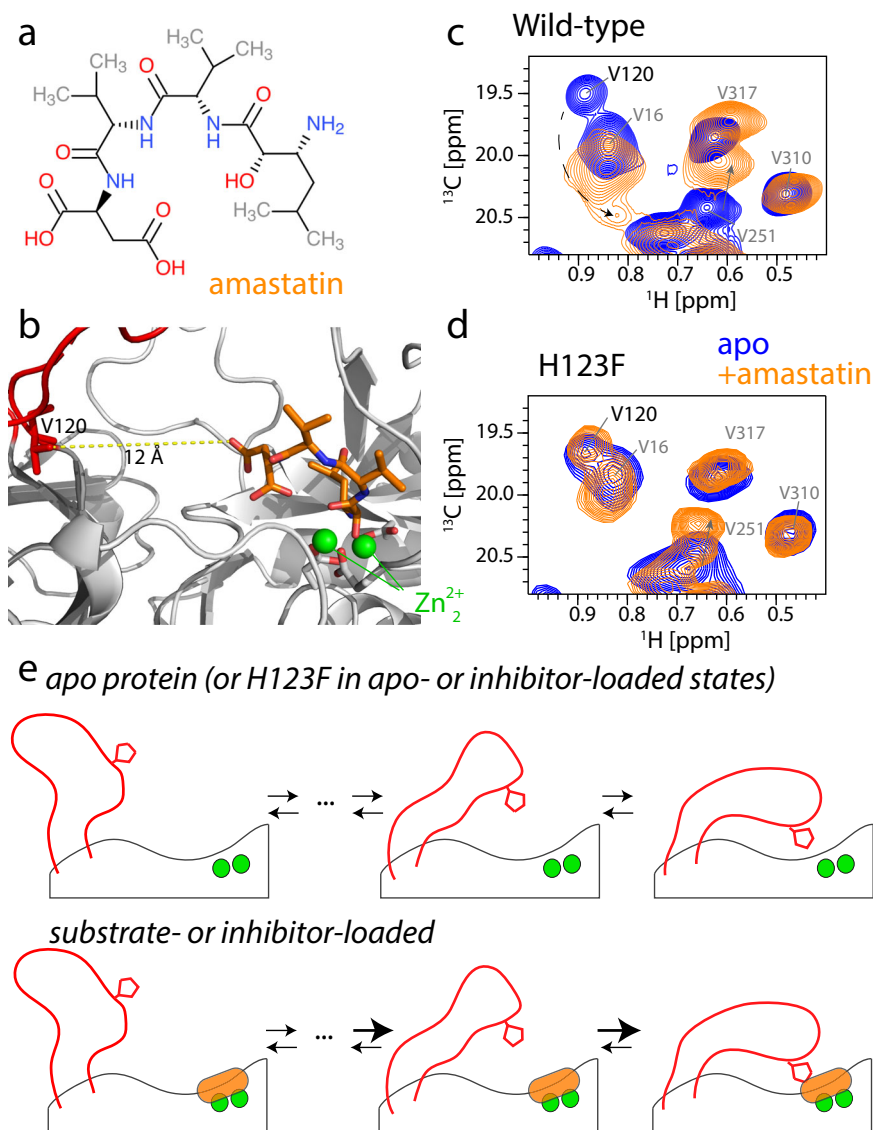
We turned to methyl  $^1H$ - $^{13}C$  correlation spectra and used the signal of V120 to monitor whether the bound inhibitor impacts the conformational equilibrium of the loop. As compared to the apo state, the cross-peak of V120 shifts significantly upon amastatin binding (Fig. 5c). V120 is ca. 6–18 Å away from the nearest atom of the inhibitor in the MD ensemble (Fig. 5b), too far to cause any impact of amastatin on the V120 signal by direct molecular contact. Because the chemical shift reports on the ensemble-averaged conformational equilibrium, the altered peak position of V120 rather reveals that the relative population levels of the loop conformers are altered; similarly, the backbone N-H signal of D135 is significantly altered upon amastatin binding (Supplementary Figure 15a). Figure 5e sketches this idea of a conformational ensemble.

Based on the activity measurements, we expected that H123, via its effect on stabilizing the substrate in the active site, plays an important role in reshuffling the loop conformational equilibrium upon ligand binding. Consequently, we expected that a H123F mutant, unable to form these contacts, would be unable to induce this population re-shuffling. We found that this is exactly the case: in H123F TET2, the reporter NMR resonance of the loop conformation, V120, was essentially unaffected by the presence of the inhibitor in the active site (Fig. 5d). We ensured that amastatin tightly binds to the H123F mutant, evidenced by significant CSPs in  $^1H$ - $^{15}N$  correlation spectra upon inhibitor binding and again a fully bound state (i.e., no residual apo-state peaks), akin to the wild-type protein (Supplementary Figure 15b).

#### Binding and release at the active site depend on the loop.

Having shown that the presence of ligand in the active site alters the loop conformation, we investigated if the reverse is equally true, i.e., if the presence of the loop alters binding and release of a substrate (or a non-cleavable weak binder) in the active site. Because substrates are rapidly degraded by TET2, it is difficult to study their binding and release under equilibrium conditions. We discovered by serendipity that the dialcohol 2-methyl-2,4-pentanediol (MPD), a common crystallization agent used also to obtain solid-state NMR samples, interacts with the active site of TET2. The inhibitory properties of aliphatic alcohols on aminopeptidases have been reported earlier<sup>63</sup>. In the presence of MPD, several H-N moieties (Gly92 (NH), Asp94 (NH), and the zinc-chelating histidine 323 ( $N\delta$ -H $\delta$ )) feature two cross-peaks of approximately equal intensity, indicating that MPD-bound and free states of TET2 co-exist in slow exchange (Fig. 6). A comparison with spectra of TET2 without MPD (sedimented rather than MPD-precipitated) shows that one of these two peaks corresponds to an MPD-bound form.

Whereas in the WT TET2 we find, thus, two peaks in the presence of MPD (MPD-bound and free), we only find a single

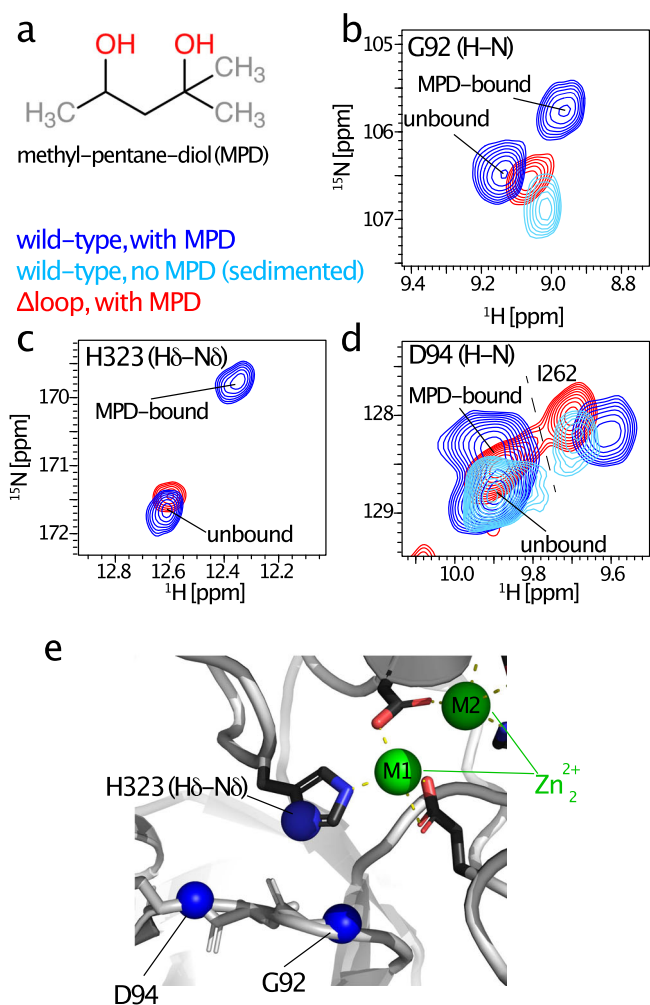


**Fig. 5** The loop conformational ensemble is altered by ligand-binding to the active site. **a** Chemical structure of the inhibitor peptide amastatin. **b** Structural view of amastatin in the active site (PDB: 1Y0Y; loops modeled with SwissModel). **c** <sup>1</sup>H-<sup>13</sup>C correlation spectrum of WT TET2 in the absence and presence of the inhibitor amastatin, showing a clear change of the peak of the loop residue V120. The dashed arrow indicates the putative shift of the V120 cross-peak to a new position. The shift of the cross-peak of V251 is ascribed to effects from direct binding of amastatin. **d** Equivalent <sup>1</sup>H-<sup>13</sup>C spectra of H123F TET2 in the apo and amastatin-bound states. The unchanged peak position and intensity of V120 indicates that the loop ensemble is not altered by amastatin, despite full occupancy of the active site with amastatin (Supplementary Figure 15). The full spectra are shown in Supplementary Figure 16. **e** Schematic representation of the loop conformational ensemble, and the impact of ligands in the active site on reweighting populations within the ensemble.

peak in the  $\Delta$  loop mutant (Fig. 6). For all three sites, G92, D94, and H323 (N $\delta$ -H $\delta$ ), the observed peak is close to the peak position that, in WT TET2, corresponds to the free (not MPD-bound) state. This finding suggests that the MPD-bound state is not significantly populated in the absence of the loop. Of note, these experiments do not provide direct evidence that the loop directly interacts with the ligand (in this case MPD). It is conceivable that the loop stabilizes the MPD-bound state more indirectly, by contacting other residues of the protein rather than the substrate itself. The precise mechanism, as well as the binding affinity, likely depends also on the nature of the substrate. Irrespective, this data shows that the observed affinity of a ligand at the active site directly depends on the presence of the loop, mirroring the reduced binding affinity of substrates that we observed in the activity measurements (Fig. 4,  $1/K_M$  values).

We questioned whether the conserved His may also play a role for the catalytic reaction itself. We first investigated whether addition of free histidine might suffice to increase the activity in a His-free mutant (H123Y). The answer is no: addition of free His not only does not increase the activity, but abolishes the function of WT and mutant TET2, possibly because of binding of the free amino acid to the active site (Supplementary Figure 17). We then performed activity measurements with WT and mutants (H123F, H123Y, H123K,  $\Delta(122,126)$ ) at higher pH values (pH 9.3). Given the fact that the mechanism involves nucleophilic attack by a hydroxide ion, increasing the pH is expected to lead to a significant acceleration. Indeed, the mutants that lack the H123 have higher activity at the higher pH of 9.3. (The increase is modest, which might be due to negative effects on the structural integrity.) However, for WT TET2, the activity is essentially





**Fig. 6** Binding of a weak inhibitor at the active site is altered in the  $\Delta$ loop mutant. **a** Chemical structure of 2-methyl-2,4-pentanediol (MPD), a previously<sup>63</sup> identified competitive inhibitor for aminopeptidases.

**b–d** Zoom on the  $^1\text{H}$ - $^{15}\text{N}$  correlation peaks of the backbone amides of G92 and D94, and the side-chain N $\delta$  position of H323, one of the chelators of the zinc active site. In the sample obtained by precipitating TET2 in 50% (vol/vol) MPD, two peaks are visible; comparison to the sample of ultracentrifuge-sedimented TET2 (light blue) reveals that one of these is MPD-bound and one unbound. In the MPD-precipitated  $\Delta$  loop mutant, only one peak is visible, which, based on the similarity of the peak position, we assign to the state not bound to MPD. **e** Location of the three  $^{15}\text{N}$  sites with peak doubling in the presence of MPD.

unaltered compared to pH 7.5 (Supplementary Figure 18). A plausible explanation of these findings is that H123 and its free electron pair on the imidazolium side chain facilitate abstraction of the proton from water (which is thought to occur on E212). This would explain why already at pH 7.5 the reaction proceeds as fast as at pH 9.3. The mutants lack this possibility, and thus are less active at pH 7.5. (Note that over the pH range the singly protonated imidazolium is predominant<sup>64</sup>.) Given the rather modest difference in pH-response for WT and mutants, we believe, however, that this role of H123, if present, is minor compared to substrate stabilization.

## Discussion

Why do the large TET aminopeptidases, present in all kingdoms of life, feature loop regions which fill up almost one-third of their

catalytic chamber? Why has evolution generated these long stretches which seem to hamper the access of substrates to the active sites, rather than having an empty spacious lumen? Our combined MAS NMR, functional and computational study clarifies the functional role of these loops, which we show to be highly flexible (Fig. 2). We demonstrated that these loops act to stabilize substrates in the active site (higher enzyme-substrate affinity, Fig. 4). The loop-substrate interaction in turn shifts the conformational ensemble of the loop (Fig. 5e), and the loop has an impact on active-site binding of a ligand, i.e., it seems to favor the bound state (Fig. 6). Crystallography (Supplementary Figure 1) and MD simulations (Supplementary Figure 12) point to motion of these loops also in other TET isoforms, although the effect appears most pronounced in TET2.

Residues in loop regions evolve rapidly, on evolutionary time scales, and are generally hardly conserved<sup>59</sup>. Evolution of the physico-chemical characteristics of the loop may have helped to widen the substrate specificity of TET peptidases. Remarkably, though, we identified a highly conserved His within a Pro-Pro-His stretch, and demonstrated its functional relevance. Interestingly, histidines in loop regions close to the substrate have been identified also in the peptidases APP and eMetAP<sup>65,66</sup>, and also proposed for thermolysin<sup>67</sup>. Although the structural scaffolds of these peptidases are unrelated to TET2, it has been found in crystal structures with ligands trapped in the active sites that the imidazolium nitrogen of histidine side chains interacts with oxygens of the bound ligand (Supplementary Figure 19). Our data indicate that H123 of TET peptidases may act similarly; the histidine is indeed able to reach the substrate for forming such interactions (Supplementary Figures 13 and 20). A possible additional role of the histidine might be to polarize water molecules, thereby increasing their nucleophilic properties. Given the modest effect (Supplementary Figure 18) we assume that this role is minor.

Why is this functionally important element highly flexible, rather than being located on a short, less flexible element in the direct vicinity of the active site? We propose that the high degree of flexibility is required to allow the passage of substrates within the chamber, particularly as the substrates can be up to 35 amino acids long. Freedom of movement within the chamber is important not only for newly entering substrates but also for substrates that were cleaved once and which remain in the chamber for further degradation at one of the 12 catalytic sites<sup>18</sup>. The length and flexibility of the loops may furthermore allow the required versatility for the interaction with a broad range of substrates of different lengths.

On the methodological side, the current study establishes that MAS NMR is highly suited for probing enzyme function, even of very large complexes such as the half-megadalton large TET assembly. For complexes of this size, solution-state NMR<sup>68,69</sup> is generally limited to methyl groups. The ability of MAS NMR to detect essentially all backbone and side chain sites allows to obtain a more comprehensive view; here, only the combination of methyl data with backbone and even side chain His resonances allowed seeing with parts are in contact with the loop (Fig. 3), or binding of the ligands (Fig. 6). We have exploited advanced MAS NMR methods to probe dynamics, including  $^{13}\text{C}$  NERRD data (Fig. 2d), which, to our knowledge is the first report of this method, and asymmetric dipolar-coupling tensor averaging (Fig. 2b), both of which are unavailable for solution-state NMR methods. The prerequisite for performing such studies at the atomic level is that the individual cross-peaks are visible and resolved. In the present case, the spectral resolution is, generally, very high, even at modest magnetic field strength (600 MHz  $^1\text{H}$  Larmor frequency). While similarly high resolution has been reported for a number of other systems by MAS

NMR, the present approach is not necessarily general for any system. Also, due to the extensive  $\mu\text{s}$  mobility, most of the backbone of the key loop here went undetected. In the general case, a combination of methods such as MAS NMR, solution-NMR and other spectroscopic and structural methods along with simulations may be required, depending on the molecular system. Such insight into conformational dynamics exploration might be decisive to reveal the connections between static structures to functional mechanisms.

## Methods

**Protein samples.** TET2 from *P. horikoshii* (UniProt entry O59196) was produced by overexpression of a pET41c plasmid encoding the TET2 sequence in *Escherichia coli* BL21(DE3) (Novagen) cells in suitably isotope-labeled M9 minimum media (for all NMR) or LB medium (for functional assays). Samples used for NMR studies were either  $u\text{-}[^2\text{H},^{13}\text{C},^{15}\text{N}]$  labeled (for all 3D H-N-C correlation experiments, Figs. 3 and 6), or  $u\text{-}[^2\text{H},^{15}\text{N}]$ , Ile- $\delta$  1- $^{13}\text{CHD}_2$ , Val- $\gamma$  2- $^{13}\text{CHD}_2$ , Leu- $\delta$  2- $^{13}\text{CHD}_2$  labeled (data in Figs. 2c and 5). The sample of amastatin-bound WT TET2 (Fig. 2d) was  $u\text{-}[^2\text{H},\text{Val-}\gamma$  2- $^{13}\text{CHD}_2$ , Leu- $\delta$  2- $^{13}\text{CHD}_2$ ] labeled. The labeling of deuterated samples was achieved by using M9 minimum culture media in 99.8%  $\text{D}_2\text{O}$ , the use of  $^{15}\text{NH}_4\text{Cl}$  as sole nitrogen source, and D-glucose (deuterated and  $^{13}\text{C}_6$ -labeled for  $u\text{-}[^2\text{H},^{13}\text{C},^{15}\text{N}]$  samples or only deuterated, not  $^{13}\text{C}$  labeled, for methyl-labeled samples). Proteins for enzymatic assays were produced in a similar manner (temperature, growth time), but in LB medium.

*E. coli* BL21(DE3) cells were transformed with the pET41c-PhTET2 plasmid (kanamycin resistance). For production of deuterated samples, the cells were adapted to M9/D<sub>2</sub>O medium in three steps (preculture in LB/H<sub>2</sub>O during day, preculture in M9/H<sub>2</sub>O over night, preculture in M9/50% H<sub>2</sub>O/50% D<sub>2</sub>O during day, M9/D<sub>2</sub>O over night). For the culture, cells were grown to an OD<sub>600</sub> of ca. 0.6. At this point, for the methyl-labeled samples, ketoacid precursors were added, according to the manufacturer's instruction (NMR-bio; [www.nmr-bio.com](http://www.nmr-bio.com)), and protein expression was induced 1 hour later by isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG, 1 mM in final culture). The culture was grown at 37 °C for 4 h before harvesting by centrifugation.

For protein purification, the cell pellet was resuspended in lysis buffer [50 mM Tris, 150 mM NaCl, 0.1% Triton X-100, lysozyme (0.25 mg/ml), deoxyribonuclease (0.05 mg/ml), 20 mM MgSO<sub>4</sub>, and ribonuclease (0.2 mg/ml) (pH 8)]. Cells were disrupted in a Microfluidizer using three passes at 15,000 psi. The extract was heated to 85 °C for 15 min followed by centrifugation at 17,500 relative centrifugal force (rcf) for 1 hour at 4 °C. The supernatant was dialyzed overnight against 20 mM Tris and 100 mM NaCl (pH 7.5) at room temperature and centrifuged at 17,500 rcf for 10 min at 4 °C. The supernatant was loaded on a Resource Q column (GE Healthcare) and TET2 was eluted with a linear gradient [0 to 1 M NaCl in 20 mM Tris (pH 8) over 10 column volumes]. The fractions containing protein with similar mass (39 kDa), according to SDS-polyacrylamide gel electrophoresis (12.5% polyacrylamide), were pooled and concentrated using an Amicon concentrator (Millipore) with a molecular mass cutoff of 30 kDa. The protein solution was then loaded onto a HiLoad 16/600 Superdex 200 column (GE Healthcare) equilibrated with buffer containing 20 mM Tris (pH 8) and 100 mM NaCl.

Samples for MAS NMR measurements were prepared as described earlier<sup>48</sup>; briefly, TET2, was concentrated to 10 mg/mL in 100% H<sub>2</sub>O buffer containing 20 mM Tris, 20 mM NaCl (pH 7.6), and mixed 1:1 (vol/vol) with 2-methyl-2,4-pentanediol (MPD), which results in appearance of white precipitate, which we filled into 1.3 mm MAS rotors (Bruker Biospin) using an ultracentrifuge device (ca. 50,000 g, in a Beckman SW32 rotor, 20,000 rpm) for at least 1 hour. We have also prepared samples by sedimenting TET2 from the buffer solution [20 mM Tris, 20 mM NaCl (pH 7.6)] with the same ultracentrifuge parameters, over night, without addition of precipitation agent (used for data shown in Fig. 6, light blue).  $^{13}\text{C}$ - $^{13}\text{C}$  spectra of MPD-precipitated, isopropanol-precipitated and sedimented samples were highly similar, and also similar to solution-state NMR spectra (Supplementary Figure 21). Note also that the MAS NMR spectra are very similar to solution-state NMR spectra (see Figure S1 of reference<sup>48</sup>).

The loop-deletion mutant plasmid, lacking residues 120–138 of the WT sequence, was prepared by the RoBioMol platform at IBS Grenoble within the Integrated Structural Biology Grenoble (ISBG) facility. The other mutants were generated by a commercial provider, GenScript.

**NMR.** MAS NMR data were acquired on a 14.1 T (600 MHz  $^1\text{H}$  Larmor frequency) Bruker Avance III HD spectrometer (Bruker Biospin) using a 1.3 mm probe tuned to  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  frequencies on the main coil, and an additional  $^2\text{H}$  coil that allows for deuterium decoupling, which greatly enhances resolution of  $^{13}\text{CHD}_2$  spectra<sup>70</sup>. One additional data set, a  $^{13}\text{C}$  R<sub>1</sub> measurement, was collected on a 22.3 T (950 MHz  $^1\text{H}$  Larmor frequency) Bruker Avance III HD spectrometer (Bruker Biospin) using a similar  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^2\text{H}$  1.3 mm probe. The effective sample temperature in all experiments was kept at ca. 28 °C, using the water frequency,  $\delta_{\text{H}_2\text{O}}$ , as chemical-shift thermometer (and an internal DSS as chemical-shift reference), whereby the temperature  $T$  (in °C) is related to the bulk water chemical

shift  $\delta_{\text{H}_2\text{O}}$  as  $T = 455 - 90 \cdot \delta_{\text{H}_2\text{O}}$ . This temperature calibration was found to be in good agreement with an independent temperature calibration via KBr chemical shifts<sup>71</sup> in an external sample; we used the water-based temperature calibration throughout this study.

Three-dimensional hCANH, hCONH and hcaCBcaNH experiments<sup>72</sup> were recorded on  $u\text{-}[^2\text{H},^{13}\text{C},^{15}\text{N}]$  labeled wild-type and  $\Delta$ loop TET2 at 55 kHz MAS frequency and 600 MHz  $^1\text{H}$  Larmor frequency. The experiments used cross-polarization steps (i) from  $^1\text{H}$  to  $^{13}\text{C}$  at typical RF field strengths of ca. 90 kHz ( $^1\text{H}$ , linear ramp 90–100%) and 35 kHz ( $^{13}\text{C}$ ) and a typical duration of 2 ms, (ii) from CA or CO to  $^{15}\text{N}$  at typical RF field strengths of ca. 40 kHz ( $^{15}\text{N}$ , linear ramp 90–100%) and 14 kHz ( $^{13}\text{C}$ ), typically for ca. 8 ms and (iii) from  $^{15}\text{N}$  to  $^1\text{H}$  at typical RF field strengths of ca. 95 kHz ( $^1\text{H}$ , linear ramp 90–100%) and 40 kHz ( $^{15}\text{N}$ ), for ca. 1 ms. The additional CA-CB (out and back) transfer step in the hcaCBcaNH was done with a INEPT transfer, using a 6 ms total transfer delay. The selective pulses for the CA-CB transfer had a REBURP shape (70 ppm bandwidth). The selective pulses for homonuclear decoupling (CO from CA and vice versa) in the indirect  $^{13}\text{C}$  dimensions were ISNOB (applied to the  $^{13}\text{C}$  spins to decouple) and REBURP (applied to the  $^{13}\text{C}$  spins in the transverse plane, for the Bloch-Siegert correction element), akin to previously reported experiments<sup>73</sup>. All experiments are implemented in the NMRlib library<sup>74</sup> and freely available for academic users (<http://www.ibs.fr/nmrlib>).

The combined  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  CSP reported in Fig. 3a was calculated as  $\text{CSP} = \sqrt{(\Delta\delta(^1\text{H}))^2 + \alpha_{\text{N}}(\Delta\delta(^{15}\text{N}))^2 + \alpha_{\text{CA}}(\Delta\delta(^{13}\text{C}))^2}$ , where  $\alpha_{\text{CA}} = 0.3$  and  $\alpha_{\text{N}} = 0.1$ , and  $\Delta\delta$  denote the chemical-shift differences in the two spectra in units of ppm.

All  $^{13}\text{C}$  relaxation experiments and the REDOR experiment described below were obtained using pulse sequences reported in Figure S2 of ref. <sup>50</sup> as a series of 2D  $^1\text{H}$ - $^{13}\text{C}$  spectra (also implemented in NMRlib<sup>74</sup>).  $^1\text{H}$ - $^{13}\text{C}$  transfers (out and back) were achieved by cross-polarization, typically using ca. 2 ms long CP transfer with a  $^1\text{H}$  RF field strength of ca. 90 kHz (linear ramp 90–100%) and matching the  $^{13}\text{C}$  RF field strength to the  $n = 1$  Hartmann-Hahn condition (i.e., ca. 35 kHz).  $^{13}\text{C}$  near-rotary-resonance relaxation dispersion (NERRD) R<sub>1 $\rho$</sub>  experiments<sup>51,53</sup> (Fig. 2d, e) were recorded at 14.1 T and a MAS frequency of 46 kHz. Relaxation delays were adapted in the different experiments, in order not to damage the hardware with extensively long high-power spin-lock duration; the delays are listed in Supplementary Table 3.

$^{13}\text{C}$  R<sub>1</sub> measurements were done at 22.3 T, using relaxation delays of 0.05, 0.2, 0.4, 0.6, 0.8, 1.0, 1.25, 1.5, 2.0, 2.5 s.

$^1\text{H}$ - $^{13}\text{C}$  rotational-echo double resonance (REDOR)<sup>75</sup> experiments (Fig. 2c), in the implementation described in ref. <sup>76</sup> were used to measure asymmetric dipolar coupling tensors. The MAS frequency was 55.555 kHz (18  $\mu\text{s}$  rotor period). The  $^1\text{H}$  and  $^{13}\text{C}$   $\pi$  pulses were 5  $\mu\text{s}$  and 6  $\mu\text{s}$  (100 kHz and 83.3 kHz RF field strength), respectively. One out of two  $^1\text{H}$   $\pi$  pulses was shifted away from the center of the rotor period, in order to scale down the dipolar-coupling evolution and thus sample it more accurately, as described earlier<sup>77</sup>, such that the short and long delays between successive  $^1\text{H}$   $\pi$  pulses were 0.5  $\mu\text{s}$  and 7.5  $\mu\text{s}$ , respectively.

NMR data were processed in the Topspin software (version 3, Bruker Biospin) and analyzed using CCPnmr<sup>78</sup> (version 2.3) and in-house written python analysis routines. In analyses of the NERRD experiment, a two-parameter monoexponential decay function was fitted to the spin-lock-duration-dependent peak intensity decays at the various RF field strengths.

The fitting procedure of the REDOR experiment was described previously<sup>76</sup>. Briefly, numerical simulations were performed with the GAMMA simulation package<sup>79</sup> (version 4.3), setting all pulse-sequence related parameters (MAS frequency, pulse durations, RF field strengths and timing) to the values used in the experiment. A series of such simulations was carried out, in which the  $^1\text{H}$ - $^{13}\text{C}$  dipolar-coupling tensor anisotropy was varied from 1030 to 15,000 Hz (where a rigid H-C pair at a distance of 1.115 Å has a tensor anisotropy of 43,588 Hz, which results in a rigid-limit value of 14,529 Hz when considering the fast methyl rotation) with a grid step size of 30 Hz, and the tensor asymmetry was varied from 0 to 1 with a grid step size of 0.05. Each experimental REDOR curve was compared to this two-dimensional grid of simulations (ca. 9800 simulations in total) and a chi-square value was calculated for each simulation. The reported best-fit tensor parameters are those that minimize the chi-square. Error estimates were obtained by a Monte Carlo approach (pages 104–109 of reference<sup>80</sup>). Briefly, for each methyl site 1000 synthetic noisy REDOR curves were generated around the best-fit simulated REDOR curve, using the spectral noise level and assuming a normal distribution for generating the noisy data points, within three times the standard deviation of the noise level of the spectra. These 1000 synthetic REDOR curves were fitted analogously to the above-described procedure, and the standard deviation over the tensor anisotropy and asymmetry is reported as error estimates. Squared order parameters (Fig. 2b) were obtained by dividing the best-fit tensor anisotropy by the rigid-limit value (14,529 Hz), and squaring the value.

