Atomic-resolution studies of structure, dynamics and interactions in chaperone assemblies by NMR spectroscopy.

Katharina Weinhäupl

Communauté UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : Biologie Structurale et Nanobiologie

Arrêté ministériel : 25 mai 2016

Présentée par

Katharina WEINHÄEUPL

Thèse dirigée par **Paul SCHANDA (EDCSV)**, CEA

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

Etudes de structure, interactions et dynamique dans des complexes de protéines "chaperone" à l'échelle atomique par spectroscopie RMN

Atomic-resolution studies of structure, dynamics and interactions in chaperone assemblies by NMR spectroscopy.

Thèse soutenue publiquement le **11 janvier 2018**, devant le jury composé de :

Monsieur PAUL SCHANDA
CADRE SCIENTIFIQUE, CEA GRENOBLE, Directeur de thèse
Monsieur SEBASTIAN HILLER
PROFESSEUR, UNIVERSITE DE BÂLE - SUISSE, Rapporteur
Monsieur MICHAEL PLEVIN
MAITRE DE CONFERENCES, UNIVERSITE D'YORK - ROYAUME-UNI, Rapporteur
Monsieur DORON RAPAPORT
PROFESSEUR, UNIVERSITE DE TÜBINGEN - ALLEMAGNE, Examinateur
Monsieur PAU BERNADO
CHARGE DE RECHERCHE, INSERM DELEGATION LANGUEDOC-ROUSSILLON, Examinateur
Monsieur FRANCK FIESCHI
PROFESSEUR, UNIVERSITE GRENOBLE ALPES, Président

0.1 Abstract

The diverse group of molecular chaperones is dedicated to accompany, fold and protect other proteins until they reach their final conformation and location inside the cell. To this end, molecular chaperones need to be specialized in performing specific tasks, like folding, transport or disaggregation, and versatile in their recognition pattern to engage many different client proteins. Moreover, molecular chaperones need to be able to interact