**MD simulations.** All MD simulations were performed with Gromacs 2018.3<sup>81</sup> using amber99sb-disp<sup>82</sup> force field for the protein and TIP4D<sup>83</sup> model for water molecules. The v-rescale<sup>84</sup> and Parrinello-Rahman<sup>85</sup> schemes were employed to control temperature ( $T = 300$  K) and pressure ( $P = 1$  atm) respectively. A cutoff of 1 nm was used to compute van der Waals interactions, while electrostatic interactions were evaluated by means of the Particle Mesh Ewald algorithm using a

cutoff of 1 nm for the real space interactions. The LINC86 algorithm was used to restrain all bond lengths to their equilibrium value. High-frequency bond-angle vibrations of hydrogen atoms were removed by substituting them by virtual sites, allowing an integration time step of 4 fs<sup>87</sup>. Initial configuration for the TET2 dodecameric complex was taken from X-ray structure (pdb code: 1Y0R) and the missing loop was modeled with Swiss Model. The TET2 dodecamer was solvated in a rhombic dodecahedron box with a volume of 2880 nm<sup>3</sup> with periodic boundary conditions. Distance restraints between protein molecules and zinc atoms were applied to preserve the local geometry of the enzymatic site. In the simulations of the substrate-bound protein, the substrate was modeled as a tetrapeptide (Leu-Leu-Val-Ala) where the N-terminal residue was modified in order to have a neutral terminus. A substrate molecule was bound to the active site of each monomer by introducing an additional set of distance restraints between substrate and zinc atoms. These restraints were modeled on the basis of the X-ray structure of amastatin-bound complex (pdb code: 1Y0Y) to preserve a correct binding geometry. Apo and substrate-bound systems were energy minimized and equilibrated for 200 ns and then 1  $\mu$ s production runs were performed for each system. Reported results were obtained by analyzing one frame every 100 ps. Residues were considered in direct contact when the minimum inter-residue distance between heavy atoms was below 5 Å, whereas a looser cutoff (7 Å) was considered when evaluating DCA predictions according to standard practices in coevolutionary analysis<sup>88</sup>. Initial structures for TET1 and TET3 dimeric assemblies were taken from X-ray structures with pdb codes 2WYR and 2WZN, respectively. Chains A and C were used for the TET1 system, while chains A and D were used for the TET3 complex. Missing residues of both systems were rebuilt with MODELLER using the interface available in UCSF Chimera 1.11<sup>89</sup>. Initial structure for the dimeric assembly of TET2 complex was generated by combining the X-ray structure (pdb code: 1Y0R) with the conformation of the 115–143 fragment obtained by electron microscopy (pdb code: 6R8N). Each dimeric complex was solvated in rhombic dodecahedron boxes with a volume of 1480 nm<sup>3</sup> with periodic boundary conditions, and simulated for 1  $\mu$ s.

**Enzymatic activity assays.** The enzymatic activity was measured by following the absorbance change induced when a para-nitroanilide (pNA) labeled substrate is enzymatically cleaved using aminoacyl-pNA compounds H-Leu-pNA and H-Leu-Val-Leu-Ala-pNA (Bachem, Bubendorf, Switzerland) as substrates. Measurements were performed on a BioTek Synergy H4 plate reader (Fisher Scientific) measuring the absorbance at 410 nm in a 384-well plate at 50 °C. In all cases, the wells were filled with 50  $\mu$ L of substrate solution at concentrations varying in the range from 0.1 to 6.4 mM for H-Leu-pNA and 1 mM for H-Leu-Val-Leu-Ala-pNA in buffer (20 mM Tris, 100 mM NaCl, pH 7.5); plates were briefly centrifuged to ensure that the solution is in the bottom of the wells. The plate loaded with the substrate solutions was pre-equilibrated for 20 min at 50 °C. Then, 10  $\mu$ L of the protein solution (in the same buffer as the substrate) was added on each well in order to reach a final protein solution concentration on each well of 5 ng/ $\mu$ L. All solutions contained 2.8% (vol/vol) dimethylsulfoxide (DMSO; Sigma-Aldrich), which increases the solubility of the substrates. In order to minimize changes in the substrate solution (e.g., temperature) upon the protein addition, the plate was kept above the plate-reader thermostat and an electronic multichannel pipette was employed to load the protein solution into the wells and gently mix the solution. We estimated the pNA concentration from the solution absorptivity (molar absorption coefficient for the pNA at 410 nm of 8800 M<sup>-1</sup> cm<sup>-1</sup>). The path length (0.375 cm) was estimated considering the shape and dimensions of the plate wells and the final volume of the solution. Before analysis, curves from blank sample (no protein) were subtracted. The time-dependent absorbance values were analyzed with in-house written python scripts, by fitting the initial rate with a linear equation. Duplicate measurements (time traces of pNA absorbance) were performed. The error estimate of these initial slopes was obtained from the python function `lmfit` (least-squares fit routine). The difference of the duplicate measurements was small (ca. 3% or less), of the same order as the error estimate. These initial-regime slopes as a function of the substrate concentration were fitted to obtain Michaelis-Menten parameters  $K_M$  and  $k_{cat}$ , reported in Figs. 1d and 4b,c and Supplementary Table 1. In this fit, all data points (including duplicates) were used in a joint fit. To determine the error estimates of the  $K_M$  and  $k_{cat}$  parameters, a Monte Carlo approach was chosen, following the principles described e.g., in ref. <sup>80</sup>. In brief, 1000 noisy data sets (initial slope vs. substrate concentration) were created, assuming a normal distribution around the experimentally obtained slopes with  $\sigma$  corresponding to the error estimate of the slope (see above). The reported error bars in Fig. 4b and Supplementary Table 1 are the standard deviations over these 1000 Monte Carlo fits.

The measurements shown in Supplementary Figures 17 and 18 were done with essentially the same approach and minute changes: measurements on the same instrument as above were done in 96-well plates, equally at 50 °C. In all cases, the wells were filled with 80  $\mu$ L of H-Leu-pNA substrate solution at 6.4 mM in either 20 mM Tris, 100 mM NaCl, pH 7.5, or in 100 mM CAPSO buffer at pH 9.3 or in 100 mM MES buffer at pH 5.3. To investigate the role of free histidine, 200 mM histidine solution was prepared in pH 7.5 buffer and the volume added to the reaction well was adjusted to have final concentration of 1, 5 or 20 mM histidine in the reaction solution.

**Bioinformatic analyses.** An initial seed for the co-evolution analysis was built using the sequences contained in the PFAM seed of the M42 Peptidase family (PFAM ID: PF05343) and aligned using the MAFFT utility. The alignment was then curated, removing overly gapped regions. This resulted in a sequence model consisting of 353 positions, covering the whole width of the *Pyrococcus horikoshii* TET2 peptidase (Uniprot ID O59196). A hmmer model of the family was then built using the hmmbuild utility and used to search the uniprot database (union of TREMBL and Swissprot datasets, release 07\_2019) for homologs using the hmmssearch utility, with standard inclusion thresholds. To remove fragments, the retrieved homologs were further filtered by coverage, keeping only sequences containing no more than 25% gapped positions. The loop region of TET was defined as lying between V120 and Q138 in the *Pyrococcus horikoshii* TET2 peptidase. Starting from this final Multiple Sequence Alignment (MSA), logo sequences considering only the mentioned loop region, including some neighboring residues due to highly conserved physicochemical properties between them (residues 115 to 139) were made using seqlogo<sup>90</sup>, a method that takes the position weight matrix of a DNA sequence motif and plots the corresponding sequence logo according to parameters. Column heights in Fig. 4a are proportional to the information content. Regarding sequence identity, no significant differences were observed between logo sequences considering full MSA versus 90% sequence identity. The frequencies at the position corresponding to His123 in TET2 across the alignment are: His 88%, Gaps 7.2%, and the remaining AA all have frequencies of  $\leq 1\%$ . The conservation of the His is very strong for close homologs, and there are near regions with high levels of conservation too in the MSA, suggesting that the remaining sequences having gaps or other amino acids in this particular position correspond to remote homologs.

Direct-Coupling Analysis (DCA) was performed using the asymmetric version of the pseudo-likelihood maximization method, implemented in the lbsDCA code<sup>88</sup>, using standard regularization parameters. To remove sampling bias, sequences were reweighted by identity, downweighting sequences with more than 90% sequence identity to homologs. DCA results were processed using utilities in the dcaTools package<sup>88</sup> (<https://gitlab.com/ducciomalinverni/lbsDCA>). To ignore uninformative very-short range predictions, all reported predictions and accuracies are for residue pairs separated by more than four residues along the chain. Structural contacts were defined by inter-atomic distances between heavy-atoms below 8 Å.

The sequence mining procedure resulted in the extraction of 26'067 TET homologs with at least 75% coverage. After reweighting by sequence identity, the number of effective sequences was of 9157.67, giving an excellent  $B_{eff}/N^{Pos}$  ratio of 25.9, where  $B_{eff}$  denotes the number of effective sequences after weighting sequences by sequence identity<sup>91</sup> and  $N^{Pos}$  denotes the number of residue positions (i.e., columns) in the MSA. DCA prediction benchmarked on the 1Y0Y structure show excellent prediction accuracies over a large range of predictions (Supplementary Figure 10a). Notably, considering the top  $2N = 706$  highest ranked DCA predictions results in a prediction accuracy of 88%. Ignoring the false-positives rising from predictions falling in regions where the PDB structure is not defined, the accuracy rises above 90%. Inspection of the predicted contacts with respect to the 1Y0Y PDB structures (Supplementary Figure 10b and Fig. 3) highlights the prediction of multiple sets of contacts involving the loop region. These can be separated in a set formed by loop-loop interactions, a set of putative intra-molecular loop contacts, and a third set of putative inter-molecular loop interactions (Supplementary Figure 10). Supplementary Table 2 reports the list of all 19 predicted contacts involving the TET loop.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Source data are provided with this paper: REDOR data and the analysis routines (GAMMA simulation program and python analysis script), <sup>13</sup>C relaxation data, the Direct-Coupling Analysis (co-evolution) data, the activity assay data, the <sup>15</sup>N relaxation data (Fig. S3) and the MD-derived contact data (Fig. S11). Data have also been deposited on Mendeley Data, <https://doi.org/10.17632/vx2xmjgmk9.3>. The Protein Data Bank (PDB) entries of the structures used in this work are: 2WZN, 2CF4, 1Y0R, 2WYR, 6R8N, 6F3K, 1Y0Y. The plasmid for expressing TET2 is deposited at Addgene under accession number 182428. Other data are available from the authors upon request. Source data are provided with this paper.

## Code availability

Python and GAMMA code for the fit of the REDOR data is available on Mendeley Data, <https://doi.org/10.17632/vx2xmjgmk9.3> and has been provided with the manuscript.

Received: 12 July 2021; Accepted: 14 March 2022;

Published online: 08 April 2022



## References

- Groll, M. & Clausen, T. Molecular shredders: how proteasomes fulfill their role. *Curr. Opin. Struct. Biol.* **13**, 665–673 (2003).
- Taylor, A. Aminopeptidases: Structure and function. *FASEB J.* **7**, 290–298 (1993).
- Alexopoulos, J. A., Guarné, A. & Ortega, J. ClpP: A structurally dynamic protease regulated by AAA+ proteins. *J. Struct. Biol.* **179**, 202–210 (2012).
- Religa, T. L., Sprangers, R. & Kay, L. E. Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science* **328**, 98–102 (2010).
- Meltzer, M. et al. Allosteric activation of HtrA protease DegP by stress signals during bacterial protein quality control. *Angew. Chemie - Int. Ed.* **47**, 1332–1334 (2008).
- Eisenmesser, E. Z. et al. Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **438**, 117–121 (2005).
- Wolf-Watz, M. et al. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat. Struct. Mol. Biol.* **11**, 945–949 (2004).
- Henzler-Wildman, K. et al. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913–916 (2007).
- Chi, C. N. et al. A structural ensemble for the enzyme cyclophilin reveals an orchestrated mode of action at atomic resolution. *Angew. Chem. Int. Ed.* **54**, 11657–11661 (2015).
- Wurm, J. P. & Sprangers, R. Dcp2: An mrna decapping enzyme that adopts many different shapes and forms. *Curr. Opin. Struct. Biol.* **59**, 115–123 (2019).
- Schuetz, A. K. & Kay, L. E. A dynamic molecular basis for malfunction in disease mutants of p97/VCP. *Elife* **5**, e20143 (2016).
- Dutoit, R. et al. M42 aminopeptidase catalytic site: The structural and functional role of a strictly conserved aspartate residue. *Proteins* **88**, 1639–1647 (2020).
- Kim, D. et al. Structural basis for the substrate specificity of pepa from streptococcus pneumoniae, a dodecameric tetrahedral protease. *Biochem. Biophys. Res. Commun.* **391**, 431–436 (2010).
- Franzetti, B. et al. Tetrahedral aminopeptidase: A novel large protease complex from archaea. *Embo J.* **21**, 2132–2138 (2002).
- Borissenko, L. & Groll, M. Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *J. Mol. Biol.* **346**, 1207–1219 (2005).
- Durá, M. A. et al. The structural and biochemical characterizations of a novel TET peptidase complex from *Pyrococcus horikoshii* reveal an integrated peptide degradation system in hyperthermophilic Archaea. *Mol. Microbiol.* **72**, 26–40 (2009).
- Petrova, T. et al. Structure of the dodecamer of the aminopeptidase apdkam598 from the archaeon *Desulfurococcus kamchatkensis*. *Acta Cryst. F* **71**, 277–285 (2015).
- Appolaire, A. et al. TET peptidases: A family of tetrahedral complexes conserved in prokaryotes. *Biochimie* **122**, 1–9 (2015).
- Berger, A. & Schechter, I. On the size of active site in proteases. I. papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162 (1967).
- Colombo, M., Girard, E. & Franzetti, B. Tuned by metals: The TET peptidase activity is controlled by 3 metal binding sites. *Sci. Rep.* **6**, 20876 (2016).
- Dutoit, R. et al. How metal cofactors drive dimer-dodecamer transition of the M42 aminopeptidase TmPep1050 of *Thermotoga maritima*. *J. Biol. Chem.* **294**, 17777–17789 (2019).
- Lowther, W. T. & Matthews, B. W. Metalloaminopeptidases: Common functional themes in disparate structural surroundings. *Chem. Rev.* **102**, 4581–4607 (2002).
- Schürer, G., Lanig, H. & Clark, T. *Aeromonas proteolytica* aminopeptidase: An investigation of the mode of action using a quantum mechanical/molecular mechanical approach. *Biochemistry* **43**, 5414–5427 (2004).
- Holz, R. C. The aminopeptidase from *aeromonas proteolytica*: Structure and mechanism of co-catalytic metal centers involved in peptide hydrolysis. *Coord. Chem. Rev.* **232**, 5–26 (2002).
- Durá, M. A. et al. Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon *Pyrococcus horikoshii*. *Biochemistry* **44**, 3477–86 (2005).
- Appolaire, A. et al. The TET2 and TET3 aminopeptidases from *Pyrococcus horikoshii* form a hetero-subunit peptidosome with enhanced peptide destruction properties. *Mol. Microbiol.* **94**, 803–814 (2014).
- Russo, S. & Baumann, U. Crystal structure of a dodecameric tetrahedral-shaped aminopeptidase. *J. Biol. Chem.* **279**, 51275–51281 (2004).
- Reaux, A. et al. Aminopeptidase A inhibitors as potential central antihypertensive agents. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 13415–13420 (1999).
- Chen, Y., Farquhar, E. R., Chance, M. R., Palczewski, K. & Kiser, P. D. Insights into substrate specificity and metal activation of mammalian tetrahedral aspartyl aminopeptidase. *J. Biol. Chem.* **287**, 13356–13370 (2012).
- Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002).
- Sheppard, D., Sprangers, R. & Tugarinov, V. Experimental approaches for NMR studies of side-chain dynamics in high-molecular-weight proteins. *Prog. Nucl. Magn. Reson. Spectr.* **56**, 1–45 (2010).
- Sprangers, R. & Kay, L. E. Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* **445**, 618–622 (2007).
- Mas, G. et al. Structural investigation of a chaperonin in action reveals how nucleotide binding regulates the functional cycle. *Sci. Adv.* **4**, eaau4196 (2018).
- Religa, T. L. & Kay, L. E. Optimal methyl labeling for studies of supra-molecular systems. *J. Biomol. NMR* **47**, 163–169 (2010).
- 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.* **130**, 16611–16621 (2008).
- Bonaccorsi, M., Le Marchand, T. & Pintacuda, G. Protein structural dynamics by Magic-Angle Spinning NMR. *Curr. Opin. Struct. Biol.* **70**, 34–43 (2021).
- Singh, H. et al. The active site of a prototypical “rigid” drug target is marked by extensive conformational dynamics. *Angew. Chem. Int. Ed.* **59**, 22916–22921 (2020).
- Byeon, I. J. L. et al. Motions on the millisecond time scale and multiple conformations of HIV-1 capsid protein: Implications for structural polymorphism of CA assemblies. *J. Am. Chem. Soc.* **134**, 6455–6466 (2012).
- Watt, E. D. & Rienstra, C. M. Recent advances in solid-state nuclear magnetic resonance techniques to quantify biomolecular dynamics. *Anal. Chem.* **86**, 58–64 (2014).
- Gelenter, M. D. et al. Water orientation and dynamics in the closed and open influenza B virus M2 proton channels. *Commun. Biol.* **4**, 1–14 (2021).
- Rovó, P. Recent advances in solid-state relaxation dispersion techniques. *Solid State Nucl. Magn. Reson.* **108**, 101665 (2020).
- Felix, J. et al. Mechanism of the allosteric activation of the ClpP protease machinery by substrates and active-site inhibitors. *Sci. Adv.* **5**, eaaw3818 (2019).
- Good, D. B. et al. Conformational dynamics of a seven transmembrane helical protein Anabaena Sensory Rhodopsin probed by solid-state NMR. *J. Am. Chem. Soc.* **136**, 2833–2842 (2014).
- Lamley, J. M. & Lewandowski, J. R. Relaxation-based magic-angle spinning NMR approaches for studying protein dynamics. *eMagRes* **5**, 1423–1434 (2016).
- Schubert, T. et al. A  $\beta$ -barrel for oil transport through lipid membranes: Dynamic nmr structures of alk. *Proc. Natl. Acad. Sci. USA* **117**, 21014–21021 (2020).
- Dregni, A. J., Duan, P. & Hong, M. Hydration and dynamics of full-length tau amyloid fibrils investigated by solid-state nuclear magnetic resonance. *Biochemistry* **59**, 2237–2248 (2020).
- 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* **65**, 171–191 (2016).
- Gauto, D. F. et al. Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nat. Commun.* **10**, 2697 (2019).
- Amero, C. et al. A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J. Biomol. NMR* **50**, 229–236 (2011).
- Gauto, D. F. et al. Aromatic ring dynamics, thermal activation, and transient conformations of a 468 kDa Enzyme by specific 1 H- 13 C labeling and fast magic-angle spinning NMR. *J. Am. Chem. Soc.* **141**, 11183–11195 (2019).
- Ma, P. et al. Probing transient conformational states of proteins by solid-state  $R_{1\rho}$  relaxation-dispersion NMR Spectroscopy. *Angew. Chem. Int. Ed.* **53**, 4312–4317 (2014).
- Schanda, P. & Ernst, M. Studying dynamics by magic-angle spinning solid-state NMR spectroscopy: Principles and applications to biomolecules. *Prog. Nucl. Magn. Reson. Spectr.* **96**, 1–46 (2016).
- Kurauskas, V. et al. Slow conformational exchange and overall rocking motion in ubiquitin protein crystals. *Nat. Commun.* **8**, 145 (2017).
- Krushelnitsky, A., Zinkevich, T., Reif, B. & Saalwächter, K. Slow motions in microcrystalline proteins as observed by MAS-dependent  $^{15}\text{N}$  rotating-frame NMR relaxation. *J. Magn. Reson.* **248**, 8–12 (2014).
- 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.* **54**, 15374–15378 (2015).
- Palmer, A. G. & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev.* **106**, 1700–1719 (2006).
- Lewandowski, J. R. et al. Measurement of site-specific  $^{13}\text{C}$  spin-lattice relaxation in a crystalline protein. *J. Am. Chem. Soc.* **132**, 8252–8254 (2010).
- Lewandowski, J. R. Advances in solid-state relaxation methodology for probing site-specific protein dynamics. *Acc. Chem. Res.* **46**, 2018–2027 (2013).



59. Echave, J., Spielman, S. J. & Wilke, C. O. Causes of evolutionary rate variation among protein sites. *Nat. Rev. Genet.* **17**, 109–121 (2016).
60. Schneider, F. Histidine in enzyme active centers. *Angew. Chem. Int. Ed.* **17**, 583–592 (1978).
61. Williamson, M. P. The structure and function of proline-rich regions in proteins. *Biochem. J.* **297**, 249–260 (1994).
62. Michaelis, L. & Menten, M. L. Die kinetik der invertinwirkung. *Biochem. Z.* **49**, 352 (1913).
63. Ustyniuk, L., Bennett, B., Edwards, T. & Holz, R. C. Inhibition of the aminopeptidase from *aeromonas proteolytica* by aliphatic alcohols. Characterization of the hydrophobic substrate recognition site  $\Delta$ C. *Biochemistry* **38**, 11433–11439 (1999).
64. Li, S. & Hong, M. Protonation, tautomerization, and rotameric structure of histidine: a comprehensive study by magic-angle-spinning solid-state NMR. *J. Am. Chem. Soc.* **133**, 1534–1544 (2011).
65. Wilce, M. et al. Structure and mechanism of a proline-specific aminopeptidase from *escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3472–3477 (1998).
66. Lowther, W. T. et al. *Escherichia coli* methionine aminopeptidase: implications of crystallographic analyses of the native, mutant, and inhibited enzymes for the mechanism of catalysis. *Biochemistry* **38**, 7678–7688 (1999).
67. Beaumont, A. et al. The role of histidine 231 in thermolysin-like enzymes. A site-directed mutagenesis study. *J. Biol. Chem.* **270**, 16803–16808 (1995).
68. Ruschak, A. M. & Kay, L. E. Methyl groups as probes of supra-molecular structure, dynamics and function. *J. Biomol. NMR* **46**, 75–87 (2010).
69. Rosenzweig, R. & Kay, L. E. Bringing Dynamic Molecular Machines into Focus by Methyl-TROSY NMR. *Annu. Rev. Biochem.* **83**, 291–315 (2014).
70. Huber, M. et al. A supplementary coil for  $^2\text{H}$  decoupling with commercial HCN MAS probes. *J. Magn. Reson.* **214**, 76–80 (2012).
71. Thurber, K. R. & Tycko, R. Measurement of sample temperatures under magic-angle spinning from the chemical shift and spin-lattice relaxation rate of  $^{79}\text{Br}$  in KBr powder. *J. Magn. Reson.* **196**, 84–87 (2009).
72. Barbet-Massin, E. et al. Rapid proton-detected NMR assignment for proteins with fast magic angle spinning. *J. Am. Chem. Soc.* **136**, 12489–12497 (2014).
73. Fraga, H. et al. Solid-State NMR H-N-(C)-H and H-N-C-C 3D/4D correlation experiments for resonance assignment of large proteins. *ChemPhysChem* **18**, 2697–2703 (2017).
74. Vallet, A., Favier, A., Brutscher, B. & Schanda, P. ssNMRlib: a comprehensive library and tool box for acquisition of solid-state nuclear magnetic resonance experiments on Bruker spectrometers. *Magn. Reson.* **1**, 331–345 (2020).
75. Gullion, T. & Schaefer, J. Detection of weak heteronuclear dipolar coupling by rotational-echo double-resonance nuclear-magnetic-resonance. *Adv. Magn. Reson.* **13**, 57–83 (1988).
76. 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.* **50**, 11005–11009 (2011).
77. 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.* **132**, 15957–15967 (2010).
78. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–96 (2005).
79. Smith, S., Levante, T., Meier, B. & Ernst, R. Computer simulations in magnetic resonance. An object-oriented programming approach. *J. Magn. Reson.* **106**, 75–105 (1994).
80. Motulsky, H. & Christopoulos, A. Fitting models to biological data using linear and nonlinear regression: A practical guide to curve fitting (Oxford University Press, 2004).
81. Abraham, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
82. Robustelli, P., Piana, S. & Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. USA* **115**, E4758–E4766 (2018).
83. Piana, S., Donchev, A. G., Robustelli, P. & Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **119**, 5113–5123 (2015).
84. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
85. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
86. Hess, B. P-lincs: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
87. Feenstra, K. A., Hess, B. & Berendsen, H. J. C. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* **20**, 786–798 (1999).
88. Malinverni, D. & Barducci, A. Coevolutionary analysis of protein sequences for molecular modeling. In *Biomolecular Simulations*, 379–397 (Springer, 2019).
89. Pettersen, E. F. et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
90. Bembom, O. seqlogo: Sequence logos for dna sequence alignments. r package version 1.32. I. R Foundation for Statistical Computing (2021).
91. Ekeberg, M., Lövkvist, C., Lan, Y., Weigt, M. & Aurell, E. Improved contact prediction in proteins: using pseudolikelihoods to infer potts models. *Phys. Rev. E* **87**, 012707 (2013).

## Acknowledgements

We are grateful to Bernhard Brutscher, Alicia Vallet, and Adrien Favier for excellent NMR platform operation and management. The plasmid coding for TET2 was kindly provided by Bruno Franzetti and Jerome Boisbouvier (IBS Grenoble). We thank Anne-Marie Villard and the RoBioMol platform for preparing the loop deletion construct. The RoBioMol platform is part of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL (ANR-10-LABX-49-01), financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). This work was supported by the European Research Council (StG-2012-311318-ProtDyn2Function to P. S.) and the French Agence Nationale de la Recherche (ANR), under grant ANR-14-ACHN-0016 (M.P. and A.B.).

## Author contributions

D.F.G. performed and analyzed enzymatic assays, collected and analyzed NMR data, prepared figures and contributed to writing the manuscript. P.M. and H.F. prepared protein samples. D.M. performed the co-evolution analysis. M.P. and A.B. performed and analyzed MD simulations. I.S. performed activity assays. A.H. produced protein samples. J.P.B. prepared the logo plots and underlying sequence alignment. P.S. designed, initiated, and supervised the study, performed and analyzed NMR experiments, prepared the figures, and wrote the manuscript with input from all co-authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-29423-0>.

**Correspondence** and requests for materials should be addressed to Alessandro Barducci or Paul Schanda.

**Peer review information** *Nature Communications* thanks Adam Lange and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

## BIOPHYSICS

## Structural basis of client specificity in mitochondrial membrane-protein chaperones

Iva Sučec<sup>1</sup>, Yong Wang<sup>2\*</sup>, Ons Dakhlaoui<sup>1</sup>, Katharina Weinhäupl<sup>1†</sup>, Tobias Jores<sup>3‡</sup>, Doriane Costa<sup>1</sup>, Audrey Hessel<sup>1§</sup>, Martha Brennich<sup>4||</sup>, Doron Rapaport<sup>3</sup>, Kresten Lindorff-Larsen<sup>2</sup>, Beate Bersch<sup>1\*</sup>, Paul Schanda<sup>1\*</sup>

Chaperones are essential for assisting protein folding and for transferring poorly soluble proteins to their functional locations within cells. Hydrophobic interactions drive promiscuous chaperone-client binding, but our understanding of how additional interactions enable client specificity is sparse. Here, we decipher what determines binding of two chaperones (TIM8-13 and TIM9-10) to different integral membrane proteins, the all-transmembrane mitochondrial carrier Ggc1 and Tim23, which has an additional disordered hydrophilic domain. Combining NMR, SAXS, and molecular dynamics simulations, we determine the structures of Tim23/TIM8-13 and Tim23/TIM9-10 complexes. TIM8-13 uses transient salt bridges to interact with the hydrophilic part of its client, but its interactions to the transmembrane part are weaker than in TIM9-10. Consequently, TIM9-10 outcompetes TIM8-13 in binding hydrophobic clients, while TIM8-13 is tuned to few clients with both hydrophilic and hydrophobic parts. Our study exemplifies how chaperones fine-tune the balance of promiscuity versus specificity.

## INTRODUCTION

Cellular survival and function fundamentally rely on an intact proteome. Proteins within cells need to be correctly folded to their functional conformation and be present at the cellular location where they function. Chaperones play a central role in maintaining this cellular protein homeostasis (1), either by helping other proteins to reach their functional three-dimensional (3D) structure after synthesis, by transporting them across the cytosol or organelles, or by sustaining their native fold along their lifetime. More than 20,000 different proteins are required to fulfill the functions of human cells, and it is believed that the majority rely on chaperones to reach and maintain their native fold (2). Given the diversity of the client proteins, many chaperones promiscuously interact with tens of different “client” proteins that may differ widely in size, structure, and physicochemical properties. However, the need for efficient binding and refolding of their clients also calls for some degree of specificity. Chaperones operate at this delicate balance of promiscuity and specificity to their clients. The interactions that determine the chaperone-client specificity are only partly understood.

Hydrophobic interactions play a crucial role for chaperone interactions, as most chaperones bind to hydrophobic patches on their clients and shield them from aggregation. Electrostatic charges also play a role in some chaperone complexes (3). The interaction motifs recognized by different chaperones differ by their physico-

chemical properties (4). For example, for interacting with the Hsp70 chaperone family, Ile, Phe, Leu, and Val residues are particularly important (5, 6); the SecB chaperone recognizes nine-residue-long stretches enriched in aromatic and basic residues (7); the chaperone Spy uses longer-range charge interactions for the formation of an initial encounter complex, followed by more tight binding mediated by hydrophobic interactions (8), whereby structurally frustrated sites on the client protein are particularly prone to binding (9).

Our understanding of the underlying principles of chaperone-client interactions is hampered by the lack of atomic-level views onto the structure and dynamics of these complexes. Their inherently dynamic and often transient nature represents a substantial experimental challenge toward structural characterization. Only a very limited number of chaperone complex structures have been reported [reviewed in (10)]. The modes of interactions that they revealed range from rather well-defined binding poses of client polypeptides in the chaperone’s binding pockets, reminiscent of complexes formed by globular proteins, to highly flexible ensembles of at least partly disordered conformations (“fuzzy complexes”). In the latter, a multitude of local chaperone-client interactions may result in a high overall affinity despite the low affinity and short lifetime of each individual intermolecular contact.

Multiple molecular chaperones are present in the cell with mutually overlapping functions and “clientomes” (2, 11, 12). It is poorly understood, however, whether a given client protein adopts a different conformation (or ensemble of conformations) when it is bound to different chaperones, and if different clients, when bound to a given chaperone, all show similar conformational properties.  $\alpha$ -Synuclein appears to have similar interaction patterns with six different chaperones (13); outer membrane proteins (OmpA, OmpX, and FhuA) have similar properties—essentially fully unfolded—when bound to SurA and Skp chaperones (14, 15), at least as judged by their nuclear magnetic resonance (NMR) fingerprint spectra. Phosphatase A displays an extended dynamic conformation, but well-defined binding poses of its interacting parts, when bound to trigger factor (16), Hsp40 (17), or SecB (18). Thus, while these reports suggest that a given protein adopts similar properties on

<sup>1</sup>Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale (IBS), 71, Avenue des Martyrs, F-38044 Grenoble, France. <sup>2</sup>Structural Biology and NMR Laboratory, the Linderstrøm-Lang Centre for Protein Science, Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark. <sup>3</sup>Interfaculty Institute of Biochemistry, University of Tübingen, 72076 Tübingen, Germany. <sup>4</sup>European Molecular Biology Laboratory, 38042 Grenoble, France.

\*Corresponding author. Email: yong.wang@bio.ku.dk (Y.W.); beate.beresch@ibs.fr (B.B.); paul.schanda@ist.ac.at (P.S.)

†Present address: Universidade do Porto, I3S - Instituto de Investigação e Inovação em Saúde, Porto, Portugal.

‡Present address: Department of Genome Sciences, University of Washington, Seattle, WA, USA.

§Present address: Institute of Pharmacology and Structural Biology, 31077 Toulouse, France.

||Present address: Bruker Optics, 76275 Ettlingen, Germany.

different chaperones, the scarcity of data and the absence of a direct comparison of complex structures leave open which interactions may confer specificity.

A pair of “holdase” chaperone complexes of the mitochondrial intermembrane space (IMS), TIM8-13 and TIM9-10, are structurally highly similar but have different substrate binding preferences. These chaperones transport precursors of membrane proteins with internal targeting sequence (henceforth denoted as “precursors”) to the membrane-insertase machineries in the inner membrane (TIM22) and outer mitochondrial membranes (SAM) (19). The TIM chaperones form hetero-hexameric structures of ca. 70 kDa, composed of an alternating arrangement of Tim9 and Tim10 or Tim8 and Tim13. TIM9-10 is essential to cellular viability (20–22); even single-point mutations in Tim9 or Tim10 that keep the chaperone structure intact but affect precursor protein binding can impair yeast growth and cause lethality (23). Although TIM8-13 is not essential in yeast (24), yeast cells depleted of Tim8 and Tim13 show conditional lethality (25). In addition, mutations in the human Tim8a protein have been identified as the cause of a neurodegenerative disorder known as Mohr-Tranebjærg syndrome or deafness-dystonia-optic neuropathy syndrome (26, 27).

In vivo experiments, predominantly in yeast, have identified mitochondrial membrane proteins whose biogenesis depends on small TIM chaperones. TIM9-10 is believed to interact with all members of the mitochondrial carrier (SLC25) family, which comprises more than 50 members in humans, such as the adenosine diphosphate (ADP)/adenosine triphosphate (ATP) carrier (Aac in yeast); furthermore, TIM9-10 transports the central components of the TIM22 and TIM23 insertion machineries (Tim23, Tim17, and Tim22) as well as outer membrane  $\beta$  barrel proteins (28). TIM8-13 has a narrower clientome and was shown to bind the precursors of the inner membrane proteins Tim23 (25, 29, 30) and  $\text{Ca}^{2+}$ -binding aspartate-glutamate carriers (31), as well as the outer membrane  $\beta$  barrel proteins VDAC/porin, Tom40 (32), and Tob55/Sam50 (33). There is evidence that TIM8-13 does not bind neither the inner membrane protein ADP/ATP carrier (Aac) nor Tim17 (25). The inner membrane proteins that have been reported to interact with TIM8-13 have a hydrophilic domain in addition to transmembrane (TM) domains (fig. S1), but this does not hold true for the outer membrane  $\beta$  barrels. Thus, the mechanisms by which TIM8-13 binds its clients remain unclear.

Recently, we obtained the first structure of a complex of a small TIM chaperone, TIM9-10, with the mitochondrial guanosine diphosphate (GDP)/guanosine triphosphate (GTP) carrier (Ggc1) (23). The structure, composed of two chaperone complexes holding one precursor protein, revealed a highly dynamic ensemble of Ggc1 conformers that form multiple short-lived and rapidly interconverting (<1 ms) interactions with a hydrophobic binding cleft of the chaperone (fig. S2). The TIM9-10-Ggc1 complex can be described as a “fuzzy complex,” in which the high overall affinity is driven by a multitude of individually weak interactions with the hydrophobic TM parts of its clients.

To understand what confers specificity in the mitochondrial IMS chaperone system, we studied chaperone complexes of TIM9-10 and TIM8-13 with two precursor proteins, the Ggc1 and the insertase component Tim23. In their native state, Ggc1 comprises six TM helices without soluble domains, and Tim23 comprises four TM helices and a ca. 100-residue-long soluble IMS domain (Fig. 1A). By solving the complex structures of the two chaperone complexes

holding Tim23, we reveal that the differential specificity of the two chaperones is based on an interplay of hydrophobic and hydrophilic interactions, which leads to different conformational properties of the precursor protein bound to these chaperones.

## RESULTS

### TIM8-13 and TIM9-10 interact differently with membrane precursor proteins

We have developed an experimental protocol (23) to prepare complexes of the inherently insoluble membrane-protein precursors and chaperones (Fig. 1, B and C). Briefly, the approach involves the recombinant production of the His-tagged precursor protein, binding it to a His-affinity column (NiNTA) in denaturing conditions, followed by removal of the denaturant and simultaneous addition of a chaperone. The chaperone-precursor complex is then eluted for further biochemical, biophysical, and structural investigations.

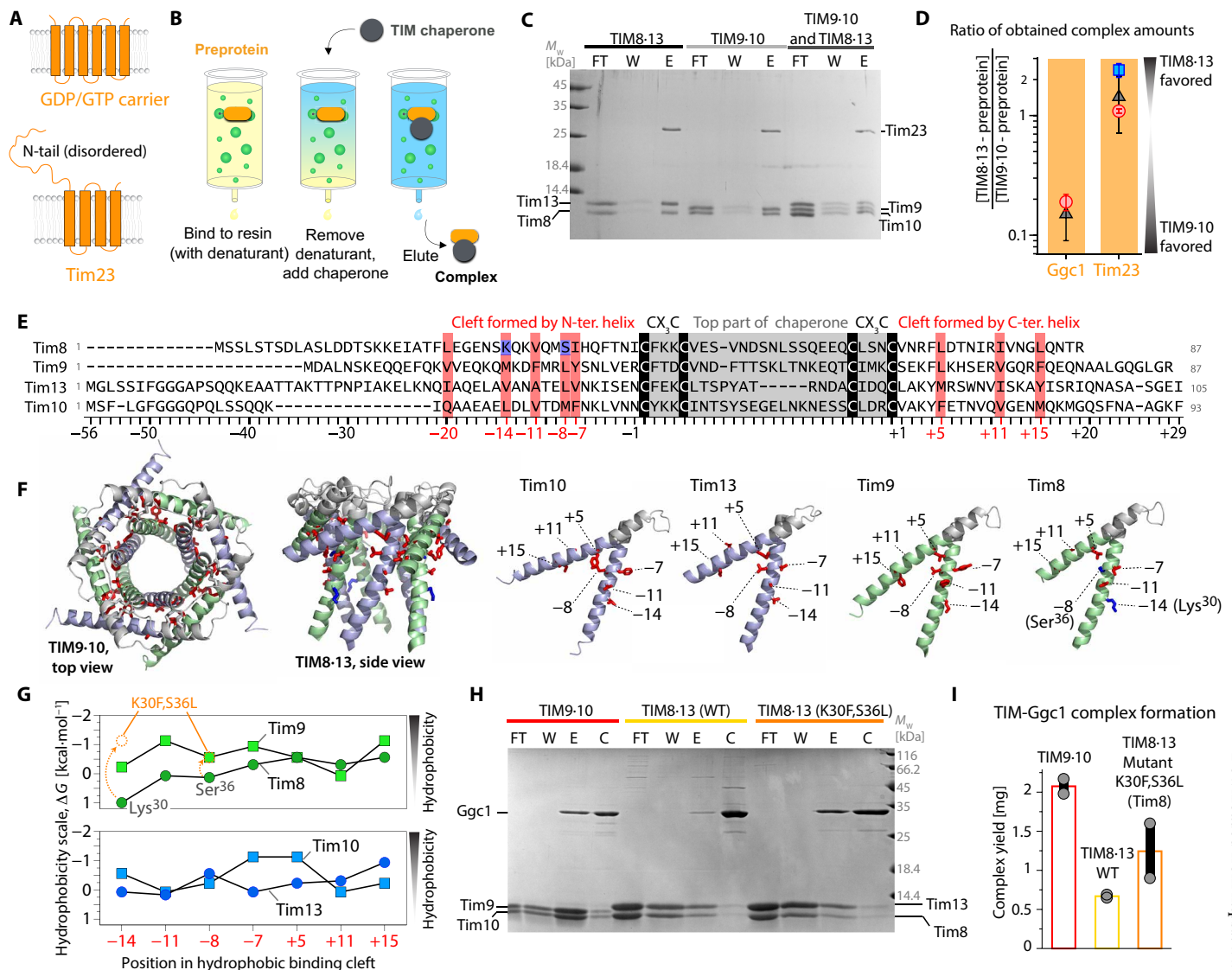
The measurement of dissociation constants of chaperones and membrane precursor proteins, using methods such as isothermal titration calorimetry (ITC) or surface plasmon resonance, is not possible, because the complexes cannot be formed in solution [e.g., flash-dilution methods, which work for other chaperones (14), failed]. Thus, to characterize the relative affinities of the precursor proteins to the two chaperones, we performed different types of competition experiments. In a first experiment, the precursor protein was bound to the affinity resin, and both chaperones were simultaneously added, before washing excess chaperone and eluting the chaperone-precursor complexes (Fig. 1C). NMR spectroscopy shows that the two chaperones do not form mixed hetero-hexameric complexes, implying that TIM9-10 and TIM8-13 stay intact in such competition experiments (fig. S3). In a second class of experiments, we prepared one type of complex (e.g., TIM9-10-Tim23) and added the other chaperone (e.g., TIM8-13) in its apo state, allowing the precursor protein to be transferred. These experiments also demonstrate that membrane precursor proteins can be transferred between these two chaperones on the time scale that we investigated (minutes to hours). We used SDS-polyacrylamide gel electrophoresis (PAGE) analyses and electrospray ionization mass spectrometry (ESI-MS) to systematically quantify the amount of obtained complexes (Fig. 1D and fig. S4). Consistently, we find that Ggc1 has a strong preference for TIM9-10 (ca. 5- to 10-fold), while Tim23 shows a slight preference for TIM8-13 (ca. 1.5-fold).

Together, we established that the two chaperones bind with different affinities to two inner membrane precursor proteins, whereby TIM8-13 is barely able to hold Ggc1, in contrast to TIM9-10, while it can hold Tim23 slightly better than TIM9-10.

### The small TIM chaperones use a conserved hydrophobic cleft for membrane precursor protein binding

To understand the different binding properties, we performed a sequence alignment of the small TIMs, which reveals a well conserved set of hydrophobic residues that point toward the binding cleft formed between the inner (N terminus) and outer tentacles (Fig. 1, E and F) (23). The overall hydrophobicity of these residues is lower in Tim8 and Tim13 than in Tim9 and Tim10 (Fig. 1G). In particular, Tim8 has a charged residue in position  $-14$  (Lys<sup>30</sup>). (The sequences are numbered starting with negative numbering at the twin CX<sub>3</sub>C motif toward the N terminus and positive numbering from the last Cys to the C terminus, and the hydrophobic motif residues are at positions





**Fig. 1. Biochemical characterization of TIM chaperone-membrane protein complexes.** (A) Native topology of the two precursor proteins used in this study. (B) Schematic view of the pull-down experiment used to prepare chaperone-precursor complexes. (C) Formation of Tim23-chaperone complexes, monitored by SDS-PAGE. Either TIM8-13, TIM9-10, or a 1:1 mixture was added to NiNTA-bound Tim23. The lanes correspond to flow-through after applying chaperone (FT), additional wash (W), and imidazole elution (E). Protein bands corresponding to Tim10 and Tim8 overlap.  $M_w$ , molecular weight. (D) Relative amounts of complexes with Ggc1 and Tim23, obtained from three different experiments: (i) a pull-down assay where both chaperones were applied to bound precursor protein (black); (ii) preparation of a TIM9-10-precursor protein complex and addition of TIM8-13, and SDS-PAGE and MS analysis after 1 and 3 hours (red); (iii) preparation of TIM8-13-Tim23 followed by TIM9-10 addition and SDS-PAGE (blue) as in (ii). The protein amounts were determined from LC-ESI-TOF-MS (fig. S4); error estimates from two or more experiments. (E) Sequence alignment of the small Tims, numbered from the N- and C-terminal conserved Cys residues ("0"). Red, conserved hydrophobic residues; blue, hydrophilic Tim8 residues, K30 and S36L. See fig. S5 for comprehensive alignment. (F) Location of the residues in the hydrophobic cleft. (G) Comparison of Kyte-Doolittle hydrophobicity of the residues in the binding cleft of wild-type (WT) native Tim proteins and Tim8<sub>K30F,S36L</sub>. (H) Pull-down experiment of Ggc1 with TIM9-10, TIM8-13, and TIM8-13 (Tim8<sub>K30F,S36L</sub>). Lane descriptions are as in (C); in addition, the fraction obtained after final wash with urea and imidazole, to control the Ggc1 initially loaded onto the column, is shown [control (C)]. (I) Amount of complex obtained from pull-down experiments of WT and mutant chaperones; the same amount of Ggc1 was applied in all three experiments, and the total amount of eluted complex was determined spectroscopically.

-20, -14, -11, -8, -7, +5, +11, and +15.) This positive charge at position -14, either Lys or Arg, is conserved among eukaryotes (fig. S5). In yeast, position -8 of the hydrophobic motif is polar (Ser<sup>36</sup>), although this position is not strictly conserved. Overall, these residues make the hydrophobic binding cleft of Tim8 less hydrophobic than in the other small Tims.

We speculated that the less hydrophobic nature of TIM8-13's binding cleft reduces its affinity to TM parts of membrane precursor

proteins. To test this hypothesis, we generated a mutant TIM8-13 with increased hydrophobicity (Tim8<sub>K30F,S36L</sub>; Fig. 1G). This more hydrophobic TIM8-13 (Tim8<sub>K30F,S36L</sub>) chaperone allowed us to obtain significantly larger amounts of complex with Ggc1 than native TIM8-13, under otherwise identical conditions (Fig. 1, H and I). This observation establishes the importance of the hydrophobic cleft for binding hydrophobic TM parts of precursor proteins. Equivalent experiments with the full-length (FL) Tim23, shown in

fig. S6, reveal that the additional hydrophobic residues in the binding cleft of Tim8<sub>K30F,S36L</sub> do not improve its capacity to bind Tim23. This observation suggests that the binding mechanisms in place for binding these two different precursor proteins differ.

To better understand the client-binding properties of the two chaperones, we turned to structural studies. Solution-NMR spectra of apo TIM8-13 (Fig. 2A) and residue-wise resonance assignments allowed the identification of the residues forming secondary structure and estimating their local flexibility. In agreement with the crystal structure, the core of rather rigid tentacles comprises the top part of the chaperone between the CX<sub>3</sub>C motifs and ca. 15 to 25 residues before and after these motifs. About 10 to 20 residues on each of the N and C termini are flexible (fig. S7).

To probe the binding of a TM segment of a membrane precursor protein, we performed NMR-detected titration experiments of TIM8-13 with a cyclic peptide corresponding to the two C-terminal strands of the  $\beta$  barrel voltage-dependent anion channel (VDAC<sub>257-279</sub>) that has a propensity to form a  $\beta$ -turn (34). Addition of this cyclic VDAC<sub>257-279</sub> induces chemical shift perturbations (CSPs) (Fig. 2B) that are primarily located in the hydrophobic cleft formed between the inner and outer rings of helices (Fig. 2, C and D). This binding site matches very closely the site on TIM9-10 to which VDAC<sub>257-279</sub> binds (Fig. 2C) (23). The VDAC<sub>257-279</sub>-induced CSP effects in TIM8-13 are overall only about half of the magnitude of CSPs found in TIM9-10, pointing to a higher population of the TIM9-10-VDAC<sub>257-279</sub> complex compared to TIM8-13-VDAC<sub>257-279</sub> at comparable conditions (Fig. 2C). This finding suggests a lower affinity of TIM8-13 to VDAC<sub>257-279</sub>, as expected from its lower hydrophobicity.

Photo-induced cross-linking experiments of a Bpa-modified VDAC<sub>257-279</sub> peptide to TIM8-13 show that only the cyclic peptide forms cross-linking adducts, while the linear, mostly disordered (34) form does not (Fig. 2E). The same behavior was also found for TIM9-10 (23) and yeast cytosolic chaperones Ssa1, Ydj1, Djp1, and Hsp104 (35). A rationale for this finding is the fact that in a  $\beta$ -turn the side chains of consecutive residues point to the two opposing faces, thus creating one hydrophobic and one more hydrophilic face (Fig. 2F). In contrast, because of its disorder, the linear VDAC<sub>257-279</sub> peptide does not have a stable hydrophobic face, reducing its affinity to the hydrophobic binding cleft on the chaperone. In line with these findings, NMR titration data with the linear peptide show small CSPs that are spread across the protein, thus pointing to unspecific interaction (fig. S8).

Collectively, the experiments with the client fragment VDAC<sub>257-279</sub> provide a first evidence that both chaperones use the same conserved binding cleft to interact with hydrophobic membrane precursor protein sequences, and that TIM9-10 interacts more efficiently with TM parts, and thus with Ggc1 and the VDAC fragment. We propose that the more hydrophobic nature of the binding cleft in TIM9-10 allows it to interact more strongly with TM parts of its clients. In light of this observation, how does TIM8-13 achieve a binding affinity to Tim23, which is slightly higher than the one of TIM9-10 (Fig. 1D)?

### Hydrophilic fragments interact differently with TIM8-13 and TIM9-10

Tim23 has a hydrophilic N-terminal segment in addition to four TM helices (Fig. 3A), and we investigated whether this part interacts with the chaperones. NMR spectra of the soluble Tim23<sub>IMS</sub> fragment (residues 1 to 98) in isolation show the hallmark features of a highly flexible intrinsically disordered protein with low spectral

dispersion of <sup>1</sup>H-<sup>15</sup>N NMR signals (Fig. 3, B and C, orange spectrum), as previously reported (36). Upon addition of TIM9-10, the Tim23<sub>IMS</sub> <sup>1</sup>H-<sup>15</sup>N spectrum (Fig. 3B, left) shows only small changes: All cross-peaks are still detectable, and small CSPs are only observed for a few residues at the N terminus, which has higher hydrophobicity (Fig. 3D). This finding suggests only very weak, possibly nonspecific interactions between the very N terminus of Tim23<sub>IMS</sub> and TIM9-10. In line with this finding, the interaction is not detectable by ITC measurements (Fig. 3E).

The interaction of the hydrophilic Tim23<sub>IMS</sub> fragment with TIM8-13 is significantly stronger, with pronounced binding effects detected by ITC, and a dissociation constant of  $K_d = 66 \pm 8 \mu\text{M}$  (Fig. 3E, right; see table S1). The <sup>1</sup>H-<sup>15</sup>N NMR spectrum of Tim23<sub>IMS</sub> in the presence of TIM8-13 shows strongly reduced peak intensities for most of the residues (Fig. 3C, left). Such a peak broadening is expected when a highly flexible polypeptide binds to a relatively large object such as TIM8-13, thereby inducing faster nuclear spin relaxation and thus broader signals of lower intensity. Analysis of the peak intensity reduction reveals two regions of Tim23 that are particularly involved in the binding: (i) the N-terminal hydrophobic residues, which are also involved in interacting with TIM9-10, and (ii) a long sequence stretch comprising residues from ca. 30 to 80 (Fig. 3F).

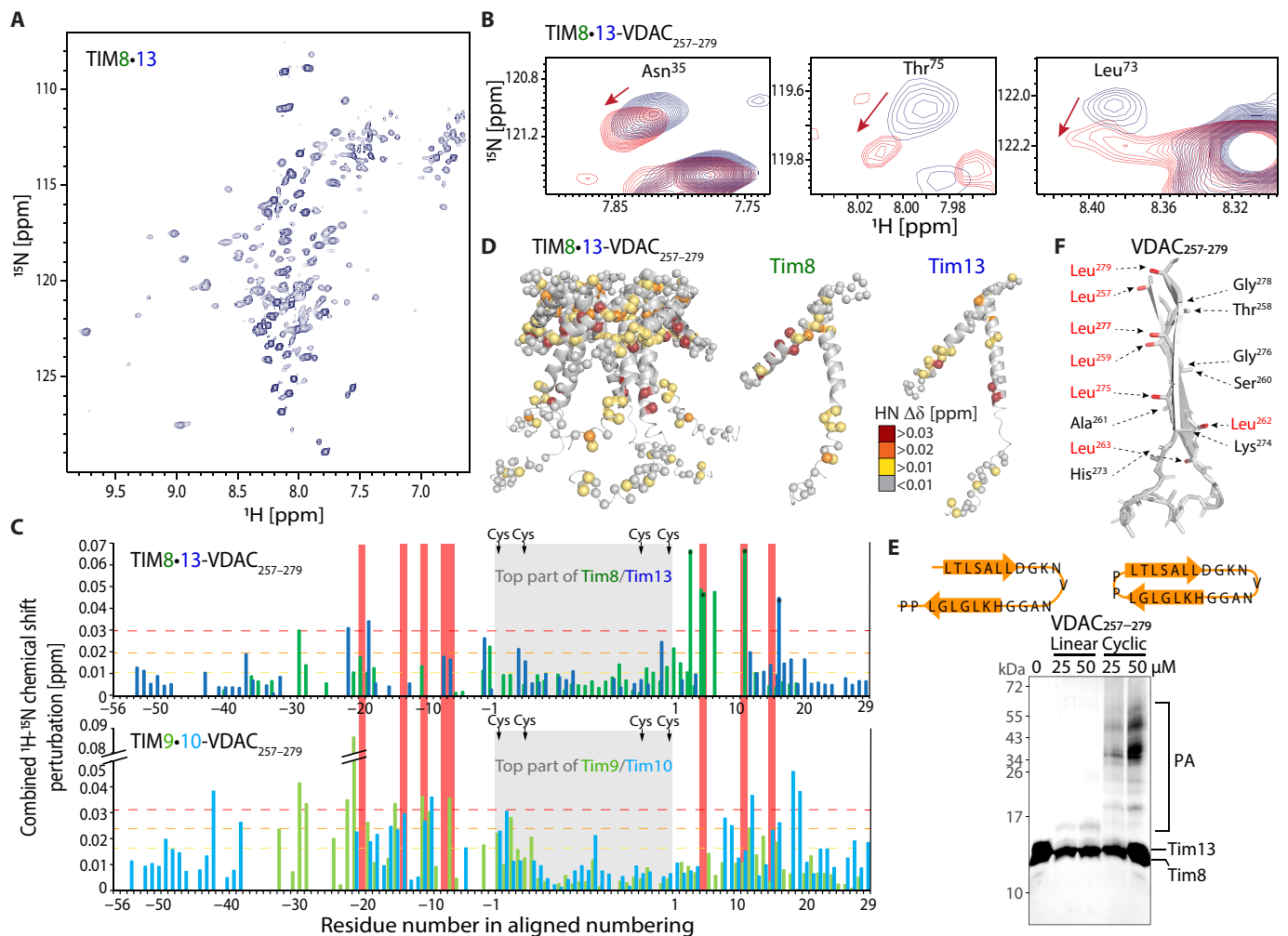
To investigate whether TIM8-13 may interact with another soluble protein from the IMS, we performed ITC experiments with the globular protein cytochrome c. No interaction could be detected (fig. S9), suggesting that the TIM8-13-Tim23<sub>IMS</sub> interaction may be related to the unfolded, flexible character of the latter.

To characterize the conformation of FL Tim23 bound to TIM8-13 and TIM9-10, we prepared Tim23<sub>FL</sub>-labeled Tim23-chaperone complexes using the method outlined in Fig. 1B. Very similar to the experiments with the Tim23<sub>IMS</sub> fragment, the signals corresponding to the N-terminal half of Tim23 are still intense in the Tim23<sub>FL</sub>-TIM9-10 complex (Fig. 3, B and G), revealing that the N-terminal half of Tim23<sub>FL</sub> does not interact strongly with TIM9-10. The small observed CSPs are localized primarily at the hydrophobic N terminus. In contrast, when Tim23<sub>FL</sub> is bound to TIM8-13, the signals corresponding to its N-terminal half are severely reduced in intensity, revealing tight contact of the flexible N-terminal half of Tim23 to TIM8-13 (Fig. 3, C and H).

In neither of the two Tim23<sub>FL</sub> complexes, any additional signals, which may correspond to Tim23's TM helices, are visible. We ascribe this lack of detectable signals of residues in the TM part to extensive line broadening. The origin of this line broadening may be ascribed to the large size of the complex and likely to additional millisecond time scale dynamics of Tim23's TM parts in the hydrophobic binding cleft of the chaperones. Such millisecond motions have been found in the TIM9-10-Ggc1 complex (23).

### TIM8-13 uses an additional hydrophilic face for protein binding

We probed the binding sites that the chaperones use to interact with Tim23<sub>IMS</sub> or Tim23<sub>FL</sub> using NMR spectroscopy on samples, in which only the chaperone was isotope-labeled. The CSPs in the two chaperones upon addition of Tim23<sub>IMS</sub> reveal distinct binding patterns (Fig. 4A): In TIM8-13, the largest effects involve residues in the hydrophilic top part of the chaperone, between the CX<sub>3</sub>C motifs, as well as a few residues toward the C-terminal outer ring of helices; in contrast, the corresponding top part of TIM9-10 does not show any significant effects, but CSPs are observed at residues in the



**Fig. 2. Solution-NMR and binding of a VDAC fragment to TIM8-13.** (A)  $^1\text{H}$ - $^{15}\text{N}$  NMR spectrum of TIM8-13 at 35°C. (B) CSP in TIM8-13 upon addition of five molar equivalents of cyclic VDAC<sub>257-279</sub>. (C) CSP effects of VDAC<sub>257-279</sub> binding. The data for TIM9-10 are from (23). (D) Plot of CSP data on the TIM8-13 structure. (E) Photo-induced cross-linking of the linear (left) and cyclic (right) VDAC<sub>257-279</sub> peptides to TIM8-13. While hardly any adducts are observed for the linear one, the cyclic peptide forms cross-linking photo-adducts (PA), including of higher molecular weight, resulting from multiple cross-links, as reported earlier (23, 34, 35). (F) Schematic structure of the two last strands of VDAC, as found in the NMR structure (61) of the full  $\beta$  barrel, showing that the hydrophobic and hydrophilic side chains cluster on the two opposite faces of the  $\beta$ -turn.

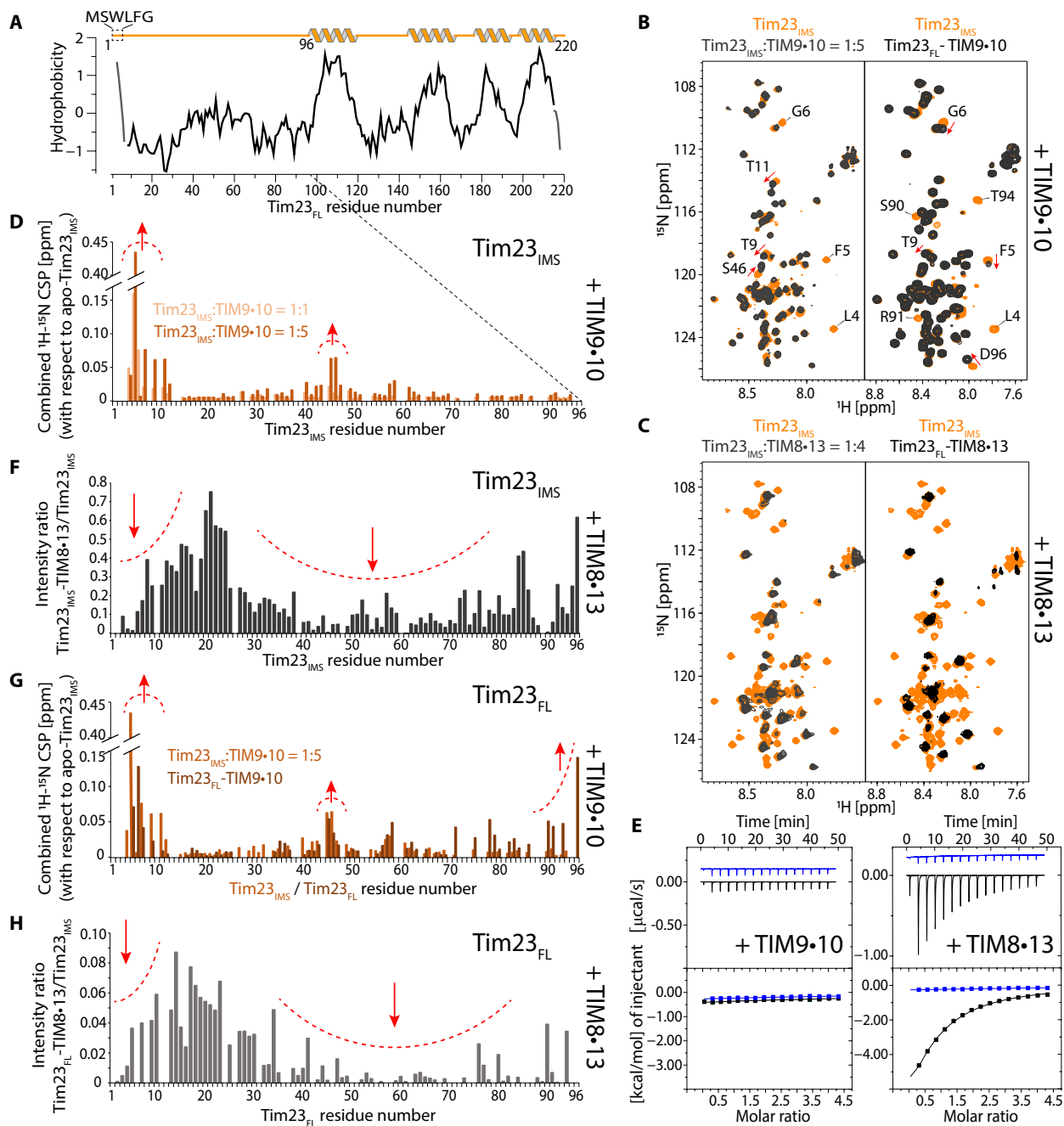
hydrophobic binding cleft and, in particular, the N-terminal helix (Fig. 4, B and C). These data, together with the Tim23<sub>TM</sub>-detected data in Fig. 3, establish that TIM8-13 uses its hydrophilic top part to bind Tim23's N-terminal half, while only a short stretch of hydrophobic residues at the very N terminus of Tim23 interacts with the hydrophobic cleft of TIM9-10, which is also the binding site of TM parts (Figs. 1 and 2).

Chaperone-labeled complexes with Tim23<sub>FL</sub> confirm these findings and point to the additional effects induced by the bound TM part: In TIM9-10-Tim23<sub>FL</sub>, large CSP effects are located primarily in the binding cleft, in line with the view that the top part of TIM9-10 is not involved in binding Tim23. In contrast, Tim23<sub>FL</sub>-induced CSPs are found across the whole TIM8-13, including the hydrophilic top and the hydrophobic cleft (Fig. 4, D and E, and fig. S10).

We furthermore prepared complexes of a truncated Tim23 fragment (Tim23<sub>TM</sub>, residues 92 to 222), which allows to selectively detect the interaction of the TM part with the chaperones. The TIM9-10-Tim23<sub>TM</sub> complex features the largest CSPs in the hydrophobic

binding cleft, qualitatively similar to the binding site detected with Tim23<sub>FL</sub> (Fig. 4F and fig. S11). The complex of Tim23<sub>TM</sub> with TIM8-13 appears to be much less stable than TIM9-10-Tim23<sub>TM</sub>: In the pull-down experiment, only a very small amount of complex could be obtained, and the complex rapidly precipitated (not shown), excluding NMR analyses. This observation reflects that the hydrophobic cleft of TIM8-13 is less capable of holding a hydrophobic polypeptide than the one of TIM9-10.

Collectively, NMR, ITC, and mutagenesis have revealed that the hydrophobic cleft of both TIM8-13 and TIM9-10 is essential to hold the hydrophobic parts of the clients and that TIM8-13, but not TIM9-10, additionally interacts with the hydrophilic part of Tim23 to increase its affinity. This interaction, which is mediated by the hydrophilic top part of TIM8-13, reduces the conformational flexibility of Tim23's N-terminal half. The observation that the interaction is driven by hydrophilic contacts supports previous findings of the protein import (25): TIM8-13 was found to interact with



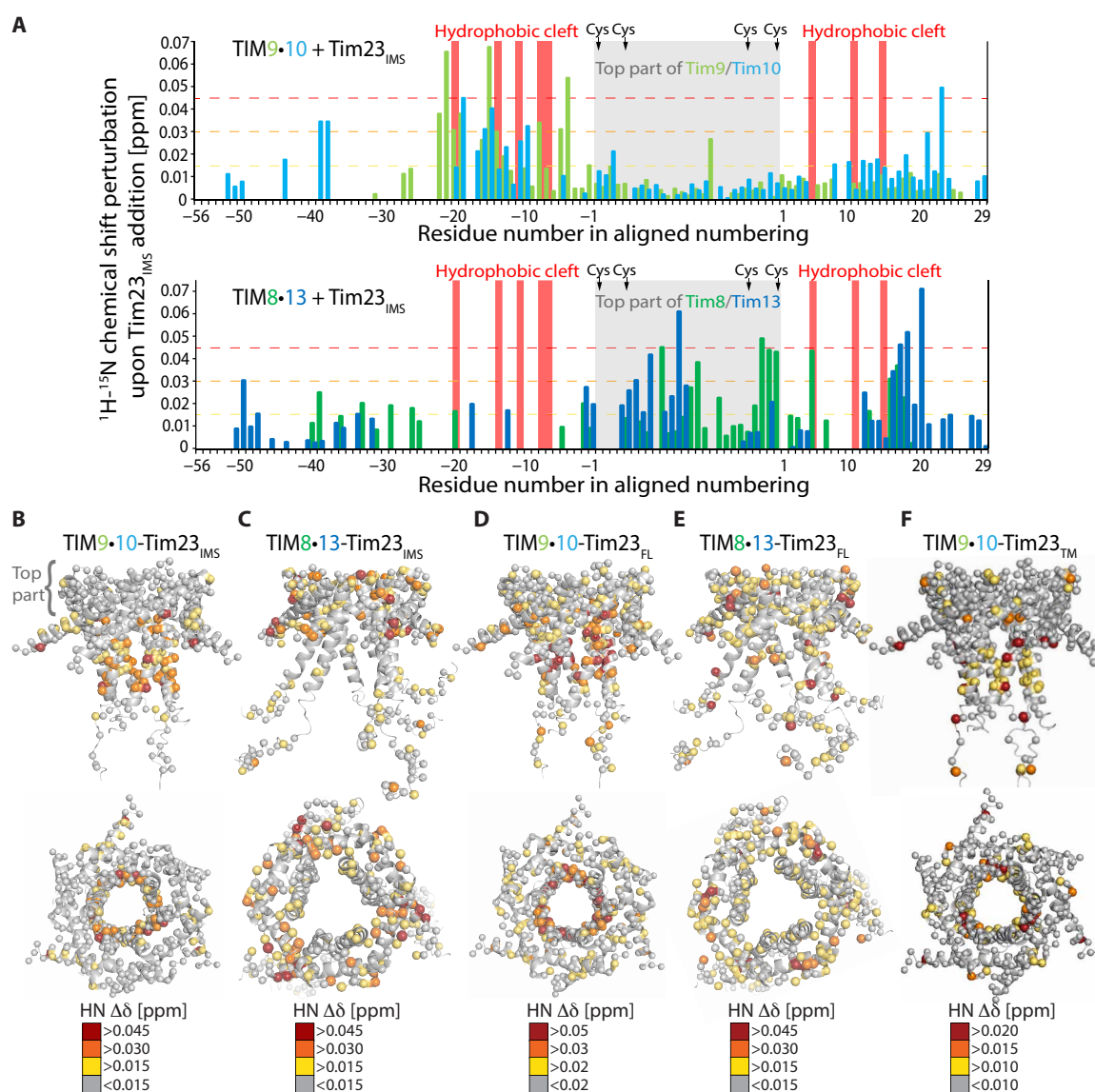
**Fig. 3. Tim23 has markedly different properties when binding to TIM8-13 and to TIM9-10.** (A) Hydrophobicity of Tim23 (Kyte-Doolittle). (B) NMR spectra of the <sup>15</sup>N-labeled soluble Tim23<sub>IMS</sub> fragment in the presence of TIM9-10 (left, black) and of FL Tim23 bound to TIM9-10 (right, black) are compared to the Tim23<sub>IMS</sub> fragment in isolation (orange), under identical buffer conditions and NMR parameters. (C) As in (B) but with TIM8-13 instead of TIM9-10. (D) CSP of residues in Tim23<sub>IMS</sub> upon addition of one (light orange) or five (dark orange) molar equivalents of TIM9-10. (E) Calorimetric titrations for the interaction of TIM9-10 or TIM8-13 (54 μM in the calorimetric cell) with Tim23<sub>IMS</sub> (1.15 mM in the injecting syringe). Thermograms are displayed in the upper plots, and binding isotherms (ligand-normalized heat effects per injection as a function of the molar ratio, [Tim23<sub>IMS</sub>]/[chaperone]) are displayed in the lower plots. Control experiments, injecting into a buffer, are shown in blue. (F) Intensity ratio of residues in Tim23<sub>IMS</sub> in the presence of four molar equivalents of TIM8-13 compared to Tim23<sub>IMS</sub> alone. (G) CSP of the detectable residues in FL Tim23 attached to TIM9-10 (brown), compared to the soluble Tim23<sub>IMS</sub> fragment. (H) Intensity ratio of detectable residues in Tim23<sub>FL</sub> attached to TIM8-13. Note that the ratio was not corrected for differences in sample concentration, and the scale cannot be compared to the one in (G).

hydrophobic membrane precursor only when they were fused to the hydrophilic Tim23<sub>IMS</sub> part.

We have also investigated whether a given FL Tim23 chain may interact simultaneously with TIM9-10 and TIM8-13, using hydrophobic

and hydrophilic interactions, respectively, to form a ternary complex. However, samples containing all three components do not contain detectable amounts of such complexes, and we conclude that the affinity is too low to simultaneously bind two chaperones (fig. S12).





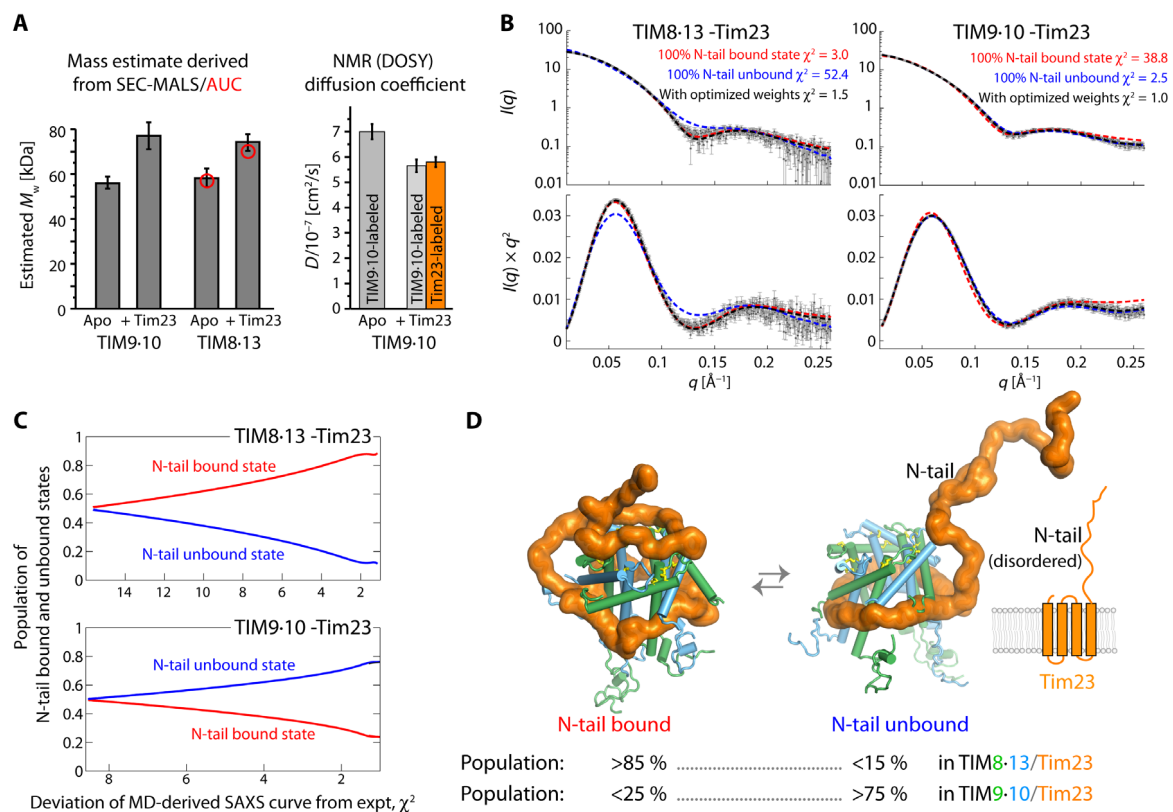
**Fig. 4. Tim23<sub>IMS</sub> and FL Tim23 differ in their interactions with TIM9-10 and TIM8-13 chaperones.** (A) CSPs observed upon addition of the Tim23<sub>IMS</sub> fragment to TIM9-10 (top) and TIM8-13 (bottom). The chaperone:Tim23<sub>IMS</sub> ratios were 1:1 (TIM8-13) and 1:3 (TIM9-10). Mapping of Tim23<sub>IMS</sub>-induced CSPs on TIM9-10 (B) and TIM8-13 (C), showing that while the top part of TIM9-10 does not show any significant CSPs, the corresponding part is the main interacting region of TIM8-13. CSP in complexes of TIM9-10 (D) and TIM8-13 (E) bound to FL Tim23. Tim23<sub>TM</sub>-induced CSP mapped on TIM9-10 (F), showing similar binding as the FL Tim23.

### Structural ensembles of chaperone-Tim23 complexes

We integrated the NMR data with further biophysical, structural, and numerical techniques to obtain a full structural and dynamical description of the complexes. We first investigated the complex stoichiometry using size exclusion chromatography coupled to multiangle light scattering (SEC-MALS), NMR-detected diffusion coefficient measurements, and analytical ultracentrifugation (AUC). These methods, which provide estimates of molecular mass (and shape) from orthogonal physical properties (gel filtration and light scattering; translational diffusion), reveal properties best compatible with a 1:1 (chaperone:precursor) stoichiometry (Fig. 5A and fig. S13). Small-angle x-ray scattering (SAXS) data of both TIM9-10-Tim23 and TIM8-13-Tim23 also point to a molecular weight corresponding to a 1:1 complex (SAXS; Fig. 5B). This stoichiometry contrasts

the 2:1 (chaperone:precursor) stoichiometry for TIM9-10 holding the 35-kDa large carrier Ggc1 (fig. S2) (23).

SAXS provides significantly more information, namely, the overall shape of the ensemble of conformations present in solution. Given the flexibility of the complex, these SAXS data are best analyzed by considering explicitly a dynamic ensemble. We used molecular dynamics (MD) simulations to account for the breadth of possible conformations that, collectively, result in the observed scattering. To effectively sample the conformational space of the chaperone-Tim23 complex, we constructed two distinct structural models, in which the N-terminal half of Tim23 is either modeled as a floppy unstructured tail or bound to the hydrophilic upper part of chaperone, denoted as “N-tail unbound” and “N-tail bound” conformations, respectively. In both models, the hydrophobic C-terminal TM



**Fig. 5. Architecture of the TIM8-13 and TIM9-10 holdases in complex with FL Tim23.** (A) Left: Apparent molecular weights of apo and holo chaperone complexes from SEC-MALS and AUC (red circles). Right: Translational diffusion coefficients of TIM9-10 (apo) and TIM9-10-Tim23<sub>FL</sub> from NMR DOSY measurements. Two independent samples were used for the complex, in which either the chaperone or the precursor protein was labeled, as indicated. See also fig. S13. (B) Small-angle x-ray scattering (SAXS) curves (top) and Kratky plot representations thereof for the two chaperone-precursor complexes. The lines are SAXS curves calculated from structural ensembles obtained over 4.25- $\mu$ s-long MD trajectories, in which the N-terminal half of Tim23 was either in a conformation bound to the top part of the chaperone (red) or in a loose unbound conformation (blue), or from an ensemble in which these two classes of states were present with optimized weights. (C) Goodness of fit of the back-calculated SAXS curves to the experimental SAXS data as a function of the relative weights of the two classes of conformations (bound/unbound). (D) Snapshots of conformations in which Tim23<sub>N-tail</sub> is either bound or unbound and the best-fit relative weights of the two classes of states as derived from SAXS/MD. More SAXS/MD data and ensemble views are provided in fig. S14 and in movies S1 and S2.

domain of Tim23 is bound to the hydrophobic cleft of the chaperone, as identified by NMR (Fig. 4 and fig. S10, C and D) (23). Initiating from both conformations, explicit-solvent atomistic MD simulations ( $\sim 4.25$   $\mu$ s in total) were performed to collect the structures for the N-tail unbound and N-tail bound ensembles. In the case of TIM8-13, the N-tail bound ensemble recapitulates the experimentally observed pattern better than the N-tail unbound ensemble (Fig. 3, C and E). We then constructed a mixed ensemble consisting of a mixture of N-tail bound and N-tail unbound states. We used this pool of conformations for further ensemble refinement, with the relative populations of these two ensembles of states as free parameter, using the Bayesian maximum entropy (BME) method guided by experimental SAXS data (23, 37, 38). We found that the experimental SAXS data of TIM8-13-Tim23 are very well reproduced when the mixed ensemble has >85% of the N-tail bound state (Fig. 5, C and D). In contrast, the experimental data of TIM9-10-Tim23 are only well reproduced when the TIM9-10-Tim23 ensemble comprises predominantly the N-tail unbound state. These refined ensembles guided by experimental SAXS data are in excellent agreement with the NMR data, which showed that (i) in the TIM9-10-Tim23 complex, the N-terminal part of Tim23 is predominantly free and flexible,

and Tim23 makes contacts only to the hydrophobic cleft of the chaperone, while (ii) in TIM8-13-Tim23, the Tim23<sub>IMS</sub> part is largely bound to the upper part of the chaperone (Figs. 3 and 4).

The amount of N-tail bound relative to N-tail unbound states is expected to depend on the affinity of the N-tail of Tim23 to the chaperone. The ITC-derived TIM8-13-Tim23<sub>IMS</sub> affinity ( $K_d = 66$   $\mu$ M; Fig. 3H) predicts that the population of N-tail bound states is of the order of 75 to 98% (see Methods for details), in excellent agreement with the MD/SAXS-derived value (>85%). This good match of data from the Tim23<sub>IMS</sub> fragment and Tim23<sub>FL</sub> suggests that the binding of Tim23's hydrophilic N-tail does not strongly depend on the presence of the TM part. The low affinity of the N-tail to TIM9-10, reflected by the inability to detect TIM9-10-Tim23<sub>IMS</sub> binding by ITC, is mirrored by the small population of the N-tail bound states in the FL complex.

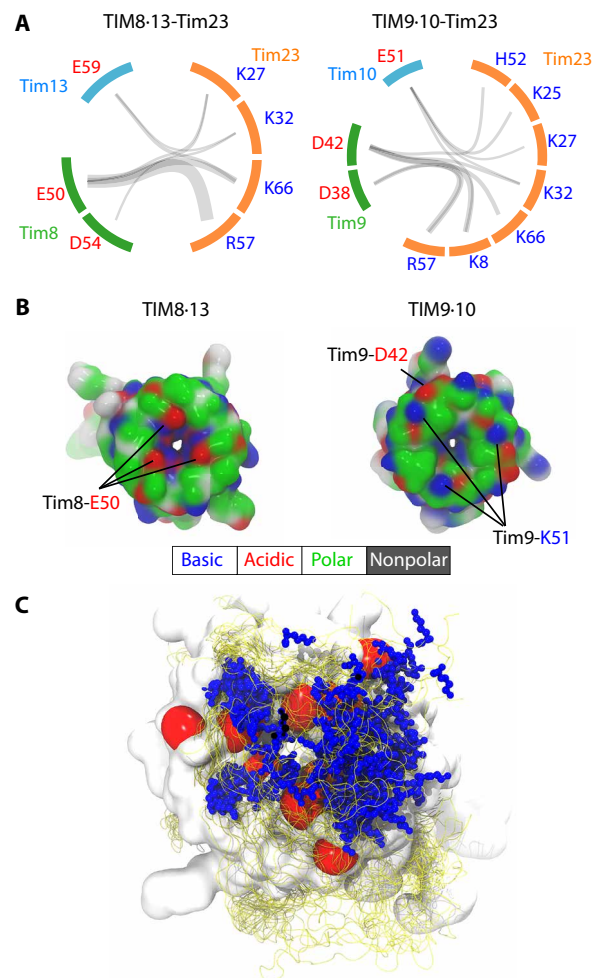
To identify the molecular mechanisms underlying the observed differences in N-tail binding, we studied the interactions formed between Tim23 and the chaperones along the MD simulation. An interesting pattern emerges from the analysis of the electrostatic interactions. The top part of TIM8-13 has predominantly polar and negatively charged residues, which are in transient contact with the

positive charges of Tim23 N-tail, within a dynamic ensemble of conformations (Fig. 6). For example, three key aspartate or glutamate residues in TIM8-13 appear to be involved in binding of lysine or arginine residues of Tim23<sub>IMS</sub> (Fig. 6A). In Tim9, a lysine (K51) is present in the top part and contributes a positive charge (the equivalent position in TIM8-13 is a noncharged, polar residue) (Fig. 6B). We hypothesized that the less complementary electrostatic properties of TIM9-10's top part and Tim23's N-tail, as compared to TIM8-13, may diminish the affinity of the N-tail to TIM9-10. We attempted to investigate the importance of these charged residues experimentally and prepared single and double mutants that invert the pattern of charged residues. In TIM8-13, we introduced lysine or arginine instead of negatively charged residues, expecting to thereby reduce the affinity to Tim23<sub>IMS</sub>; conversely, in TIM9-10, we introduced negative charges to promote the Tim23<sub>IMS</sub> interaction. However, ITC experiments show that most of these mutants do not significantly differ in their binding affinity to Tim23<sub>IMS</sub> (fig. S9 and table S1). In one of the TIM8-13 mutants, the binding affinity even increases, despite the additional positive charge in the chaperone. These findings suggest that due to the disordered nature of Tim23's N-tail, its binding with the chaperones might not be dominated by a few strong interactions but instead be contributed by a complex interaction network with many weak and widely distributed interactions, which are tolerant to introduction of the single point mutants that we explored. The MD ensemble (Fig. 6C), due to its limited time scale and force field imperfections (39), may only be able to identify a rather small number of key interaction sites. Unlike the case of this hydrophilic interaction, we were able to identify several single-point mutations in the TIM9-10 hydrophobic motif that abrogate the binding, with a strong phenotype (23).

## DISCUSSION

Transfer chaperones (holdases) need to fulfill two contradicting requirements, holding their clients very tightly to avoid their premature release and aggregation while, at the same time, allowing release at the downstream factor. This apparent contradiction is solved by a subtle balance of multiple individually weak interactions and a resulting dynamic complex, wherein the precursor protein samples a wide range of different conformations. This ensemble of conformations results in a high overall affinity; yet, a downstream foldase/insertase can detach the precursor protein from the chaperone without significant energy barrier (40). Balancing the interaction strengths is, thus, crucial to chaperone function. Here, we have revealed a fine-tuning of chaperone-client specificity that involves hydrophobic interactions with the chaperone's binding cleft and additional hydrophilic interactions, mostly mediated by charged residues, with the chaperone's top part. Lower hydrophobicity within the binding cleft of TIM8-13 compared to TIM9-10 arises by overall less hydrophobic residues and a positively charged residue (Lys/Arg) that is highly conserved in Tim8. As a consequence, TIM8-13 is less able to hold the TM parts of its clients than TIM9-10 by ca. one order of magnitude. As we showed, replacement of two charged/polar side chains in TIM8-13's cleft brings TIM8-13 to a similar level as TIM9-10 for holding an all-TM client.

For binding of its native client Tim23, TIM8-13 uses additional hydrophilic interactions to its client's IMS segment, which is ineffective in the TIM9-10-Tim23 interaction. The additional interaction effectively compensates for the lower affinity of TIM8-13 to the



**Fig. 6. Tentative identification of electrostatic interactions from the MD ensemble.** (A) The charged residue pairs forming salt bridges are connected by gray semitransparent lines whose thickness linearly scales with the frequency of the corresponding salt bridge observed in MD simulations. Although more diverse salt bridges were observed in TIM9-10-Tim23 (10 in TIM9-10-Tim23 and 7 in TIM8-13-Tim23), these salt bridges were, on average, less stable than the ones in TIM8-13-Tim23, likely resulting in overall weaker interactions. (B) Snapshots of top views of the two chaperones along MD simulations of their holo forms in complex with TIM23. The top views of the chaperones in the apo forms are shown in fig. S14 (E and F). Residues are color-coded according to the scheme reflected below the figure. (C) Ensemble view of the N-tail bound state of TIM8-13-Tim23. The red surface represents the negatively charged E59 of Tim13 and E50 and D54 of Tim8. Blue stick and ball represents the side chain of positively charged residues (K8, K25, K27, K32, R57, and K66) of Tim23, which is shown as an ensemble of 25 structures. Ensemble view of the N-tail bound state of TIM9-10-Tim23 is shown in fig. S14G.

client's TM part. In the case of Tim23, this additional interaction involves a sequence stretch of at least 35 to 40 residues (Fig. 3, E and G). TIM8-13 has also been shown to be involved in the transport of a Ca<sup>2+</sup>-regulated mitochondrial carrier, the Asp/Glu carrier (31), which has an additional soluble calmodulin-like domain. Whether this soluble domain is folded or disordered while the TM domain is attached to the hydrophobic chaperone cleft remains to be investigated. It is tempting to speculate that interactions between TIM8-13's top part and the calmodulin-like hydrophilic part of these carriers are important for this binding, similarly as for the case of Tim23 (fig. S1).



Membrane precursor proteins that have been shown not to interact with TIM8-13, such as mitochondrial carriers (Ggc and Aac) and Tim17, lack extended hydrophilic stretches, underlining the importance of those parts in binding (fig. S1). From the sequences of known clients and known “nonclients” of TIM8-13, we propose that a minimum sequence length of about 20 to 25 residues is required for binding.

The nature of these additional hydrophilic interactions appears to involve primarily charged residues that form a complex and wide-connected interaction network that could be hard to suppress by mutating individual sites.

This study provides a rationale why mitochondria contain two very similar IMS chaperone complexes, the essential TIM9-10 and the nonessential TIM8-13 complex. The observation that this dual system is conserved even in humans suggests that the presence of the TIM8-13 system is not just the result of gene duplication, which appears rather often in yeast. The current results propose that for some substrates (like Tim23, or Asp-Glu carrier; see fig. S1), TIM8-13 can contribute stabilizing interactions with the hydrophilic soluble parts. Our competition experiments have also revealed that mitochondrial membrane precursor proteins may be transferred from one TIM chaperone to the other, opening the possibility that these two chaperones truly cooperate in precursor protein transfer to downstream insertases.

Together, our study reveals how a subtle balance of hydrophobic and hydrophilic interactions is used to tune promiscuity versus specificity in molecular chaperones. We propose that a similar balance of interactions determines the clientome of the cellular chaperones.

## METHODS

### Plasmids

Genes coding for *Saccharomyces cerevisiae* Tim8 and Tim13 were cloned in the coexpression plasmid pETDuet1. The expressed protein sequences were MSSLSTDLASLDDTSKKEIATFLEGENSKQKVQM-SIHQFTNICFKKCVESVNDNLSSQEEQCLSNVCVNRFLDT-NIRIVNGLQNTNTR (Tim8) and MGSSHHHHHSQDPSQDPEN-LYFQGGSSIFGGGAPSQQKEAATTAKTTPNPIAKELKNQI-AQELAVANATELVNKISENCFEKLTSPYATRNDACIDQCLA-KYMRSWNVISKAYISRIQNASASGEI (Tim13). A tobacco etch virus (TEV) cleavage site on Tim13 allows the generation of the final construct starting with GGLSS (the native Tim13 sequence starts with MGLSS). The same approach was used for preparing TIM9-10, including coexpression of the two proteins, with a cleavable His<sub>6</sub>-tag on one of the proteins (Tim10), as described elsewhere (23). The gene coding for FL *S. cerevisiae* Tim23 (C98S, C209S, and C213A) with a C-terminal His<sub>6</sub>-tag was cloned in the expression plasmid pET21b(+). The plasmid for expression of the intrinsically disordered N-terminal domain of *S. cerevisiae* Tim23<sub>IMS</sub> (residues 1 to 98) with an N-terminal glutathione *S*-transferase (GST) tag is described in (36), and the pET10N plasmid encoding Tim23<sub>TM</sub> is described in (41). The *S. cerevisiae* Ggc1 (C222S) construct was designed with a C-terminal His<sub>6</sub>-tag in pET21a expression plasmid, reported earlier (23).

### Protein expression and purification

We found that TIM9-10 and TIM8-13 chaperone complexes can be obtained by overexpression in either SHuffle T7 or BL21(DE3) *Escherichia coli* cells. Expression in the former results in soluble

protein with correctly formed disulfide bonds, while the latter requires refolding from inclusion bodies. The proteins obtained with either method have indistinguishable properties (SEC, NMR). For TIM9-10, SHuffle expression results in better yield, while we obtain higher TIM8-13 yields with refolding from BL21(DE3). Accordingly, TIM9-10 and unlabeled TIM8-13 were overexpressed in the SHuffle T7 *E. coli* cells and purified as described previously (23). Overexpression of the isotope-labeled TIM8-13 chaperone complex from the BL21(DE3) *E. coli* cells was induced with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG), and the cells were incubated for 4 hours at 37°C. Cell pellets were sonicated, and the inclusion body fraction was resuspended sequentially, first in buffer A [50 mM tris(tris(hydroxymethyl)aminomethane), 150 mM NaCl (pH 7.4)] supplemented with 1% lauryldimethylamine oxide (LDAO) and 1% Triton X-100, then in buffer A supplemented with 1 M NaCl and 1 M urea, and, lastly, in buffer B [50 mM tris, 250 mM NaCl (pH 8.5)]. The last pellet fraction was solubilized in buffer B supplemented with 50 mM dithiothreitol and 3 M guanidine-HCl at 4°C overnight. The TIM8-13 complex was refolded by rapid dilution in buffer B containing 5 mM glutathione and 0.5 mM glutathione disulfide. The complex was purified on a NiNTA affinity column, and the affinity tag was removed with TEV protease and an additional NiNTA purification step. FL precursor proteins, Tim23 and Ggc1, were expressed as inclusion bodies from BL21(DE3) cells, at 37°C during 1.5 and 3 hours, respectively, after adding 1 mM IPTG. Precursor proteins were solubilized in buffer A supplemented with 4 M guanidine-HCl for Tim23 and 6 M guanidine-HCl for Ggc1 at 4°C overnight. Precursor proteins were purified by affinity chromatography in the same denaturing conditions used for solubilization. Imidazole was removed from the precursor protein sample with dialysis in buffer A supplemented with 4 M guanidine-HCl.

GST-tagged Tim23<sub>IMS</sub> was expressed in the soluble protein fraction from BL21(DE3)Ril+ cells during 4 hours at 25°C, after adding 1 mM IPTG. After sonication of the cell pellets, the soluble protein fraction was incubated with glutathione-agarose resin for 2 hours at 4°C. After washing the unspecifically bound proteins with 10 column volumes (CVs) of buffer A, the GST-tag was cleaved from the Tim23<sub>IMS</sub> by incubating the resin with 1 mg of TEV protease per 50 mg of the precursor protein, at 4°C overnight. Cleaved Tim23<sub>IMS</sub> and the protease were collected in the flow-through, and an additional NiNTA purification step was applied to remove the TEV protease from the protein sample. Soluble Tim23<sub>IMS</sub> was subjected to gel filtration on a Superdex 75 10/300 column and stored in buffer A.

Tim23<sub>TM</sub>, comprising residues 92 to 222 (41), was produced in *E. coli* BL21(DE3)Ril+ during 3 hours at 37°C and purified in denaturing conditions as described for the FL Tim23. Chaperone proteins used for detection by NMR experiments were expressed in D<sub>2</sub>O M9 minimal medium and either labeled with <sup>15</sup>NH<sub>4</sub>Cl (1 g/liter) and D-[<sup>2</sup>H,<sup>13</sup>C]glucose (2 g/liter) or specifically labeled on isoleucine, alanine, leucine, and valine side chains using a QLAM-A  $\beta$ <sup>1</sup> $\delta$ <sup>1</sup>L<sup>PRO</sup>R<sup>PRO</sup>R kit from NMR-Bio (www.nmr-bio.com) according to the manufacturer's instructions. The proteins not detected by NMR in complex samples (i.e., the precursor proteins in complexes directed toward chaperone detection or the chaperone in preprotein-detected experiments) were unlabeled and produced in LB medium. Chaperone-bound Tim23<sub>FL</sub> was deuterated (produced in D<sub>2</sub>O M9 medium), while the Tim23<sub>IMS</sub> fragment was prepared in H<sub>2</sub>O M9 medium.

The fragments of human VDAC1 peptide (cyclic or linear VDAC<sub>257-279</sub>) were prepared by solid-phase synthesis as described elsewhere (34),

lyophilized, and resolubilized first in dimethyl sulfoxide (DMSO) and then stepwise diluted into buffer, as described elsewhere (23). The peptide used for photo-induced cross-linking differed from the one used for NMR by the substitution of L263 by a Bpa side chain, as used earlier (23, 34).

### Preparation of chaperone-precursor protein complexes

Purified precursor protein, i.e., either FL Tim23, the TM Tim23<sup>TM</sup> fragment, or Ggc1, was bound to NiNTA resin in 4 M guanidine-HCl. The column was washed with five CVs of buffer A supplemented with 4 M guanidine-HCl and with five CVs of buffer A. A twofold excess of the chaperone complex was passed through the column twice. The column was washed with 10 CVs of buffer A, and the precursor-chaperone complex was eluted in 5 CVs of buffer A supplemented with 300 mM imidazole. The precursor-chaperone complex was immediately subjected to dialysis against buffer A before concentrating on Amicon 30-kDa molecular weight cut-off (MWCO) centrifugal filters (1000g). Immediate removal of imidazole was particularly important for the preparation of the less stable Tim23<sup>FL</sup>-TIM9-10 and Tim23<sup>TM</sup>-TIM8-13 complex. Complexes of Tim23<sup>IMS</sup> with TIM8-13 or TIM9-10 were prepared by mixing two purified protein samples and dialysis against buffer A. Formation of the precursor-chaperone complex was verified by SEC on a Superdex 200 column. The resulting complex was further characterized by SEC-MALS. TIM8-13 and TIM8-13-Tim23 were analyzed by AUC. Both experiments were performed at 10°C in buffer A. The amount of eluted complex was estimated from the protein concentration, measured absorbance of the sample at 280 nm, and the sum of the molecular weights and extinction coefficients of the chaperone and precursor protein.

### Competition assays

The first competition assay was performed by adding an equimolar mixture of TIM8-13 and TIM9-10 chaperones to the NiNTA-bound precursor protein, Tim23<sup>FL</sup> or Ggc1. After washing the column, precursor-chaperone complex was eluted in buffer A supplemented with 300 mM imidazole. In the time-dependent competition assay, the complex of a precursor protein and one of the chaperones (TIM8-13 or TIM9-10) was prepared, and then an equimolar amount of the other chaperone was added (time point 0). The reaction mixture was incubated at 30°C. After 0.5, 1, and 3 hours, an aliquot of the reaction mixture was taken and the (newly formed) precursor-chaperone complex was isolated on a NiNTA affinity column. The difference in the amount of specific chaperone, TIM8-13 or TIM9-10, bound to the precursor protein was analyzed by SDS-PAGE and liquid chromatography coupled with mass spectrometry (LC-ESI-TOF-MS, 6210, Agilent Technologies, at the MS platform, IBS Grenoble). Samples for analysis by MS were heat-shocked for 15 min at 90°C, resulting in the dissociation and precipitation of the precursor protein, while the apo-chaperones were recovered in the supernatant after cooling the sample and centrifugation for 10 min at 39 kg. As a reference, samples of precursor proteins, Tim23<sup>FL</sup> and Ggc1, bound to individual chaperone, TIM8-13 or TIM9-10, were prepared and analyzed in parallel. To be noted, preparation of the TIM8-13-Ggc1 complex, in quantity sufficient for the analysis, was unsuccessful. To calculate the difference in the amount of specific chaperone bound to the precursor protein, normalized areas under the chromatography peaks corresponding to each Tim monomer were used.

### ITC experiments

Calorimetric binding experiments of Tim23<sup>IMS</sup> and TIM chaperones were performed using a MicroCal ITC200 instrument (GE Healthcare). Sixteen successive 2.5- $\mu$ l aliquots of 1.15 mM Tim23<sup>IMS</sup> were injected into a sample cell containing 55  $\mu$ M TIM9-10 or TIM8-13. All ITC data were acquired in buffer A at 20°C. Control experiments included titrating Tim23<sup>IMS</sup> into buffer A. The enthalpy accompanying each injection was calculated by integrating the resultant exotherm, which corresponds to the released heat as a function of ligand concentration added at each titration point. ITC data were analyzed via the MicroCal Origin software using a single-site binding model and nonlinear least squares fit of thermodynamic binding parameters ( $\Delta H$ ,  $K$ , and  $n$ ). An identical procedure was performed for TIM8-13-cytochrome c and TIM9-10-cytochrome c ITC experiments. Cytochrome c was from horse heart (Merck/Sigma-Aldrich). We also performed ITC experiments with the VDAC peptides; no effects could be detected, in line with a millimolar affinity, as already reported for the TIM9-10-cyclic VDAC<sub>257-279</sub> peptide (23).

### Cross-linking of VDAC<sub>257-279</sub>

In vitro cross-linking of VDAC<sub>257-279</sub> used precisely the protocol described in (23) for TIM9-10. Briefly, 5  $\mu$ M TIM8-13 was mixed with VDAC<sub>257-279</sub> at 0, 25, or 50  $\mu$ M; incubated for 10 min on ice; and ultraviolet (UV)-illuminated (30 min, 4°C).

We detected several cross-linking adducts of the  $\beta$ -hairpin peptide and the TIM components. Such multiband behavior is similar to the pattern of cross-linking products of this peptide with either TOM (translocase of the outer membrane) subunits or cytosolic chaperones (34, 35). We suppose that this variability can result from a variable number of peptides bound to one molecule of protein and from different cross-linking sites on the protein, which, in turn, can cause different migration behavior in the SDS-PAGE.

### SEC-MALS experiments

SEC-MALS experiments were performed at the Biophysical platform (AUC-PAOL) in Grenoble. The experimental setup comprised a high-performance liquid chromatography (HPLC) system (Shimadzu, Kyoto, Japan) consisting of a DGU-20 AD degasser, an LC-20 AD pump, a SIL20-ACHT autosampler, an XL-Therm column oven (WynSep, Sainte Foy d'Aigrefeuille, France), a CBM-20A communication interface, an SPD-M20A UV-visible detector, a miniDAWN TREOS static light scattering detector (Wyatt, Santa Barbara, USA), a DynaPro NanoStar dynamic light scattering detector, and an Optilab rEX refractive index detector. The samples were stored at 4°C, and a volume of 20, 40, 50, or 90  $\mu$ l was injected on a Superdex 200, equilibrated at 4°C; the buffer was 50 mM Tris and 150 mM NaCl filtered at 0.1  $\mu$ m, at a flow rate of 0.5 ml/min. Bovine serum albumin was used for calibration. Two independent sets of experiments conducted with two different batches of protein samples were highly similar.

### Analytical ultracentrifugation

AUC experiments of TIM8-13 and TIM8-13-Tim23 were performed at 50,000 rpm and 10°C, on an XLI analytical ultracentrifuge, with An-60 Ti and An-50 Ti rotors (Beckman Coulter, Palo Alto, USA) and double-sector cells of optical path length 12 and 3 mm equipped with Sapphire windows (Nanolytics, Potsdam, DE). Acquisitions were made using absorbance at 250- and 280-nm wavelength and interference optics. The reference is the buffer 50 mM Tris and

150 mM NaCl. The data were processed by Redate software version 1.0.1. The c(s) and Non Interacting Species (NIS) analysis was done with SEDFIT software version 15.01b and Gussi 1.2.0, and the multiwavelength analysis was done with SEDPHAT software version 12.1b.

### NMR spectroscopy

All NMR experiments were performed on Bruker Avance III spectrometers operating at 600-, 700-, 850-, or 950-MHz  $^1\text{H}$  Larmor frequency. The samples were in the NMR buffer [50 mM NaCl, 50 mM tris (pH 7.4)] with 10% (v/v)  $\text{D}_2\text{O}$ , unless stated differently. All multidimensional NMR data were analyzed with CCPN (version 2) (42). Diffusion ordered spectroscopy (DOSY) data were analyzed with in-house written python scripts. For calculating CSP data, the contribution of each different nuclei was weighted by the gyromagnetic ratios of the respective nucleus: e.g., the combined  $^1\text{H}$ - $^{15}\text{N}$  CSP was calculated as  $\sqrt{[\text{CSP}_{\text{H}}^2 + \text{CSP}_{\text{N}}^2 \cdot (\gamma_{\text{N}}/\gamma_{\text{H}})]}$ , where  $\gamma$  are the gyromagnetic ratios.

### TIM8-13 and Tim23<sub>IMS</sub> resonance assignments

For the resonance assignment of TIM8-13, the following experiments were performed: 2D  $^{15}\text{N}$ - $^1\text{H}$ -BEST-TROSY heteronuclear single-quantum coherence (HSQC), 3D BEST-TROSY HNCO, 3D BEST-TROSY HNcaCO, 3D BEST-TROSY HNCA, 3D BEST-TROSY HNcoCA, 3D BEST-TROSY HNcocaCB, and 3D BEST-TROSY HNcaCB (43, 44) and a 3D  $^{15}\text{N}$ -NOESY HSQC. The experiments were performed with a 0.236 mM [ $^2\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ ]-labeled TIM8-13 at 308 and 333 K. The NMR resonance assignment of TIM9-10 was reported earlier (23). We collected BEST-TROSY HNCA, HNCO, and HNcoCA experiments to assign Tim23<sub>IMS</sub>, aided by the previously reported assignment (36).

### VDAC titration experiments

Cyclic hVDAC<sub>1257-279</sub> peptide was synthesized and lyophilized as described elsewhere (34). The peptide was dissolved in DMSO, and the DMSO concentration was reduced to 10% by stepwise addition of NMR buffer (1:1 in each step). Chaperone, TIM9-10 or TIM8-13, in buffer A was added to yield a final DMSO concentration of 6% and a chaperone concentration of 0.15 mM (TIM9-10) or 0.1 mM (TIM8-13). Combined  $^{15}\text{N}$ - $^1\text{H}$  CSP was calculated from the chemical shifts obtained from the  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra of the complex samples with a molar ratio of 1:4 for TIM9-10:VDAC and a molar ratio of 1:5 for TIM8-13:VDAC, in comparison to the chemical shifts from the apo-chaperone spectrum. The NMR experiments were performed at 308 K.

### Tim23<sub>IMS</sub> titration experiments

For each titration point, individual samples were prepared by mixing two soluble protein samples and monitored using  $^{15}\text{N}$ - $^1\text{H}$ -BEST-TROSY HSQC experiments at 283 K (for Tim23 observed experiment) or at 308 K (for chaperone observed experiments). Titration samples with 100  $\mu\text{M}$  [ $^{15}\text{N}$ ]-labeled Tim23<sub>IMS</sub> with molar ratios for Tim23<sub>IMS</sub>:TIM8-13 from 1:0 to 1:4 and molar ratios for Tim23<sub>IMS</sub>:TIM9-10 from 1:0 to 1:5 were used. For the chaperone observed experiments, used samples contained 200  $\mu\text{M}$  [ $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled TIM8-13 with 1:0 and 1:1 molar ratios of Tim23<sub>IMS</sub> and 350  $\mu\text{M}$  [ $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled TIM9-10 with 1:0 and 1:3 molar ratios of Tim23<sub>IMS</sub>.

### NMR experiments with the Tim23<sub>FL</sub>

Complexes of the chaperones with the FL Tim23 were prepared as indicated above (preparation of chaperone-precursor protein complexes). Peak positions (chemical shifts) of the amide backbone sites

of TIM8-13, apo and in complex with Tim23<sub>FL</sub>, were obtained from the  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments at 308 K, with 120  $\mu\text{M}$  [ $^{13}\text{CH}_3$ -ILV]-TIM8-13-Tim23<sub>FL</sub> sample. Similarly, to calculate combined  $^{15}\text{N}$ - $^1\text{H}$  and  $^{13}\text{C}$ - $^1\text{H}$  CSPs, chemical shifts of the amide backbone and ILVA- $^{13}\text{CH}_3$  groups of TIM9-10, apo and in complex with Tim23<sub>FL</sub>, were obtained from the  $^1\text{H}$ - $^{15}\text{N}$  HSQC and  $^1\text{H}$ - $^{13}\text{C}$  heteronuclear multiple-quantum coherence (HMQC) experiments at 288 K. Sample of [ $^{13}\text{CH}_3$ -ILVA]-TIM9-10 with Tim23<sub>FL</sub> was at 140  $\mu\text{M}$  concentration. For the CSP calculations with the complexes of [ $^2\text{H}$ - $^{15}\text{N}$ ]-labeled Tim23<sub>FL</sub> and the chaperones (190  $\mu\text{M}$  complex with TIM8-13 and 61  $\mu\text{M}$  complex with TIM9-10), chemical shifts from  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments at 288 K were used in comparison to the chemical shifts of apo-Tim23<sub>IMS</sub>.

### Diffusion ordered spectroscopy

Diffusion-ordered NMR spectroscopy (DOSY) experiments were performed at 288 K and 600-MHz  $^1\text{H}$  Larmor frequency. Diffusion constants were derived from a series of 1D  $^1\text{H}$  spectra either over the methyl region (methyl-selective DOSY experiments, for  $^{13}\text{CH}_3$ -ILVA-labeled apo and Tim23<sub>FL</sub>-bound TIM9-10) or over the amide region (for [ $^{15}\text{N}$ ]-Tim23<sub>FL</sub>-TIM9-10). Diffusion coefficients were obtained from fitting integrated 1D intensities as a function of the gradient strength at constant diffusion delay.

### SAXS data collection and analysis

SAXS data were collected at ESRF (European Synchrotron Radiation Facility) BM29 beamline (45) with a Pilatus 1M detector (Dectris) at a distance of 2.872 m from the 1.8-mm-diameter flow-through capillary. Data on TIM8-13 were collected in a batch mode. The x-ray energy was 12.5 keV, and the accessible  $q$  range was 0.032 to 4.9  $\text{nm}^{-1}$ . The incoming flux at the sample position was in the order of 1012 photons/s in 700 mm  $\times$  700 mm. All images were automatically azimuthally averaged with pyFAI (46). SAXS data of pure TIM8-13 were collected at 1, 2.5, and 5 mg/ml using the BioSAXS sample changer (47). Ten frames of 1 s were collected for each concentration. Exposures with radiation damage were discarded, the remaining frames were averaged, and the background was subtracted by an online processing pipeline (48). Data from the three concentrations were merged following standard procedures to create an idealized scattering curve, using PRIMUS from the ATSAS package (49). The pair distribution function  $p(r)$  was calculated using GNOM (50).

Online purification of the TIM8-13-Tim23<sub>FL</sub> and TIM9-10-Tim23<sub>FL</sub> complexes using gel-filtration column (HiLoad 16/600 Superdex 200 PG) was performed with an HPLC system (Shimadzu, France), as described in (51). The HPLC system was directly coupled to the flow-through capillary of SAXS exposure unit. The flow rate for all online experiments was 0.2 ml/min. Data collection was performed continuously throughout the chromatography run at a frame rate of 1 Hz. All SAXS data have been deposited on SASBDB (Small Angle Scattering Biological Data Bank).

### MD simulations and fitting of SAXS data

The initial model of Tim23 was built using I-TASSER (Iterative Threading ASSEMBLY Refinement) (52) and QUARK web servers (53), which predicted a long unstructured N-terminal tail and four/five helical structures in the TM domain. The structure of TIM9-10 hexamer built in our previous work (23) was used as the initial model of TIM9-10 chaperone and as the template to build the model of TIM8-13 chaperone based on the sequence of yeast Tim8 and



Tim13 (UniProt IDs: P57744 and P53299) by homology modeling with MODELLER (54). [Note that in the crystal structure of TIM8-13 (PDB-ID 3CJH), more than 75 residues are missing in each Tim8-Tim13 pair, thus requiring model building.] The disulfide bonds related to the twin CX<sub>3</sub>C motif were also kept in these models. The structures of the TIM8-13 hexamer and Tim23 were subsequently used to build the full structure of the TIM8-13-Tim23 complex by manually wrapping the helical structures of the TM domain of Tim23 around the hydrophobic cleft of TIM8-13, which has been identified by NMR, and leaving the unstructured N terminus of Tim23 as a floppy tail. The complex structure was further optimized by energy minimization and relaxation in 100-ns MD simulations using the simulation protocol as described in the following section. This model was used to generate the so-called N-tail unbound ensemble, in which the N-terminal half of Tim23 is free in solution. On the basis of the N-tail unbound model of the TIM8-13-Tim23 complex, we further constructed the N-tail bound model, in which the N-terminal half of Tim23 is in contact with the upper part of TIM8-13. This was achieved by adding a restraint term in the force field using PLUMED plugin (55),  $V_{\text{restraints}}$ , which is a half-harmonic potential of the form of  $k(R - R_0)^2$  when  $R$  is larger than  $R_0$ , and zero when  $R$  is less than  $R_0$ . Here,  $R$  is the distance between the center of mass of Tim23Nter and the top part of the chaperone. We used  $R_0 = 1$  nm and  $k = 400$  kJ mol<sup>-1</sup>. The N-tail bound and unbound models for the TIM9-10-Tim23 complex were constructed by replacing TIM8-13 with TIM9-10 based on the corresponding TIM8-13-Tim23 models.

The TIM8-13-Tim23 complex in the N-tail unbound conformation was placed into a periodic cubic box with sides of 17.5 nm solvated with TIP3P water molecules containing Na<sup>+</sup> and Cl<sup>-</sup> ions at 0.10 M, resulting in ~700,000 atoms in total. To reduce the computational cost, the complex in N-tail bound conformation was placed in a smaller cubic box with sides of 12.9 nm, resulting in ~300,000 atoms in total. The systems of the TIM9-10-Tim23 complex have similar size as the TIM8-13-Tim23 systems in the corresponding states. The apo TIM8-13 chaperone was placed into a periodic cubic box with sides of 12.0 nm, containing ~230,000 atoms.

The Amber ff99SB-disp force field (56) was used for all simulations. The temperature and pressure were kept constant at 300 K using the v-rescale thermostat and at 1.0 bar using the Parrinello-Rahman barostat with a 2-ps time coupling constant, respectively. Neighbor searching was performed every 10 steps. The particle-meshed-Ewald (PME) algorithm was used for electrostatic interactions. A single cutoff of 1.0 nm was used for both the PME algorithm and van der Waals interactions. A reciprocal grid of 96 × 96 × 96 cells was used with fourth-order B-spline interpolation. The hydrogen mass repartitioning technique (57) was used with a single linear constraint solver (LINCS) iteration (expansion order 6) (58), allowing simulations to be performed with an integration time step of 4 fs. MD simulations were performed using GROMACS 2018 or 2019 (59).

A total of 4.25-μs trajectories was collected to sample the conformational space of the chaperone-Tim23 complexes in both N-tail bound and N-tail unbound states. Four-microsecond trajectories were also collected to sample the ensemble of apo TIM8-13 chaperone. These sampled conformations were used for further ensemble refinement using the BME method (37, 38) guided by experimental SAXS data as described in our previous work (23). The distribution of both states, in principle, could be identified from the force field but needs substantial sampling. Therefore, instead of estimating the prior by large-scale MD simulations, we assigned equal weight (50%)

for both states by inputting equal number of conformations (5000) into the mixed ensemble so without bias to either state. By tuning the regularization parameter in the BME reweighting algorithm, we adjusted the conformational weights in variant degrees to improve the fitting with experimental SAXS data.

The hydrogen bond and salt bridge formation between the N-terminal tail of Tim23 (residues 1 to 100) and the top surface of the chaperones was analyzed by GetContacts scripts (<https://getcontacts.github.io/>) and visualized using Flareplot (<https://gpcrviz.github.io/flareplot/>). Protein structures were visualized with PyMOL and VMD.

## Calculations of affinities and populations

### Estimation of the population of N-tail bound states from ITC-derived $K_d$

We attempted to link the ITC-derived dissociation constant of the Tim23<sub>IMS</sub> fragment to the populations of bound and unbound states in the Tim23<sub>FL</sub>-chaperone complexes, using a rationale akin to the one outlined earlier for binding of disordered proteins to two sub-sites (60). Briefly, we treat the N-terminal tail of Tim23 as a ligand and the remaining bound complex as the target protein, and then the relationship between the population of the bound state ( $P_{\text{bound}}$ ) and the binding affinity ( $K_d$ ) can be written as  $P_{\text{bound}}/(1 - P_{\text{bound}}) = C_{\text{eff}}/K_d$ , where  $C_{\text{eff}}$  is the effective concentration of the disordered N-tail, which was estimated to be between 0.2 and 3 mM from the MD simulations, resulting in the estimation of  $P_{\text{bound}}$  to be between 75 and 98%.

### Estimation of the $K_d$ ratio from competition assays

Determining dissociation constants of TIM chaperones to its insoluble client proteins is hampered by the impossibility to form the complexes by solution methods such as titration, as it requires the pull-down method outlined in Fig. 1A. Nonetheless, the amount of TIM8-13-Tim23 and TIM9-10-Tim23 complexes obtained in the competition assays (Fig. 1) can provide an estimate of the relative affinities. The dissociation constants can be written from the concentrations as follows

$$K_d^{\text{TIM8-13-TIM23}} = \frac{[\text{TIM8} \cdot 13] \times [\text{TIM23}]}{[\text{TIM8} \cdot 13 - \text{TIM23}]}$$

$$K_d^{\text{TIM9-10-TIM23}} = \frac{[\text{TIM9} \cdot 10] \times [\text{TIM23}]}{[\text{TIM9} \cdot 10 - \text{TIM23}]}$$

where [TIM8-13 - Tim23] denotes the concentration of the formed chaperone-precursor complex, and [TIM8-13] and [Tim23] are the concentrations of free chaperone and precursor protein in solution. The latter is negligible, as no free precursor protein is eluted from the column (some aggregated precursor protein was removed from the equilibrium). Both chaperones have been applied at the same concentration  $c_0 = [\text{TIM8} \cdot 13] + [\text{TIM8} \cdot 13 - \text{Tim23}] = [\text{TIM9} \cdot 10] + [\text{TIM9} \cdot 10 - \text{Tim23}]$  to the resin-bound precursor protein that was present at a concentration  $b_0 = [\text{Tim23}] + [\text{TIM9} \cdot 10 - \text{Tim23}] + [\text{TIM8} \cdot 13 - \text{Tim23}]$ .

Using the ratio of formed complex obtained in the competition assay,  $r = [\text{TIM9} \cdot 10 - \text{Tim23}]/[\text{TIM8} \cdot 13 - \text{Tim23}]$  leads to

$$\frac{K_d^{\text{TIM9-10-TIM23}}}{K_d^{\text{TIM8-13-TIM23}}} = \frac{c_0 \times (1 + r) - r \times b_0}{[c_0 \times (1 + r) - b_0] \times r}$$



The experimental protocol does not allow to determine with precision the concentrations of precursor protein ( $b_0$ ) and each chaperone ( $c_0$ ), as the former is bound to a resin. As the chaperone was added in excess, and some of the precursor protein precipitated on the column, we can safely assume  $c_0 \leq b_0$ . With an experimentally found ratio of formed complexes of  $r = 5$  and assuming that  $c_0/b_0$  assumes the values of 1 to 5, the  $K_d$  ratio falls in the range of 1:25 to 1:6, i.e., ca. one order of magnitude.

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/51/eabd0263/DC1>

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

- E. T. Powers, R. I. Morimoto, A. Dillin, J. W. Kelly, W. E. Balch, Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem.* **78**, 959–991 (2009).
- Y. E. Kim, M. S. Hipp, A. Bracher, M. Hayer-Hartl, F. U. Hartl, Molecular chaperone functions in protein folding and proteostasis. *Annu. Rev. Biochem.* **82**, 323–355 (2013).
- D. H. Kim, J.-E. Lee, Z.-Y. Xu, K. R. Geem, Y. Kwon, J. W. Park, I. Hwang, Cytosolic targeting factor AKR2A captures chloroplast outer membrane-localized client proteins at the ribosome during translation. *Nat. Commun.* **6**, 6843 (2015).
- D. Bose, A. Chakrabarti, Substrate specificity in the context of molecular chaperones. *IUBMB Life* **69**, 647–659 (2017).
- S. Rüdiger, L. Germeroth, J. Schneider-Mergener, B. Bukau, Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **16**, 1501–1507 (1997).
- S. Blond-Elguindi, S. E. Cwirla, W. J. Dower, R. J. Lipshutz, S. R. Sprang, J. F. Sambrook, M. J. Gething, Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. *Cell* **75**, 717–728 (1993).
- N. T. Knoblauch, S. Rüdiger, H. J. Schönfeld, A. J. Driessen, J. Schneider-Mergener, B. Bukau, Substrate specificity of the SecB chaperone. *J. Biol. Chem.* **274**, 34219–34225 (1999).
- P. Koldewey, F. Stull, S. Horowitz, R. Martin, J. C. A. Bardwell, Forces driving chaperone action. *Cell* **166**, 369–379 (2016).
- L. He, T. Sharpe, A. Mazur, S. Hiller, A molecular mechanism of chaperone-client recognition. *Sci. Adv.* **2**, e1601625 (2016).
- S. Hiller, B. M. Burmann, Chaperone–client complexes: A dynamic liaison. *J. Magn. Reson.* **289**, 142–155 (2018).
- P. Genevax, F. Keppel, F. Schwager, P. S. Langendijk-Genevax, F. U. Hartl, C. Georgopoulos, In vivo analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* **5**, 195–200 (2004).
- B. Bukau, E. Deuerling, C. Pfund, E. A. Craig, Getting newly synthesized proteins into shape. *Cell* **101**, 119–122 (2000).
- B. M. Burmann, J. A. Gerez, I. Matečko-Burmann, S. Campioni, P. Kumari, D. Ghosh, A. Mazur, E. E. Aspholm, D. Śulskis, M. Wawrzyniuk, T. Bock, A. Schmidt, S. G. D. Rüdiger, R. Riek, S. Hiller, Regulation of  $\alpha$ -synuclein by chaperones in mammalian cells. *Nature* **577**, 127–132 (2020).
- B. M. Burmann, C. Wang, S. Hiller, Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. *Nat. Struct. Mol. Biol.* **20**, 1265–1272 (2013).
- J. Thoma, B. M. Burmann, S. Hiller, D. J. Müller, Impact of holdase chaperones Skp and SurA on the folding of  $\beta$ -barrel outer-membrane proteins. *Nat. Struct. Mol. Biol.* **22**, 795–802 (2015).
- T. Saio, X. Guan, P. Rossi, A. Economou, C. G. Kalodimos, Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science* **344**, 1250494 (2014).
- Y. Jiang, P. Rossi, C. G. Kalodimos, Structural basis for client recognition and activity of Hsp40 chaperones. *Science* **365**, 1313–1319 (2019).
- C. Huang, P. Rossi, T. Saio, C. G. Kalodimos, Structural basis for the antifolding activity of a molecular chaperone. *Nature* **537**, 202–206 (2016).
- N. Wiedemann, N. Pfanner, Mitochondrial machineries for protein import and assembly. *Annu. Rev. Biochem.* **86**, 685–714 (2017).
- C. Sirrenberg, M. Endres, H. Fölsch, R. A. Stuart, W. Neupert, M. Brunner, Carrier protein import into mitochondria mediated by the intermembrane proteins Tim10/Mrs11 and Tim12/Mrs5. *Nature* **391**, 912–915 (1998).
- C. M. Koehler, E. Jarosch, K. Tokatlidis, K. Schmid, R. J. Schweyen, G. Schatz, Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. *Science* **279**, 369–373 (1998).
- E. Lionaki, C. de Marcos Lousa, C. Baud, M. Vougioukalaki, G. Panayotou, K. Tokatlidis, The essential function of Tim12 in vivo is ensured by the assembly interactions of its C-terminal domain. *J. Biol. Chem.* **283**, 15747–15753 (2008).
- K. Weinhäupl, C. Lindau, A. Hessel, Y. Wang, C. Schütze, T. Jores, L. Melchionda, B. Schönfisch, H. Kalbacher, B. Bersch, D. Rapaport, M. Brennich, K. Lindorff-Larsen, N. Wiedemann, P. Schanda, Structural basis of membrane protein chaperoning through the mitochondrial intermembrane space. *Cell* **175**, 1365–1379.e25 (2018).
- C. M. Koehler, D. Leuenberger, S. Merchant, A. Renold, T. Junne, G. Schatz, Human deafness dystonia syndrome is a mitochondrial disease. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2141–2146 (1999).
- S. A. Paschen, U. Rothbauer, K. Káldi, M. F. Bauer, W. Neupert, M. Brunner, The role of the TIM8–13 complex in the import of Tim23 into mitochondria. *EMBO J.* **19**, 6392–6400 (2000).
- K. Roesch, S. P. Curran, L. Tranebjærg, C. M. Koehler, Human deafness dystonia syndrome is caused by a defect in assembly of the DDP1/TIMM8a–TIMM13 complex. *Hum. Mol. Genet.* **11**, 477–486 (2002).
- Y. Kang, A. J. Anderson, T. D. Jackson, C. S. Palmer, D. P. De Souza, K. M. Fujihara, T. Stait, A. E. Frazier, N. J. Clemons, D. Tull, D. R. Thorburn, M. J. Mc Conville, M. T. Ryan, D. A. Stroud, D. Stojanovski, Function of hTim8a in complex IV assembly in neuronal cells provides insight into pathomechanism underlying Mohr-Tranebjærg syndrome. *eLife* **8**, e48828 (2019).
- M. Morgenstern, S. B. Stiller, P. Lübbert, C. D. Peikert, S. Dannenmaier, F. Drepper, U. Weill, P. Höb, R. Feuerstein, M. Gebert, M. Bohnert, M. van der Laan, M. Schuldiner, C. Schütze, S. Oeljeklaus, N. Pfanner, N. Wiedemann, B. Warscheid, Definition of a high-confidence mitochondrial proteome at quantitative scale. *Cell Rep.* **19**, 2836–2852 (2017).
- D. Leuenberger, N. A. Bally, G. Schatz, C. M. Koehler, Different import pathways through the mitochondrial intermembrane space for inner membrane proteins. *EMBO J.* **18**, 4816–4822 (1999).
- A. J. Davis, N. B. Sepuri, J. Holder, A. E. Johnson, R. E. Jensen, Two intermembrane space TIM complexes interact with different domains of Tim23p during its import into mitochondria. *J. Cell Biol.* **150**, 1271–1282 (2000).
- K. Roesch, P. J. Hynds, R. Varga, L. Tranebjærg, C. M. Koehler, Karin Roesch, P. J. Hynds, R. Varga, L. Tranebjærg, C. M. Koehler, The calcium-binding aspartate/glutamate carriers, citrin and aralar1, are new substrates for the DDP1/TIMM8a–TIMM13 complex. *Hum. Mol. Genet.* **13**, 2101–2111 (2004).
- S. C. Hoppins, F. E. Nargang, The Tim8–Tim13 complex of *Neurospora crassa* functions in the assembly of proteins into both mitochondrial membranes. *J. Biol. Chem.* **279**, 12396–12405 (2004).
- S. J. Habib, T. Waizenegger, M. Lech, W. Neupert, D. Rapaport, Assembly of the TOB complex of mitochondria. *J. Biol. Chem.* **280**, 6434–6440 (2005).
- T. Jores, A. Klinger, L. E. Groß, S. Kawano, N. Finner, E. Duchardt-Ferner, J. Wöhnert, H. Kalbacher, T. Endo, E. Schleiff, D. Rapaport, Characterization of the targeting signal in mitochondrial  $\beta$ -barrel proteins. *Nat. Commun.* **7**, 12036 (2016).
- T. Jores, J. Lawatschek, V. Beke, M. Franz-Wachtel, K. Yunoki, J. C. Fitzgerald, B. Macek, T. Endo, H. Kalbacher, J. Buchner, D. Rapaport, Cytosolic Hsp70 and Hsp40 chaperones enable the biogenesis of mitochondrial  $\beta$ -barrel proteins. *J. Cell Biol.* **217**, 3091–3108 (2018).
- L. de la Cruz, R. Bajaj, S. Becker, M. Zweckstetter, The intermembrane space domain of Tim23 is intrinsically disordered with a distinct binding region for presequences. *Protein Sci.* **19**, 2045–2054 (2010).
- S. Orioli, A. H. Larsen, S. Bottaro, K. Lindorff-Larsen, How to learn from inconsistencies: Integrating molecular simulations with experimental data. *Prog. Mol. Biol. Transl. Sci.* **170**, 123–176 (2020).
- S. Bottaro, T. Bengtsen, K. Lindorff-Larsen, Integrating molecular simulation and experimental data: A Bayesian/maximum entropy reweighting approach. *Methods Mol. Biol.* **2112**, 219–240 (2020).
- M. C. Ahmed, E. Papaleo, K. Lindorff-Larsen, How well do force fields capture the strength of salt bridges in proteins? *PeerJ* **6**, e4967 (2018).
- S. Hiller, Chaperone-bound clients: The importance of being dynamic. *Trends Biochem. Sci.* **44**, 517–527 (2019).
- K. N. Truscott, P. Kovermann, A. Geissler, A. Merlin, M. Meijer, A. J. Driessen, J. Rassow, N. Pfanner, R. Wagner, A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. *Nat. Struct. Biol.* **8**, 1074–1082 (2001).
- W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **59**, 687–696 (2005).
- P. Schanda, H. Van Melckebeke, B. Brutscher, Speeding up three-dimensional protein NMR experiments to a few minutes. *J. Am. Chem. Soc.* **128**, 9042–9043 (2006).
- A. Favier, B. Brutscher, Recovering lost magnetization: Polarization enhancement in biomolecular NMR. *J. Biomol. NMR* **49**, 9–15 (2011).

45. P. Pernot, A. Round, R. Barrett, A. De Maria Antolinos, A. Gobbo, E. Gordon, J. Huet, J. Kieffer, M. Lentini, M. Mattenet, C. Morawe, C. Mueller-Dieckmann, S. Ohlsson, W. Schmid, J. Surr, P. Theveneau, L. Zerrad, S. McSweeney, Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664 (2013).
46. G. Ashiotis, A. Deschildre, Z. Nawaz, J. P. Wright, D. Karkoulis, F. E. Picca, J. Kieffer, The fast azimuthal integration Python library: PyFAl. *J. Appl. Crystallogr.* **48**, 510–519 (2015).
47. A. Round, F. Felisaz, L. Fodinger, A. Gobbo, J. Huet, C. Villard, C. E. Blanchet, P. Pernot, S. M. Sweeney, M. Roessle, D. I. Svergun, F. Cipriani, BioSAXS sample changer: A robotic sample changer for rapid and reliable high-throughput X-ray solution scattering experiments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 67–75 (2015).
48. M. E. Brennich, J. Kieffer, G. Bonamis, A. De Maria Antolinos, S. Hutin, P. Pernot, A. Round, Online data analysis at the ESRF bioSAXS beamline, BM29. *J. Appl. Crystallogr.* **49**, 203–212 (2016).
49. M. V. Petoukhov, D. Franke, A. V. Shkumatov, G. Tria, A. G. Kikhney, M. Gajda, C. Gorba, H. D. T. Mertens, P. V. Konarev, D. I. Svergun, New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
50. D. I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
51. M. E. Brennich, A. R. Round, S. Hutin, Online size-exclusion and ion-exchange chromatography on a SAXS beamline. *J. Vis. Exp.* **2017**, e54861 (2017).
52. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The i-tasser suite: Protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
53. D. Xu, Y. Zhang, Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* **80**, 1715–1735 (2012).
54. B. Webb, A. Sali, Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics* **47**, 5–6 (2014).
55. M. Bonomi, G. Bussi, C. Camilloni, G. A. Tribello, P. Banáš, A. Barducci, M. Bernetti, P. G. Bolhuis, S. Bottaro, D. Branduardi, R. Capelli, P. Carloni, M. Ceriotti, A. Cesari, H. Chen, W. Chen, F. Colizzi, S. De, M. De La Pierre, D. Donadio, V. Drobot, B. Ensing, A. L. Ferguson, M. Filizola, J. S. Fraser, H. Fu, P. Gasparotto, F. L. Gervasio, F. Giberti, A. Gil-Ley, T. Giorgino, G. T. Heller, G. M. Hocky, M. Iannuzzi, M. Invernizzi, K. E. Jelfs, A. Jussupow, E. Kirilin, A. Laio, V. Limongelli, K. Lindorff-Larsen, T. Löhner, F. Marinelli, L. Martin-Samos, M. Masetti, R. Meyer, A. Michaelides, C. Molteni, T. Morishita, M. Nava, C. Paissoni, E. Papaleo, M. Parrinello, J. Pfäendner, P. Piaggi, G. M. Piccini, A. Pietropaolo, F. Pietrucci, S. Pipolo, D. Provasi, D. Quigley, P. Raiteri, S. Raniolo, J. Ryzewski, M. Salvalaglio, G. C. Sosso, V. Spiwok, J. Šponer, D. W. H. Swenson, P. Tiwary, O. Valsson, M. Vendruscolo, G. A. Voth, A. White, Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **16**, 670–673 (2019).
56. P. Robustelli, S. Piana, D. E. Shaw, Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4758–E4766 (2018).
57. C. W. Hopkins, S. L. Grand, R. C. Walker, A. E. Roitberg, Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).
58. K. A. Feenstra, B. Hess, H. J. C. Berendsen, Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* **20**, 786–798 (1999).
59. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
60. H.-X. Zhou, X. Pang, C. Lu, Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets. *Phys. Chem. Chem. Phys.* **14**, 10466–10476 (2012).
61. S. Hiller, R. G. Garces, T. J. Malia, V. Y. Orehkov, M. Colombini, G. Wagner, Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science* **321**, 1206–1210 (2008).

**Acknowledgments:** We thank N. Wiedemann and C. Lindau for many insightful discussions and exploratory in vivo experiments. We thank A. Le Roy and C. Ebel (SEC-MALS; AUC), L. Signor (MS), and C. Mas (ITC) for excellent support. We thank H. Kalbacher (Universität Tübingen) for synthesis of the VDAC peptides, K. Giandoreggio-Barranco for preparing Tim23 samples, and U. Guilleim for support with protein production. We are grateful to B. Brutscher, A. Vallet, and A. Favier for excellent NMR platform operation and management. We are grateful to M. Zweckstetter and N. Pfanner for providing the Tim23<sub>IMS</sub> and Tim23<sub>TM</sub> plasmids, respectively. **Funding:** This study was supported by the European Research Council (StG-2012-311318-ProtDyn2Function) and the Agence Nationale de la Recherche (ANR-18-CE92-0032-MitoMemProtImp). This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-05-02 and ANR-10-INBS-0005) and GRAL (ANR-10-LABX-49-01), financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). Y.W. and K.L.-L. were supported by the BRAINSTRUC structural biology initiative from the Lundbeck Foundation. We acknowledge access to computational resources from the Danish National Supercomputer for Life Sciences (Computerome) and the ROBUST Resource for Biomolecular Simulations (supported by the Novo Nordisk Foundation). This work was furthermore supported by the Deutsche Forschungsgemeinschaft (RA 1028/8-1,2 to D.R.) **Author contributions:** K.W. and P.S. initiated this project. I.S., A.H., O.D., D.C., B.B., and K.W. prepared protein samples. I.S. performed the pull-down and competition assays. I.S., K.W., B.B., O.D., and P.S. performed and analyzed NMR data. K.W. and M.B. performed and analyzed SAXS experiments. Y.W. and K.L.-L. performed and analyzed MD simulations. T.J. and D.R. performed/analyzed cross-linking experiments. All authors contributed ideas and discussion. I.S., Y.W., and P.S. made the figures. P.S., I.S., and Y.W. wrote the manuscript with input from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The chemical shift assignments of TIM8-13 have been deposited in the BioMagResBank ([www.bmrb.wisc.edu](http://www.bmrb.wisc.edu)) under accession number 50213. All MD models and SAXS data have been deposited in the SASBDB ([www.sasbdb.org](http://www.sasbdb.org)) under accession numbers SASDH89 (TIM8-13-Tim23), SASDJP4 (TIM9-10-Tim23), SASDQ4 (TIM8-13), and SASDEF2 [TIM9-10 (23)]. All CSP data have been deposited on Mendeley data (<http://dx.doi.org/10.17632/8cr8vtdm.1>). Additional data related to this paper may be requested from the authors.

Submitted 9 June 2020  
Accepted 5 November 2020  
Published 18 December 2020  
10.1126/sciadv.abd0263

**Citation:** I. Sućec, Y. Wang, O. Dakhlaoui, K. Weinhäupl, T. Jores, D. Costa, A. Hessel, M. Brennich, D. Rapaport, K. Lindorff-Larsen, B. Bersch, P. Schanda, Structural basis of client specificity in mitochondrial membrane-protein chaperones. *Sci. Adv.* **6**, eabd0263 (2020).

## Structural basis of client specificity in mitochondrial membrane-protein chaperones

Iva Su#ecYong WangOns DakhlaouiKatharina WeinhäuplTobias JoresDoriane CostaAudrey HesselMartha BrennichDoron RapaportKresten Lindorff-LarsenBeate BerschPaul Schanda

*Sci. Adv.*, 6 (51), eabd0263. • DOI: 10.1126/sciadv.abd0263

### View the article online

<https://www.science.org/doi/10.1126/sciadv.abd0263>

### Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS. Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).



# References

- Abe, Yoshito et al. (2000). "Structural Basis of Presequence Recognition by the Mitochondrial Protein Import Receptor Tom20". In: *Cell* 100.5, pp. 551–560.
- Ahmed, M. C., E. Papaleo, and K. Lindorff-Larsen (2018). "How well do force fields capture the strength of salt bridges in proteins?" In: *PeerJ* 6, e4967.
- Ahting, Uwe et al. (1999). "The TOM core complex: The general protein import pore of the outer membrane of mitochondria". In: *Journal of Cell Biology* 147.5, pp. 959–968. ISSN: 00219525. DOI: [10.1083/jcb.147.5.959](https://doi.org/10.1083/jcb.147.5.959).
- Alam, Athar et al. (2021). "The Role of ClpB in Bacterial Stress Responses and Virulence". In: *Front. Mol. Biosci.* 8, p. 668910. ISSN: 2296889X. DOI: [10.3389/fmolb.2021.668910](https://doi.org/10.3389/fmolb.2021.668910).
- Araiso, Yuhei, Kenichiro Imai, and Toshiya Endo (2020). "Structural snapshot of the mitochondrial protein import gate". In: *FEBS J.*, febs.15661.
- (2022). "Role of the TOM Complex in Protein Import into Mitochondria: Structural Views". In: pp. 1–25.
- Araiso, Yuhei et al. (2019). "Structure of the mitochondrial import gate reveals distinct preprotein paths". In: *Nature* 7.October, pp. 1–20.
- Ashiotis, Giannis et al. (2015). "The fast azimuthal integration Python library: PyFAI". In: *J. Appl. Crystallogr.* 48, pp. 510–519.
- Backes, Sandra et al. (2018). "Tom70 enhances mitochondrial preprotein import efficiency by binding to internal targeting sequences". In: *Journal of Cell Biology* 217.4, pp. 1369–1382. ISSN: 15408140. DOI: [10.1083/jcb.201708044](https://doi.org/10.1083/jcb.201708044).
- Baek, Minkyung et al. (2021). "Accurate prediction of protein structures and interactions using a 3-track network". In: *Science* 8754.July, p. 2021.06.14.448402. URL: <https://doi.org/10.1101/2021.06.14.448402>.
- Bakthisaran, Raman, Ramakrishna Tangirala, and Ch Mohan Rao (2015). "Small heat shock proteins: Role in cellular functions and pathology". In: *BBA- Proteins Proteom.* 1854.4, pp. 291–319. ISSN: 18781454. DOI: [10.1016/j.bbapap.2014.12.019](https://doi.org/10.1016/j.bbapap.2014.12.019). URL: <http://dx.doi.org/10.1016/j.bbapap.2014.12.019>.
- Balchin, David, Manajit Hayer-Hartl, and F. Ulrich Hartl (2020). "Recent advances in understanding catalysis of protein folding by molecular chaperones". In: *FEBS Letters* 594.17, pp. 2770–2781. ISSN: 18733468. DOI: [10.1002/1873-3468.13844](https://doi.org/10.1002/1873-3468.13844).
- Bartlett, Alice I. and Sheena E. Radford (2009). "An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms". In: *Nature Structural and Molecular Biology* 16.6, pp. 582–588. ISSN: 15459993. DOI: [10.1038/nsmb.1592](https://doi.org/10.1038/nsmb.1592).
- Bauer, Matthias F. et al. (1996). "Role of Tim23 as voltage sensor and presequence receptor in protein import into mitochondria". In: *Cell* 87.1, pp. 33–41. ISSN: 00928674. DOI: [10.1016/S0092-8674\(00\)81320-3](https://doi.org/10.1016/S0092-8674(00)81320-3).
- Bauer, Matthias F. et al. (2000). "Protein translocation into mitochondria: the role of TIM complexes". In: *Trends in Cell Biology* 10.1, pp. 25–31. ISSN: 0962-8924. DOI: [10.1016/S0962-8924\(99\)01684-0](https://doi.org/10.1016/S0962-8924(99)01684-0). URL: [https://doi.org/10.1016/S0962-8924\(99\)01684-0](https://doi.org/10.1016/S0962-8924(99)01684-0).
- Bausewein, Thomas et al. (2017). "Cryo-EM Structure of the TOM Core Complex from *Neurospora crassa*". In: *Cell*. ISSN: 10974172. DOI: [10.1016/j.cell.2017.07.012](https://doi.org/10.1016/j.cell.2017.07.012).

- Bender, Tom and Jean Claude Martinou (2016). "The mitochondrial pyruvate carrier in health and disease: To carry or not to carry?" In: *Biochimica et Biophysica Acta - Molecular Cell Research* 1863.10, pp. 2436–2442. ISSN: 18792596. DOI: [10.1016/j.bbamcr.2016.01.017](https://doi.org/10.1016/j.bbamcr.2016.01.017). URL: <http://dx.doi.org/10.1016/j.bbamcr.2016.01.017>.
- Bender, Tom, Gabrielle Pena, and Jean-Claude Martinou (2015). "Regulation of mitochondrial pyruvate uptake by alternative pyruvate carrier complexes". In: *The EMBO Journal* 34.7, pp. 911–924. ISSN: 0261-4189. DOI: [10.15252/embj.201490197](https://doi.org/10.15252/embj.201490197).
- Beverly, Kristen N et al. (2008). "The Tim8 – Tim13 Complex Has Multiple Substrate Binding Sites and Binds Cooperatively to Tim23". In: pp. 1144–1156. DOI: [10.1016/j.jmb.2008.07.069](https://doi.org/10.1016/j.jmb.2008.07.069).
- Bhagawati, Maniraj et al. (2021). "The receptor subunit Tom20 is dynamically associated with the TOM complex in mitochondria of human cells". In: *Mol. Biol. Cell (MBoC)* 32.20.
- Blond-Elguindi, Sylvie et al. (1993). "Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP". In: *Cell* 75.4, pp. 717–728.
- Bolliger, Luca et al. (1995). "Acidic receptor domains on both sides of the outer membrane mediate translocation of precursor proteins into yeast mitochondria". In: *EMBO Journal* 14.24, pp. 6318–6326. ISSN: 02614189. DOI: [10.1002/j.1460-2075.1995.tb00322.x](https://doi.org/10.1002/j.1460-2075.1995.tb00322.x).
- Borgia, Alessandro et al. (2018). "Extreme disorder in an ultrahigh-affinity protein complex". In: *Nature* 555.7694, pp. 61–66. ISSN: 14764687. DOI: [10.1038/nature25762](https://doi.org/10.1038/nature25762).
- Bose, Dipayan and Abhijit Chakrabarti (2017). "Substrate specificity in the context of molecular chaperones". In: *IUBMB Life* 69.9, pp. 647–659.
- Bracher, A and F U Hartl (2013). "Chaperonins". In: *Encyclopedia of Biological Chemistry (Second Edition)*. Ed. by William J Lennarz and M Daniel Lane. Second Edi. Waltham: Academic Press, pp. 456–460. ISBN: 978-0-12-378631-9. DOI: <https://doi.org/10.1016/B978-0-12-378630-2.00080-3>. URL: <https://www.sciencedirect.com/science/article/pii/B9780123786302000803>.
- Brennich, M. E., A. R. Round, and S. Hutin (2017). "Online Size-exclusion and Ion-exchange Chromatography on a SAXS Beamline". In: *J. Vis. Exp.* 119, e54861.
- Brennich, M. E. et al. (2016). "Online data analysis at the ESRF bioSAXS beamline, BM29". In: *J. Appl. Crystallogr.* 49, pp. 203–212. ISSN: 16005767.
- Bricker, Daniel K et al. (2012). "A Mitochondrial Pyruvate Carrier Required for Pyruvate Uptake in Yeast, Drosophila, and Humans". In: *Science* 337.6090, pp. 96–100. DOI: [10.1126/science.1218099.A](https://doi.org/10.1126/science.1218099.A).
- Brix, Jan, Klaus Dietmeier, and Nikolaus Pfanner (1997). "Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70". In: *Journal of Biological Chemistry* 272.33, pp. 20730–20735. ISSN: 00219258. DOI: [10.1074/jbc.272.33.20730](https://doi.org/10.1074/jbc.272.33.20730).
- Brix, Jan et al. (1999). "Distribution of binding sequences for the mitochondrial import receptors Tom20, Tom22, and Tom70 in a presequence-carrying preprotein and a non-cleavable preprotein". In: *Journal of Biological Chemistry* 274.23, pp. 16522–16530. ISSN: 00219258. DOI: [10.1074/jbc.274.23.16522](https://doi.org/10.1074/jbc.274.23.16522).
- Brix, Jan et al. (2000). "The mitochondrial import receptor Tom70: identification of a 25 kda core domain with a specific binding site for preproteins 1 Edited by M. Yaniv". In: *Journal of Molecular Biology* 303.4, pp. 479–488. ISSN: 00222836. DOI: [10.1006/jmbi.2000.4120](https://doi.org/10.1006/jmbi.2000.4120). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0022283600941202>.
- Bukau, Bernd et al. (2000). "Getting Newly Synthesized Proteins into Shape". In: *Cell* 101.2, pp. 119–122.
- Burgess, Richard R. (2009). "Refolding Solubilized Inclusion Body Proteins". In: *Methods in Enzymology*. Ed. by Richard R. Burgess and Murray P. Deutscher. 1st ed. Vol. 463. Elsevier Inc. Chap. Chapter 17, pp. 259–282. DOI: [10.1016/S0076-6879\(09\)63017-2](https://doi.org/10.1016/S0076-6879(09)63017-2). URL: [http://dx.doi.org/10.1016/S0076-6879\(09\)63017-2](http://dx.doi.org/10.1016/S0076-6879(09)63017-2).



- Burmann, Björn M. and Sebastian Hiller (2015). “Chaperones and chaperone-substrate complexes: Dynamic playgrounds for NMR spectroscopists”. In: *Prog. Nucl. Magn. Reson. Spectrosc.* 86-87, pp. 41–64. ISSN: 00796565. DOI: [10.1016/j.pnmrs.2015.02.004](https://doi.org/10.1016/j.pnmrs.2015.02.004).
- Burmann, Björn M., Congwei Wang, and Sebastian Hiller (2013). “Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp”. In: *Nat. Struct. Mol. Biol.* 20.11, pp. 1265–1272. ISSN: 15459993. DOI: [10.1038/nsmb.2677](https://doi.org/10.1038/nsmb.2677).
- Burmann, Björn M. et al. (2020). “Regulation of  $\alpha$ -synuclein by chaperones in mammalian cells”. In: *Nature* 577.7788, pp. 127–132. ISSN: 14764687. DOI: [10.1038/s41586-019-1808-9](https://doi.org/10.1038/s41586-019-1808-9).
- Bykov, Yury S et al. (2020). “Cytosolic Events in the Biogenesis of Mitochondrial Proteins”. In: *Trends in Biochemical Sciences* xx.xx, pp. 1–18. ISSN: 0968-0004. DOI: [10.1016/j.tibs.2020.04.001](https://doi.org/10.1016/j.tibs.2020.04.001). URL: <https://doi.org/10.1016/j.tibs.2020.04.001>.
- Callon, Morgane, Björn M Burmann, and Sebastian Hiller (2014). “Structural mapping of a chaperone-substrate interaction surface”. In: *Angew. Chemie - Int. Ed.* 53.20, pp. 5069–5072.
- Capaldi, Andrew P., Colin Kleanthous, and Sheena E. Radford (2002). “Im7 folding mechanism: Misfolding on a path to the native state”. In: *Nat. Struct. Biol.* 9.3, pp. 209–216. ISSN: 10728368. DOI: [10.1038/nsb757](https://doi.org/10.1038/nsb757).
- Capaldi, Andrew P. et al. (2001). “Ultraprapid mixing experiments reveal that Im7 folds via an on-pathway intermediate”. In: *Nat. Struct. Biol.* 8.1, pp. 68–72. ISSN: 10728368. DOI: [10.1038/83074](https://doi.org/10.1038/83074).
- Chakraborty, Kausik et al. (2010). “Chaperonin-Catalyzed Rescue of Kinetically Trapped States in Protein Folding”. In: *Cell* 142.1, pp. 112–122. ISSN: 00928674. DOI: [10.1016/j.cell.2010.05.027](https://doi.org/10.1016/j.cell.2010.05.027). URL: <http://dx.doi.org/10.1016/j.cell.2010.05.027>.
- Charitou, Vicky, Siri C Van Keulen, and Alexandre M J J Bonvin (2022). “A Cyclisation and Docking Protocol for Cyclic Peptide-Protein Modelling using HADDOCK2.4”. In: *bioRxiv*. DOI: [10.1101/2022.01.21.477251](https://doi.org/10.1101/2022.01.21.477251). URL: [https://www.biorxiv.org/content/10.1101/2022.01.21.477251](https://www.biorxiv.org/content/10.1101/2022.01.21.477251v1). URL: <https://www.biorxiv.org/content/10.1101/2022.01.21.477251v1/abstract>.
- Chen, Dong-Hua et al. (2013). “Visualizing GroEL/ES in the Act of Encapsulating a Folding Protein”. In: *Cell* 153.6, pp. 1354–1365.
- Clerico, Eugenia M. et al. (2021). “Selective promiscuity in the binding of E. coli Hsp70 to an unfolded protein”. In: *PNAS USA* 118.41, e2016962118. ISSN: 0027-8424. DOI: [10.1073/pnas.2016962118](https://doi.org/10.1073/pnas.2016962118).
- Cruz, Laura de la et al. (2010). “The intermembrane space domain of Tim23 is intrinsically disordered with a distinct binding region for presequences”. In: *Protein Sci.* 19.11, pp. 2045–2054.
- C.Y. Fan, Anna and Jason C. Young (2011). “Function of Cytosolic Chaperones in Tom70-Mediated Mitochondrial Import”. In: *Protein & Peptide Letters* 18.2, pp. 122–131. ISSN: 09298665. DOI: [10.2174/092986611794475020](https://doi.org/10.2174/092986611794475020).
- Davis, Alison J et al. (2000). “Two intermembrane space TIM complexes interact with different domains of Tim23p during its import into mitochondria”. In: *J. Cell Biol.* 150.6, pp. 1271–1282.
- Deville, Célia et al. (2017). “Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase”. In: *Sci. Adv.* 3, e1701726. ISSN: 23752548. DOI: [10.1126/sciadv.1701726](https://doi.org/10.1126/sciadv.1701726).
- Deville, Célia et al. (2019). “Two-Step Activation Mechanism of the ClpB Disaggregase for Sequential Substrate Threading by the Main ATPase Motor”. In: *Cell Rep.* 27, pp. 3433–3446. ISSN: 22111247. DOI: [10.1016/j.celrep.2019.05.075](https://doi.org/10.1016/j.celrep.2019.05.075).
- Díaz-Villanueva, José Fernando, Raúl Díaz-Molina, and Victor García-González (2015). “Protein folding and mechanisms of proteostasis”. In: *Int. J. Mol. Sci.* 16.8, pp. 17193–17230. ISSN: 14220067. DOI: [10.3390/ijms160817193](https://doi.org/10.3390/ijms160817193).
- Doron Rapaport and Walter Neupert (1999). “Biogenesis of Tom40, Core Component of the TOM Complex of Mitochondria”. In: *The Journal of Cell Biology* 146, pp. 321–331. ISSN: 00281298. DOI: [10.1007/BF00195069](https://doi.org/10.1007/BF00195069).

- Dukanovic, Jovana and Doron Rapaport (2011). "Multiple pathways in the integration of proteins into the mitochondrial outer membrane." In: *Biochim. Biophys. Acta* 1808.3, pp. 971–980.
- Ellis, R. John and Allen P. Minton (2006). "Protein aggregation in crowded environments". In: *Biol. Chem.* 387.5, pp. 485–497. ISSN: 14316730. DOI: [10.1515/BC.2006.064](https://doi.org/10.1515/BC.2006.064).
- Endo, Toshiya, Hayashi Yamamoto, and Masatoshi Esaki (2003). "Functional cooperation and separation of translocators in protein import into mitochondria, the double-membrane bounded organelles". In: *Journal of Cell Science* 116.16, pp. 3259–3267. DOI: [10.1242/jcs.00667](https://doi.org/10.1242/jcs.00667).
- Esaki, Masatoshi et al. (2003). "Tom40 protein import channel binds to non-native proteins and prevents their aggregation". In: *Nature Structural Biology* 10.12, pp. 988–994. ISSN: 10728368. DOI: [10.1038/nsb1008](https://doi.org/10.1038/nsb1008).
- Esaki, Masatoshi et al. (2004). "Mitochondrial protein import. Requirement of presequence elements and TOM components for precursor binding to the TOM complex". In: *Journal of Biological Chemistry* 279.44, pp. 45701–45707. ISSN: 00219258. DOI: [10.1074/jbc.M404591200](https://doi.org/10.1074/jbc.M404591200).
- Espargaró, Alba et al. (2008). "The in Vivo and in Vitro Aggregation Properties of Globular Proteins Correlate With Their Conformational Stability: The SH3 Case". In: *J. Mol. Biol.* 378.5, pp. 1116–1131. ISSN: 00222836. DOI: [10.1016/j.jmb.2008.03.020](https://doi.org/10.1016/j.jmb.2008.03.020).
- Evans, Richard et al. (2022). "Protein complex prediction with AlphaFold-Multimer". In: *bioRxiv*. DOI: [10.1101/2021.10.04.463034](https://doi.org/10.1101/2021.10.04.463034).
- Fan, Anna C.Y. et al. (2011). "Interaction between the human mitochondrial import receptors Tom20 and Tom70 in vitro suggests a chaperone displacement mechanism". In: *Journal of Biological Chemistry* 286.37, pp. 32208–32219. ISSN: 00219258. DOI: [10.1074/jbc.M111.280446](https://doi.org/10.1074/jbc.M111.280446).
- Favier, Adrien and Bernhard Brutscher (2011). "Recovering lost magnetization: polarization enhancement in biomolecular NMR". In: *J. Biomol. NMR* 49.1, pp. 9–15.
- Fei, Xue et al. (2014). "Formation and structures of GroEL:GroES2 chaperonin footballs, the protein-folding functional form." In: *Proc. Natl. Acad. Sci. U. S. A.* 111.35, pp. 12775–12780.
- Ferbitz, Lars et al. (2004). "Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins". In: *Nature* 431.7008, pp. 590–596. ISSN: 00280836. DOI: [10.1038/nature02899](https://doi.org/10.1038/nature02899).
- Foshag, Daniel et al. (2018). "The E. coli S30 lysate proteome: A prototype for cell-free protein production". In: *New Biotechnology* 40.May 2017, pp. 245–260. ISSN: 18764347. DOI: [10.1016/j.nbt.2017.09.005](https://doi.org/10.1016/j.nbt.2017.09.005).
- Friel, Claire T. et al. (2009). "The mechanism of folding of Im7 reveals competition between functional and kinetic evolutionary constraints". In: *Nat. Struct. Mol. Biol.* 16.3, pp. 318–324. ISSN: 15459993. DOI: [10.1038/nsmb.1562](https://doi.org/10.1038/nsmb.1562).
- Gao, Xiaopan et al. (2021). "Crystal structure of SARS-CoV-2 Orf9b in complex with human TOM70 suggests unusual virus-host interactions". In: *Nat. Commun.* 12.1, p. 2843.
- Gebert, Natalia et al. (2008). "Assembly of the three small Tim proteins precedes docking to the mitochondrial carrier translocase." In: *EMBO Rep.* 9.6, pp. 548–554.
- Genevaux, Pierre et al. (2004). "In vivo analysis of the overlapping functions of DnaK and trigger factor". In: *EMBO Rep.* 5.2, pp. 195–200.
- Gentle, Ian et al. (2004). "The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria". In: *Journal of Cell Biology* 164.1, pp. 19–24. ISSN: 00219252. DOI: [10.1083/jcb.200310092](https://doi.org/10.1083/jcb.200310092).
- Gianni, Stefano et al. (2003). "Unifying features in protein-folding mechanisms". In: *PNAS USA* 100.23, pp. 13286–13291. ISSN: 00278424. DOI: [10.1073/pnas.1835776100](https://doi.org/10.1073/pnas.1835776100).
- Gordon, D E et al. (2020a). "Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms". In: 9403.October.
- Gordon, David E. et al. (2020b). "Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms". In: *Science* 370.6521. ISSN: 10959203. DOI: [10.1126/science.abe9403](https://doi.org/10.1126/science.abe9403).

- Grevel, Alexander, Nikolaus Pfanner, and Thomas Becker (2019). "Coupling of import and assembly pathways in mitochondrial protein biogenesis". In.
- Habib, Shukry J. et al. (2005). "Assembly of the TOB complex of mitochondria". In: *Journal of Biological Chemistry* 280.8, pp. 6434–6440. ISSN: 00219258. DOI: [10.1074/jbc.M411510200](https://doi.org/10.1074/jbc.M411510200).
- Hansen, Katja G. and Johannes M. Herrmann (2019). "Transport of Proteins into Mitochondria". In: *Protein Journal* 0.0, p. 0. ISSN: 15734943. DOI: [10.1007/s10930-019-09819-6](https://doi.org/10.1007/s10930-019-09819-6). URL: <http://dx.doi.org/10.1007/s10930-019-09819-6>.
- Hartl, F. Ulrich, Andreas Bracher, and Manajit Hayer-Hartl (2011). "Molecular chaperones in protein folding and proteostasis". In: *Nature* 475.7356, pp. 324–332. ISSN: 00280836. DOI: [10.1038/nature10317](https://doi.org/10.1038/nature10317).
- He, Chenhui et al. (2020). "Molecular mechanism of SurA's chaperoning function to outer membrane proteins revealed by purification-after-crosslinking single-molecule FRET". In: *Science China Chemistry* 63.8, pp. 1142–1152. ISSN: 18691870. DOI: [10.1007/s11426-020-9758-2](https://doi.org/10.1007/s11426-020-9758-2).
- He, Lichun and Sebastian Hiller (2018). "Common Patterns in Chaperone Interactions with a Native Client Protein". In: *Angew Chem Int. Ed.* 57.20, pp. 5921–5924. ISSN: 15213773. DOI: [10.1002/anie.201713064](https://doi.org/10.1002/anie.201713064).
- (2019). "Frustrated Interfaces Facilitate Dynamic Interactions between Native Client Proteins and Holdase Chaperones". In: *ChemBioChem* 20.22, pp. 2803–2806.
- He, Lichun et al. (2016a). "A molecular mechanism of chaperone-client recognition". In: *Sci. Adv.* 2.11, pp. 1–12. ISSN: 23752548. DOI: [10.1126/sciadv.1601625](https://doi.org/10.1126/sciadv.1601625).
- He, Lichun et al. (2016b). "Structure determination of helical filaments by solid-state NMR spectroscopy". In: *Proc. Natl. Acad. Sci.* 113.3, E272–E281.
- Herzig, Sébastien (2012). "Identification and Functional Expression of the Mitochondrial Pyruvate Carrier". In: *Science* 93. DOI: [10.1126/science.1218530](https://doi.org/10.1126/science.1218530).
- Heyden, Matthias et al. (2012). "Assembly and stability of  $\alpha$ -helical membrane proteins". In: *Soft Matter* 8.30, pp. 7742–7752. ISSN: 17446848. DOI: [10.1039/c2sm25402f](https://doi.org/10.1039/c2sm25402f).
- Hiller, Matthias et al. (2008). "[2,3-<sup>13</sup>C]-labeling of aromatic residues - Getting a head start in the magic-angle-spinning NMR assignment of membrane proteins". In: *Journal of the American Chemical Society* 130.2, pp. 408–409. ISSN: 00027863. DOI: [10.1021/ja077589n](https://doi.org/10.1021/ja077589n).
- Hiller, Sebastian (2019). "Chaperone-Bound Clients: The Importance of Being Dynamic". In: *Trends Biochem. Sci.* 44.6, pp. 517–527. ISSN: 13624326. DOI: [10.1016/j.tibs.2018.12.005](https://doi.org/10.1016/j.tibs.2018.12.005). URL: <https://doi.org/10.1016/j.tibs.2018.12.005>.
- Hiller, Sebastian and Björn M. Burmann (2018). "Chaperone-client complexes: A dynamic liaison". In: *J. Magn. Reson.* 289, pp. 142–155. ISSN: 10960856. DOI: [10.1016/j.jmr.2017.12.008](https://doi.org/10.1016/j.jmr.2017.12.008).
- Hines, V. and G. Schatz (1993). "Precursor binding to yeast mitochondria. A general role for the outer membrane protein Mas70p". In: *Journal of Biological Chemistry* 268.1, pp. 449–454. ISSN: 00219258. DOI: [10.1016/s0021-9258\(18\)54172-7](https://doi.org/10.1016/s0021-9258(18)54172-7). URL: [http://dx.doi.org/10.1016/S0021-9258\(18\)54172-7](http://dx.doi.org/10.1016/S0021-9258(18)54172-7).
- Hönlinger, A et al. (1995). "The mitochondrial receptor complex: Mom22 is essential for cell viability and directly interacts with preproteins." In: *Molecular and Cellular Biology* 15.6, pp. 3382–3389. ISSN: 0270-7306. DOI: [10.1128/mcb.15.6.3382](https://doi.org/10.1128/mcb.15.6.3382).
- Hoppins, Suzanne C. and Frank E. Nargang (2004). "The Tim8-Tim13 Complex of *Neurospora crassa* Functions in the Assembly of Proteins into Both Mitochondrial Membranes". In: *Journal of Biological Chemistry* 279.13, pp. 12396–12405. ISSN: 00219258. DOI: [10.1074/jbc.M313037200](https://doi.org/10.1074/jbc.M313037200).
- Horowitz, Scott et al. (2016). "Visualizing chaperone-assisted protein folding". In: *Nat. Struct. Mol. Biol.* 23.7, pp. 691–697. ISSN: 15459985. DOI: [10.1038/nsmb.3237](https://doi.org/10.1038/nsmb.3237).
- Huang, Chengdong et al. (2016). "Structural basis for the antifolding activity of a molecular chaperone". In: *Nature* 537.7619, pp. 202–206. ISSN: 14764687. DOI: [10.1038/nature18965](https://doi.org/10.1038/nature18965). URL: <http://dx.doi.org/10.1038/nature18965>.

- Imbert, Lionel et al. (2021). "In Vitro Production of Perdeuterated Proteins in H<sub>2</sub>O for Biomolecular NMR Studies". In: *Meth. Mol. Biol.* Vol. 2199, pp. 127–149. ISBN: 9781071608920.
- Irwin, Rose et al. (2021). "Hsp40s play complementary roles in the prevention of tau amyloid formation". In: *eLife* 10, pp. 1–24. ISSN: 2050084X. DOI: [10.7554/ELIFE.69601](https://doi.org/10.7554/ELIFE.69601).
- Iwahashi, Jun et al. (1997). "Analysis of the functional domain of the rat liver mitochondrial import receptor Tom20". In: *Journal of Biological Chemistry* 272.29, pp. 18467–18472. ISSN: 00219258. DOI: [10.1074/jbc.272.29.18467](https://doi.org/10.1074/jbc.272.29.18467). URL: <http://dx.doi.org/10.1074/jbc.272.29.18467>.
- Jiang, Yajun, Paolo Rossi, and Charalampos G. Kalodimos (2019). "Structural basis for client recognition and activity of Hsp40 chaperones". In: *Science* 365.6459, pp. 1313–1319. ISSN: 10959203. DOI: [10.1126/science.aax1280](https://doi.org/10.1126/science.aax1280).
- Jin, Xing and Seok Hoon Hong (2018). "Cell-free protein synthesis for producing 'difficult-to-express' proteins". In: *Biochem. Eng. J* 138, pp. 156–164. ISSN: 1873295X. DOI: [10.1016/j.bej.2018.07.013](https://doi.org/10.1016/j.bej.2018.07.013). URL: <https://doi.org/10.1016/j.bej.2018.07.013>.
- Jores, Tobias et al. (2016). "Characterization of the targeting signal in mitochondrial  $\beta$ -barrel proteins". In: *Nature Communications*. ISSN: 20411723. DOI: [10.1038/ncomms12036](https://doi.org/10.1038/ncomms12036).
- Jores, Tobias et al. (2018). "Cytosolic Hsp70 and Hsp40 chaperones enable the biogenesis of mitochondrial  $\beta$ -barrel proteins". In: *J. Cell Biol.* 217.9, jcb.201712029.
- Jumper, John et al. (2021). "Highly accurate protein structure prediction with AlphaFold". In: *Nature* 596.7873, pp. 583–589. ISSN: 14764687. DOI: [10.1038/s41586-021-03819-2](https://doi.org/10.1038/s41586-021-03819-2). URL: <http://dx.doi.org/10.1038/s41586-021-03819-2>.
- Kang, Yilin et al. (2019). "Function of hTim8a in complex IV assembly in neuronal cells provides insight into pathomechanism underlying Mohr-Tranebjærg syndrome". In: *Elife* 8, pp. 1–32.
- Karagöz, G. Elif et al. (2014). "Hsp90-tau complex reveals molecular basis for specificity in chaperone action". In: *Cell* 156.5, pp. 963–974. ISSN: 10974172. DOI: [10.1016/j.cell.2014.01.037](https://doi.org/10.1016/j.cell.2014.01.037).
- Kellner, Ruth et al. (2014). "Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein". In: *PNAS USA* 111.37, pp. 13355–13360.
- Kiebler, Michael et al. (1993). "The mitochondrial receptor complex: A central role of MOM22 in mediating preprotein transfer from receptors to the general insertion pore". In: *Cell* 74.3, pp. 483–492. ISSN: 00928674. DOI: [10.1016/0092-8674\(93\)80050-0](https://doi.org/10.1016/0092-8674(93)80050-0).
- Kim, Dae Heon et al. (2015). "Cytosolic targeting factor AKR2A captures chloroplast outer membrane-localized client proteins at the ribosome during translation". In: *Nat. Commun.* 6, pp. 1–13.
- Kim, Yujin E. et al. (2013). "Molecular Chaperone Functions in Protein Folding and Proteostasis". In: *Annu. Rev. Biochem.* 82.1, pp. 323–355. ISSN: 0066-4154. DOI: [10.1146/annurev-biochem-060208-092442](https://doi.org/10.1146/annurev-biochem-060208-092442).
- Knoblauch, Nicola T.M. et al. (1999). "Substrate specificity of the SecB chaperone". In: *Journal of Biological Chemistry* 274.48, pp. 34219–34225. ISSN: 00219258. DOI: [10.1074/jbc.274.48.34219](https://doi.org/10.1074/jbc.274.48.34219).
- Koehler, C M et al. (1998). "Import of Mitochondrial Carriers Mediated by Essential Proteins of the Inter-membrane Space". In: *Science* 279.5349, pp. 369–373.
- Koehler, Carla M. et al. (1999). "Human deafness dystonia syndrome is a mitochondrial disease". In: *Proceedings of the National Academy of Sciences of the United States of America* 96.5, pp. 2141–2146. ISSN: 00278424. DOI: [10.1073/pnas.96.5.2141](https://doi.org/10.1073/pnas.96.5.2141).
- Koh, Julie Y., Petr Hájek, and David M. Bedwell (2001). "Overproduction of PDR3 Suppresses Mitochondrial Import Defects Associated with a TOM70 Null Mutation by Increasing the Expression of TOM72 in *Saccharomyces cerevisiae*". In: *Molecular and Cellular Biology* 21.22, pp. 7576–7586. ISSN: 0270-7306. DOI: [10.1128/mcb.21.22.7576-7586.2001](https://doi.org/10.1128/mcb.21.22.7576-7586.2001).
- Koldewey, Philipp et al. (2016). "Forces Driving Chaperone Action". In: *Cell* 166.2, pp. 369–379. ISSN: 10974172. DOI: [10.1016/j.cell.2016.05.054](https://doi.org/10.1016/j.cell.2016.05.054). URL: <http://dx.doi.org/10.1016/j.cell.2016.05.054>.



- Kondo-Okamoto, Noriko, Janet M. Shaw, and Koji Okamoto (2008). "Tetratricopeptide repeat proteins Tom70 and Tom71 mediate yeast mitochondrial morphogenesis". In: *EMBO Reports* 9.1, pp. 63–69. ISSN: 1469221X. DOI: [10.1038/sj.embor.7401113](https://doi.org/10.1038/sj.embor.7401113).
- Kozjak, Vera et al. (2003). "An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane". In: *Journal of Biological Chemistry* 278.49, pp. 48520–48523. ISSN: 00219258. DOI: [10.1074/jbc.C300442200](https://doi.org/10.1074/jbc.C300442200).
- Kragelund, Birthe B and Karen Skriver (2020). *Methods in Molecular Biology: Intrinsically Disordered Proteins*. ISBN: 9781071605233. URL: <http://www.springer.com/series/7651>.
- Kreimendahl, Sebastian et al. (2020). "The selectivity filter of the mitochondrial protein import machinery". In: *BMC Biol.* 18.1, p. 156.
- Krojer, Tobias et al. (2008). "Interplay of PDZ and protease domain of DegP ensures efficient elimination of misfolded proteins". In: *Proceedings of the National Academy of Sciences of the United States of America* 105.22, pp. 7702–7707. ISSN: 00278424. DOI: [10.1073/pnas.0803392105](https://doi.org/10.1073/pnas.0803392105).
- Kyte, Jack and Russell F Doolittle (1982). "A simple method for displaying the hydropathic character of a protein". In: *J. Mol. Biol.* 157.1, pp. 105–132.
- Lemarie, A. and S. Grimm (2011). "Mitochondrial respiratory chain complexes: Apoptosis sensors mutated in cancer". In: *Oncogene* 30.38, pp. 3985–4003. ISSN: 09509232. DOI: [10.1038/onc.2011.167](https://doi.org/10.1038/onc.2011.167).
- Leuenberger, Danielle et al. (1999). "Different import pathways through the mitochondrial intermembrane space for inner membrane proteins". In: *EMBO J.* 18.17, pp. 4816–4822.
- Li, Jingzhi, Wenjun Cui, and Bingdong Sha (2010). "The structural plasticity of Tom71 for mitochondrial precursor translocations". In: *Acta Crystallographica Section F* 66.9, pp. 985–989. DOI: [10.1107/S1744309110025522](https://doi.org/10.1107/S1744309110025522). URL: <https://doi.org/10.1107/S1744309110025522>.
- Li, Jingzhi et al. (2009). "Molecular chaperone Hsp70/Hsp90 prepares the mitochondrial outer membrane translocon receptor Tom71 for preprotein loading". In: *Journal of Biological Chemistry* 284.35, pp. 23852–23859. ISSN: 00219258. DOI: [10.1074/jbc.M109.023986](https://doi.org/10.1074/jbc.M109.023986).
- Lill, Roland and Ulrich Mühlenhoff (2008). "Maturation of iron-sulfur proteins in eukaryotes: Mechanisms, connected processes, and diseases". In: *Annual Review of Biochemistry* 77, pp. 669–700. ISSN: 00664154. DOI: [10.1146/annurev.biochem.76.052705.162653](https://doi.org/10.1146/annurev.biochem.76.052705.162653).
- Lill, Roland and Walter Neupert (1996). "Mechanisms of protein import across the mitochondrial outer membrane". In: *Trends in Cell Biology* 6.2, pp. 56–61. DOI: [10.1016/0962-8924\(96\)81015-4](https://doi.org/10.1016/0962-8924(96)81015-4).
- Lionaki, Eirini et al. (2008). "The essential function of Tim12 in vivo is ensured by the assembly interactions of its C-terminal domain." In: *J. Biol. Chem.* 283.23, pp. 15747–15753.
- Lithgow, Trevor and André Schneider (2010). "Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes". In: *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 365.1541, pp. 799–817.
- Lithgow, Trevor et al. (1994). "Yeast mitochondria lacking the two import receptors Mas20p and Mas70p can efficiently and specifically import precursor proteins". In: *Journal of Biological Chemistry* 269.21, pp. 15325–15330. ISSN: 00219258.
- Maciejewski, Mark W. et al. (2017). "NMRbox: A Resource for Biomolecular NMR Computation". In: *Biophysical Journal* 112.8, pp. 1529–1534. DOI: [10.1016/j.bpj.2017.03.011](https://doi.org/10.1016/j.bpj.2017.03.011).
- Macošek, Jakub, Guillaume Mas, and Sebastian Hiller (2021). "Redefining Molecular Chaperones as Chaotropes". In: *Front. Mol. Biosci.* 8.June, pp. 1–11.
- Maduke, Merritt and David Roise (1996). "Structure and function of mitochondrial presequences". In: *Membrane Protein Transport*. Vol. 3. C, pp. 49–79. ISBN: 1559389893.
- Malet, H el ane et al. (2012). "Newly folded substrates inside the molecular cage of the HtrA chaperone DegQ". In: *Nature Structural and Molecular Biology* 19.2, pp. 152–157. ISSN: 15459993. DOI: [10.1038/nsmb.2210](https://doi.org/10.1038/nsmb.2210).

- Martinez-Hackert, Erik and Wayne A Hendrickson (2009). "Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone". In: *Cell* 138.5, pp. 923–934.
- Mas, Guillaume et al. (2018). "Structural investigation of a chaperonin in action reveals how nucleotide binding regulates the functional cycle". In: *Sci. Adv.* 4.9, pp. 1–10. ISSN: 23752548. DOI: [10.1126/sciadv.aau4196](https://doi.org/10.1126/sciadv.aau4196).
- Mashaghi, Alireza et al. (2016). "Alternative modes of client binding enable functional plasticity of Hsp70". In: *Nature* 539.7629, pp. 448–451. ISSN: 14764687. DOI: [10.1038/nature20137](https://doi.org/10.1038/nature20137). URL: <http://dx.doi.org/10.1038/nature20137>.
- Mattoo, Rayees U.H. and Pierre Goloubinoff (2014). "Molecular chaperones are nanomachines that catalytically unfold misfolded and alternatively folded proteins". In: *Cell. Mol. Life Sci.* 71.17, pp. 3311–3325. ISSN: 14209071. DOI: [10.1007/s00018-014-1627-y](https://doi.org/10.1007/s00018-014-1627-y).
- Mauldin, Randall V and Robert T Sauer (2013). "Allosteric regulation of DegS protease subunits through a shared energy landscape". In: *Nat Chem Biol* 9.2, pp. 90–96.
- Meisinger, Chris et al. (2001). "Protein Import Channel of the Outer Mitochondrial Membrane: a Highly Stable Tom40-Tom22 Core Structure Differentially Interacts with Preproteins, Small Tom Proteins, and Import Receptors". In: *Molecular and Cellular Biology* 21.7, pp. 2337–2348. ISSN: 0270-7306. DOI: [10.1128/mcb.21.7.2337-2348.2001](https://doi.org/10.1128/mcb.21.7.2337-2348.2001).
- Melin, Jonathan et al. (2015). "A presequence-binding groove in Tom70 supports import of Mdl1 into mitochondria". In: *Biochimica et Biophysica Acta - Molecular Cell Research* 1853.8, pp. 1850–1859. ISSN: 18792596. DOI: [10.1016/j.bbamcr.2015.04.021](https://doi.org/10.1016/j.bbamcr.2015.04.021). URL: <http://dx.doi.org/10.1016/j.bbamcr.2015.04.021>.
- Mills, Ryan D. et al. (2009). "Domain Organization of the Monomeric Form of the Tom70 Mitochondrial Import Receptor". In: *Journal of Molecular Biology* 388.5, pp. 1043–1058. ISSN: 00222836. DOI: [10.1016/j.jmb.2009.03.070](https://doi.org/10.1016/j.jmb.2009.03.070). URL: <http://dx.doi.org/10.1016/j.jmb.2009.03.070>.
- Model, Kirstin, Chris Meisinger, and Werner Kühlbrandt (2008). "Cryo-Electron Microscopy Structure of a Yeast Mitochondrial Preprotein Translocase". In: *Journal of Molecular Biology* 383.5, pp. 1049–1057. ISSN: 00222836. DOI: [10.1016/j.jmb.2008.07.087](https://doi.org/10.1016/j.jmb.2008.07.087). URL: <http://dx.doi.org/10.1016/j.jmb.2008.07.087>.
- Model, Kirstin et al. (2001). "Multistep assembly of the protein import channel of the mitochondrial outer membrane". In: *Nat. Struct. Mol. Biol.* 8.4, pp. 361–370.
- Morgenstern, Marcel et al. (2017). "Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale". In: *Cell Reports* 19.13, pp. 2836–2852. ISSN: 2211247. DOI: [10.1016/j.celrep.2017.06.014](https://doi.org/10.1016/j.celrep.2017.06.014).
- Mossmann, Dirk, Chris Meisinger, and F. Nora Vögtle (2012). "Processing of mitochondrial presequences". In: *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1819.9-10, pp. 1098–1106. DOI: [10.1016/j.bbagr.2011.11.007](https://doi.org/10.1016/j.bbagr.2011.11.007).
- Muto, Takanori et al. (2001). "NMR identification of the Tom20 binding segment in mitochondrial presequences". In: *Journal of Molecular Biology* 306.2, pp. 137–143. ISSN: 00222836. DOI: [10.1006/jmbi.2000.4397](https://doi.org/10.1006/jmbi.2000.4397).
- Nargang, Frank E. et al. (1998). "Role of the Negative Charges in the Cytosolic Domain of TOM22 in the Import of Precursor Proteins into Mitochondria". In: *Molecular and Cellular Biology* 18.6, pp. 3173–3181. ISSN: 0270-7306. DOI: [10.1128/mcb.18.6.3173](https://doi.org/10.1128/mcb.18.6.3173).
- Neupert, Walter (2015). "A perspective on transport of proteins into mitochondria: a myriad of open questions." In: *J. Mol. Biol.* 427.6 Pt A, pp. 1135–1158.
- Newmeyer, Donald D. and Shelagh Ferguson-Miller (2003). "Mitochondria: Releasing power for life and unleashing the machineries of death". In: *Cell* 112.4, pp. 481–490. ISSN: 00928674. DOI: [10.1016/S0092-8674\(03\)00116-8](https://doi.org/10.1016/S0092-8674(03)00116-8).



- Nielsen, Jakob Toudahl and Frans AA Mulder (2018). "Potenci: Prediction of temperature, neighbor and ph-corrected chemical shifts for intrinsically disordered proteins". In: *Journal of Biomolecular NMR* 70.3, pp. 141–165. DOI: [10.1007/s10858-018-0166-5](https://doi.org/10.1007/s10858-018-0166-5).
- Nielsen, Jakob Toudahl and Frans A.A. Mulder (2021). "CheSPI: chemical shift secondary structure population inference". In: *Journal of Biomolecular NMR* 75.6-7, pp. 273–291. DOI: [10.1007/s10858-021-00374-w](https://doi.org/10.1007/s10858-021-00374-w). URL: <https://doi.org/10.1007/s10858-021-00374-w>.
- Ollerenshaw, Jason E, Vitali Tugarinov, and Lewis E Kay (2003). "Methyl TROSY: explanation and experimental verification". In: *Magn. Reson. Chem.* 41.10, pp. 843–852.
- Orioli, Simone et al. (2020). "How to learn from inconsistencies: Integrating molecular simulations with experimental data". In: *Prog. Mol. Biol. Transl. Sci.* 1, pp. 123–176. arXiv: [1909.06780](https://arxiv.org/abs/1909.06780).
- Paschen, S. A. (2000). "The role of the TIM8-13 complex in the import of Tim23 into mitochondria". In: *The EMBO Journal* 19.23, pp. 6392–6400. ISSN: 14602075. DOI: [10.1093/emboj/19.23.6392](https://doi.org/10.1093/emboj/19.23.6392). URL: <http://emboj.embopress.org/cgi/doi/10.1093/emboj/19.23.6392>.
- Paschen, Stefan A et al. (2000). "The role of the TIM8-13 complex in the import of Tim23 into mitochondria". In: *EMBO J.* 19.23, pp. 6392–6400.
- Paschen, Stefan A et al. (2003). "Evolutionary conservation of biogenesis of". In: 426.December, pp. 862–866. DOI: [10.1038/nature02222](https://doi.org/10.1038/nature02222). Published.
- Pashley, Clare L. et al. (2012). "Conformational properties of the unfolded state of Im7 in non-denaturing conditions". In: *J. Mol. Biol.* 416.2, pp. 300–318. ISSN: 10898638. DOI: [10.1016/j.jmb.2011.12.041](https://doi.org/10.1016/j.jmb.2011.12.041). URL: <http://dx.doi.org/10.1016/j.jmb.2011.12.041>.
- Pernot, P. et al. (2013). "Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution". In: *J. Synchrotron Radiat.* 20.4, pp. 660–664.
- Petoukhov, M. V. et al. (2012). "New developments in the ATSAS program package for small-angle scattering data analysis". In: *J. Appl. Crystallogr.* 45.2, pp. 342–350.
- Pfanner, Nikolaus (2000). "Protein sorting: Recognizing mitochondrial presequences". In: *Current Biology* 10.11, pp. 412–415. ISSN: 09609822. DOI: [10.1016/S0960-9822\(00\)00507-8](https://doi.org/10.1016/S0960-9822(00)00507-8).
- Pfanner, Nikolaus and Andreas Geissler (2001). "Versatility of the mitochondrial protein import machinery". In: *Nature Reviews* 2.May.
- Pfanner, Nikolaus, Bettina Warscheid, and Nils Wiedemann (2019). "Mitochondrial Protein Organization: From Biogenesis to Networks and Function". In: *Nature Reviews Molecular Cell Biology* 20.5, pp. 267–284. DOI: [10.1038/s41580-018-0092-0](https://doi.org/10.1038/s41580-018-0092-0). Mitochondrial.
- Powers, Evan T. et al. (2009). "Biological and Chemical Approaches to Diseases of Proteostasis Deficiency". In: *Annual Review of Biochemistry* 78.1, pp. 959–991. ISSN: 0066-4154. DOI: [10.1146/annurev.biochem.052308.114844](https://doi.org/10.1146/annurev.biochem.052308.114844).
- Qu, Jian et al. (2007). "The Trimeric Periplasmic Chaperone Skp of Escherichia coli Forms 1:1 Complexes with Outer Membrane Proteins via Hydrophobic and Electrostatic Interactions". In: *J. Mol. Biol.* 374.1, pp. 91–105. ISSN: 00222836. DOI: [10.1016/j.jmb.2007.09.020](https://doi.org/10.1016/j.jmb.2007.09.020).
- Quan, Shu et al. (2011). "Genetic selection designed to stabilize proteins uncovers a chaperone called Spy". In: *Nat. Struct. Mol. Biol.* 18.3, pp. 262–269. ISSN: 15459993. DOI: [10.1038/nsmb.2016](https://doi.org/10.1038/nsmb.2016).
- Rampelt, Heike et al. (2020). "The mitochondrial carrier pathway transports non-canonical substrates with an odd number of transmembrane segments". In: *BMC Biology* 18.1, pp. 1–15. ISSN: 17417007. DOI: [10.1186/s12915-019-0733-6](https://doi.org/10.1186/s12915-019-0733-6).
- Rapaport, Doron (2003). "Finding the right organelle. Targeting signals in mitochondrial outer-membrane proteins". In: *EMBO Reports* 4.10, pp. 948–952. ISSN: 1469221X. DOI: [10.1038/sj.embor.embor937](https://doi.org/10.1038/sj.embor.embor937).

- Rapaport, Doron, Walter Neupert, and Roland Lill (1997). "Mitochondrial protein import: Tom40 PLAYS A MAJOR ROLE IN TARGETING AND TRANSLOCATION OF PREPROTEINS BY FORMING A SPECIFIC BINDING SITE FOR THE PRESEQUENCE". In: *The Journal of biological chemistry* 272, pp. 18725–18731. ISSN: 03044157. DOI: [10.1016/0304-4157\(89\)90002-6](https://doi.org/10.1016/0304-4157(89)90002-6).
- Rehling, Peter et al. (2001). "The mitochondrial import machinery for preproteins". In: *Critical Reviews in Biochemistry and Molecular Biology* 36.3, pp. 291–336. DOI: [10.1080/20014091074200](https://doi.org/10.1080/20014091074200).
- Roesch, Karin et al. (2002). "Human deafness dystonia syndrome is caused by a defect in assembly of the DDP1/TIMM8a-TIMM13 complex." In: *Hum. Mol. Genet.* 11.5, pp. 477–486.
- Roesch, Karin et al. (2004). "The calcium-binding aspartate/glutamate carriers, citrin and aralar1, are new substrates for the DDP1/TIMM8a-TIMM13 complex". In: *Human Molecular Genetics* 13.18, pp. 2101–2111. ISSN: 09646906. DOI: [10.1093/hmg/ddh217](https://doi.org/10.1093/hmg/ddh217).
- Rosenzweig, Rina et al. (2017). "Promiscuous binding by Hsp70 results in conformational heterogeneity and fuzzy chaperone-substrate ensembles". In: *eLife* 6, pp. 1–22. ISSN: 2050084X. DOI: [10.7554/eLife.28030](https://doi.org/10.7554/eLife.28030).
- Round, A. et al. (2015). "BioSAXS Sample Changer: A robotic sample changer for rapid and reliable high-throughput X-ray solution scattering experiments". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 71, pp. 67–75.
- Rüdiger, Stefan et al. (1997). "Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries respect to the bound folding conformer are only partly". In: *The EMBO Journal* 16.7, pp. 1501–1507. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1169754/pdf/001501.pdf>.
- Sadqi, Mourad, Lisa J. Lapidus, and Victor Muñoz (2003). "How fast is protein hydrophobic collapse?" In: *Proceedings of the National Academy of Sciences of the United States of America* 100.21, pp. 12117–12122. ISSN: 00278424. DOI: [10.1073/pnas.2033863100](https://doi.org/10.1073/pnas.2033863100).
- Saio, Tomohide et al. (2014). "Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone". In: *Science* 344.6184, p. 1250494. arXiv: [NIHMS150003](https://arxiv.org/abs/NIHMS150003).
- Saitoh, Takashi et al. (2007). "Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states". In: *EMBO Journal*. ISSN: 02614189. DOI: [10.1038/sj.emboj.7601888](https://doi.org/10.1038/sj.emboj.7601888).
- Scheufler, Clemens et al. (2000). "Structure of TPR domain-peptide complexes: Critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine". In: *Cell* 101.2, pp. 199–210. ISSN: 00928674. DOI: [10.1016/S0092-8674\(00\)80830-2](https://doi.org/10.1016/S0092-8674(00)80830-2).
- Schiffirin, Bob, David J. Brockwell, and Sheena E. Radford (2017). "Outer membrane protein folding from an energy landscape perspective". In: *BMC Biology* 15.1, pp. 1–16. ISSN: 17417007. DOI: [10.1186/s12915-017-0464-5](https://doi.org/10.1186/s12915-017-0464-5).
- Schneider, Birgit et al. (2010). "Membrane Protein Expression in Cell-Free Systems". In: *Methods Mol. Biol.* 601, pp. 165–186. DOI: [10.1007/978-1-60761-344-2](https://doi.org/10.1007/978-1-60761-344-2). URL: <http://link.springer.com/10.1007/978-1-60761-344-2>.
- Schneider, Gisbert et al. (1998). "Feature-extraction from endopeptidase cleavage sites in mitochondrial targeting peptides". In: *Proteins: Structure, Function and Genetics* 30.1, pp. 49–60. ISSN: 08873585. DOI: [10.1002/\(SICI\)1097-0134\(19980101\)30:1<49::AID-PROT5>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-0134(19980101)30:1<49::AID-PROT5>3.0.CO;2-F).
- Sekhar, Ashok et al. (2015). "Mapping the conformation of a client protein through the Hsp70 functional cycle." In: *PNAS USA* 112.33, pp. 10395–10400.
- Sekhar, Ashok et al. (2018). "Conserved conformational selection mechanism of Hsp70 chaperone-substrate interactions". In: *Elife* 7, e32764. arXiv: [NIHMS150003](https://arxiv.org/abs/NIHMS150003).
- Shen, Yang and Ad Bax (2013). "Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks". In: *J. Biomol. NMR* 56.3, pp. 227–241.

- Shiota, T. et al. (2011). "In vivo protein-interaction mapping of a mitochondrial translocator protein Tom22 at work". In: *Proceedings of the National Academy of Sciences* 108.37, pp. 15179–15183. ISSN: 0027-8424. DOI: [10.1073/pnas.1105921108](https://doi.org/10.1073/pnas.1105921108).
- Shiota, Takuya et al. (2015). "Molecular architecture of the active mitochondrial protein gate". In: *Science* 349.6255, pp. 1544–1548.
- Sinnige, Tessa, Anan Yu, and Richard I. Morimoto (2020). "Challenging Proteostasis: Role of the Chaperone Network to Control Aggregation-Prone Proteins in Human Disease". In: *HSF1 and Molecular Chaperones in Biology and Cancer, Advances in Experimental Medicine and Biology*. Ed. by M.L. Mendillo and Et Al. Vol. 1243. Springer Nature Switzerland AG. Chap. 4, pp. 53–68. ISBN: 9783030402044. DOI: [10.1007/978-3-030-40204-4\\_4](https://doi.org/10.1007/978-3-030-40204-4_4).
- Sirrenberg, C. et al. (1998). "Carrier protein import into mitochondria mediated by the intermembrane proteins Tim10/Mrs11 and Tim12/Mrs5". In: *Nature*. ISSN: 00280836. DOI: [10.1038/36136](https://doi.org/10.1038/36136).
- Skjærven, Lars et al. (2015). "Dynamics, flexibility, and allostery in molecular chaperonins". In: *FEBS Lett.* 589.19, pp. 2522–2532. ISSN: 18733468. DOI: [10.1016/j.febslet.2015.06.019](https://doi.org/10.1016/j.febslet.2015.06.019).
- Street, Timothy O., Laura A. Lavery, and David A. Agard (2011). "Substrate Binding Drives Large-Scale Conformational Changes in the Hsp90 Molecular Chaperone". In: *Mol. Cell* 42.1, pp. 96–105. ISSN: 10972765. DOI: [10.1016/j.molcel.2011.01.029](https://doi.org/10.1016/j.molcel.2011.01.029). URL: <http://dx.doi.org/10.1016/j.molcel.2011.01.029>.
- Stull, Frederick et al. (2016). "Protein folding occurs while bound to the ATP-independent chaperone Spy". In: *Nat Struct Mol Biol.* 23.1, pp. 53–58. DOI: [10.1038/nsmb.3133](https://doi.org/10.1038/nsmb.3133). Protein.
- Sučec, Iva et al. (2020). "Structural basis of client specificity in mitochondrial membrane-protein chaperones". In: *Sci. Adv.* 6.51. ISSN: 23752548. DOI: [10.1126/SCIADV.ABD0263](https://doi.org/10.1126/SCIADV.ABD0263).
- Sučec, Iva et al. (2021). "Deciphering the dynamic chaperone – client interactions at the atomic level". In: Svergun, D. I. (1992). "Determination of the regularization parameter in indirect-transform methods using perceptual criteria". In: *J. Appl. Crystallogr.* 25.pt 4, pp. 495–503.
- Swie Goping, Ing, Douglas G. Millar, and Gordon C. Shore (1995). "Identification of the human mitochondrial protein import receptor, huMas20p. Complementation of  $\Delta$ mas20 in yeast". In: *FEBS Letters* 373.1, pp. 45–50. ISSN: 00145793. DOI: [10.1016/0014-5793\(95\)01010-C](https://doi.org/10.1016/0014-5793(95)01010-C).
- Takeda, Hironori et al. (2020). *Mitochondrial sorting and assembly machinery operates by  $\beta$ -barrel switching*. April. Springer US. ISBN: 4158602003113. DOI: [10.1038/s41586-020-03113-7](https://doi.org/10.1038/s41586-020-03113-7). URL: <http://dx.doi.org/10.1038/s41586-020-03113-7>.
- Tavoulari, Sotiria et al. (2019). "The yeast mitochondrial pyruvate carrier is a hetero-dimer in its functional state". In: pp. 1–13. DOI: [10.15252/embj.2018100785](https://doi.org/10.15252/embj.2018100785).
- Terada, Takaho et al. (2016). "Cell-Free Protein Production for Structural Biology". In: *Advanced Methods in Structural Biology*. Ed. by T. Senda and K. Maenaka. Tokyo: Springer Protocols Handbooks, pp. 83–102. ISBN: 9784431560302. DOI: [10.1007/978-4-431-56030-2\\_5](https://doi.org/10.1007/978-4-431-56030-2_5).
- Thoma, Johannes et al. (2015). "Impact of holdase chaperones Skp and SurA on the folding of  $\beta$ -barrel outer-membrane proteins". In: *Nat. Struct. Mol. Biol.* 22.10, pp. 795–802. ISSN: 15459985. DOI: [10.1038/nsmb.3087](https://doi.org/10.1038/nsmb.3087).
- Truscott, Kaye N. et al. (2001). "A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23". In: *Nature Structural Biology* 8.12, pp. 1074–1082. ISSN: 10728368. DOI: [10.1038/nsb726](https://doi.org/10.1038/nsb726).
- Tsirigotaki, Alexandra et al. (2018). "Long-Lived Folding Intermediates Predominate the Targeting-Competent Secretome". In: *Structure* 26.5, 695–707.e5. ISSN: 18784186. DOI: [10.1016/j.str.2018.03.006](https://doi.org/10.1016/j.str.2018.03.006). URL: <https://doi.org/10.1016/j.str.2018.03.006>.
- Tucker, Kyle (2019). "Cryo-EM structure of the mitochondrial protein-import channel TOM complex from *Saccharomyces cerevisiae*". In: 40, pp. 11–13.

- Van Wilpe, Sandra et al. (1999). "Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase". In: *Nature* 401.6752, pp. 485–489. ISSN: 00280836. DOI: [10.1038/46802](https://doi.org/10.1038/46802).
- Vaughan, Cara K. et al. (2006). "Structure of an Hsp90-Cdc37-Cdk4 Complex". In: *Mol. Cell* 23.5, pp. 697–707. ISSN: 10972765. DOI: [10.1016/j.molcel.2006.07.016](https://doi.org/10.1016/j.molcel.2006.07.016). URL: <http://dx.doi.org/10.1016/j.molcel.2006.07.016>.
- Verba, Kliment A. et al. (2016). "Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase". In: *Science* 352.6293, pp. 1542–1547. ISSN: 10959203. DOI: [10.1126/science.aaf5023](https://doi.org/10.1126/science.aaf5023).
- Vitali, Daniela G. et al. (2018). "Independent evolution of functionally exchangeable mitochondrial outer membrane import complexes". In: *eLife* 7, pp. 1–22. ISSN: 2050084X. DOI: [10.7554/eLife.34488](https://doi.org/10.7554/eLife.34488).
- Vranken, Wim F et al. (2005). "The CCPN data model for NMR spectroscopy: development of a software pipeline." In: *Proteins* 59.4, pp. 687–696.
- Wang, Mingzhang et al. (2018). "Fast Magic Angle Spinning <sup>1</sup>F NMR of HIV-1 Capsid Protein Assemblies". In: *Angew. Chem. Int. Ed.*, pp. 5–10.
- Wang, Qiang et al. (2021). "Structural insight into the SAM-mediated assembly of the mitochondrial TOM core complex". In: *Science* 373.6561, pp. 1377–1381.
- Waudby, Christopher A. et al. (2016). "Two-Dimensional NMR Lineshape Analysis". In: *Scientific Reports* 6. February, pp. 1–8. ISSN: 20452322. DOI: [10.1038/srep24826](https://doi.org/10.1038/srep24826).
- Webb, Chaille T. et al. (2006). "Crystal structure of the mitochondrial chaperone TIM9•10 reveals a six-bladed  $\alpha$ -propeller". In: *Molecular Cell* 21.1, pp. 123–133. ISSN: 10972765. DOI: [10.1016/j.molcel.2005.11.010](https://doi.org/10.1016/j.molcel.2005.11.010).
- Weickert, S. et al. (2020). "The mechanism of Hsp90-induced oligomerization of Tau". In: *Sci. Adv.* 6.11. ISSN: 23752548. DOI: [10.1126/sciadv.aax6999](https://doi.org/10.1126/sciadv.aax6999).
- Weinhäupl, Katharina et al. (2018). "Structural Basis of Membrane Protein Chaperoning through the Mitochondrial Intermembrane Space". In: *Cell* 175, pp. 1–15. ISSN: 10974172. DOI: [10.1016/j.cell.2018.10.039](https://doi.org/10.1016/j.cell.2018.10.039).
- Weinhäupl, Katharina et al. (2021). "Architecture and assembly dynamics of the essential mitochondrial chaperone complex TIM9-10-12". In: *Structure*, pp. 1–9. ISSN: 09692126. DOI: [10.1016/j.str.2021.04.009](https://doi.org/10.1016/j.str.2021.04.009).
- Whittaker, Sara B.M. et al. (2007). "NMR Analysis of the Conformational Properties of the Trapped on-pathway Folding Intermediate of the Bacterial Immunity Protein Im7". In: *J. Mol. Biol.* 366.3, pp. 1001–1015. ISSN: 00222836. DOI: [10.1016/j.jmb.2006.11.012](https://doi.org/10.1016/j.jmb.2006.11.012).
- Wiedemann, Nils and Nikolaus Pfanner (2017). "Mitochondrial Machineries for Protein Import and Assembly". In: DOI: [10.1146/annurev](https://doi.org/10.1146/annurev). URL: <https://doi.org/10.1146/annurev->.
- Wiedemann, Nils, Nikolaus Pfanner, and Michael T. Ryan (2001). "The three modules of ADP/ATP carrier cooperate in receptor recruitment and translocation into mitochondria". In: *EMBO Journal* 20.5, pp. 951–960. ISSN: 02614189. DOI: [10.1093/emboj/20.5.951](https://doi.org/10.1093/emboj/20.5.951).
- Wu, Yunkun and Bingdong Sha (2006). "Crystal structure of yeast mitochondrial outer membrane translocon member Tom70p". In: *Nature Structural & Molecular Biology*. ISSN: 1545-9993. DOI: [10.1038/nsmb1106](https://doi.org/10.1038/nsmb1106).
- Xie, Yansheng and Donald B. Wetlaufer (1996). "Control of aggregation in protein refolding: The temperature-leap tactic". In: *Prot. Sci.* 5.3, pp. 517–523. ISSN: 09618368. DOI: [10.1002/pro.5560050314](https://doi.org/10.1002/pro.5560050314).
- Yamamoto, Hayashi et al. (2009). "Roles of Tom70 in import of presequence-containing mitochondrial proteins." In: *The Journal of biological chemistry* 284.46, pp. 31635–46. ISSN: 1083-351X. DOI: [10.1074/jbc.M109.041756](https://doi.org/10.1074/jbc.M109.041756). URL: <http://www.ncbi.nlm.nih.gov/pubmed/19767391><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2797234>.

- Yamano, Koji et al. (2008a). "Step-size analyses of the mitochondrial Hsp70 import motor reveal the Brownian ratchet in operation". In: *J. Biol. Chem.* 283.40, pp. 27325–27332.
- Yamano, Koji et al. (2008b). "Tom20 and Tom22 share the common signal recognition pathway in mitochondrial protein import". In: *Journal of Biological Chemistry* 283.7, pp. 3799–3807. ISSN: 00219258. DOI: [10.1074/jbc.M708339200](https://doi.org/10.1074/jbc.M708339200).
- Yin, Yanting et al. (2021). "Structural basis for aggregate dissolution and refolding by the Mycobacterium tuberculosis ClpB-DnaK bi-chaperone system". In: *Cell Rep.* 35.8, p. 109166. ISSN: 22111247. DOI: [10.1016/j.celrep.2021.109166](https://doi.org/10.1016/j.celrep.2021.109166). URL: <https://doi.org/10.1016/j.celrep.2021.109166>.
- Young, Jason C., Nicholas J. Hoogenraad, and F. Ulrich Hartl (2003). "Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70". In: *Cell* 112.1, pp. 41–50. ISSN: 00928674. DOI: [10.1016/S0092-8674\(02\)01250-3](https://doi.org/10.1016/S0092-8674(02)01250-3).
- Young, Jason C et al. (2004). "Pathways of chaperone-mediated protein folding in the cytosol." In: *Nat. Rev. Mol. Cell Biol.* 5.10, pp. 781–791.
- Zhou, Huan Xiang, Xiaodong Pang, and Cai Lu (2012). "Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets". In: *Physical Chemistry Chemical Physics* 14.30, pp. 10466–10476. ISSN: 14639076. DOI: [10.1039/c2cp41196b](https://doi.org/10.1039/c2cp41196b).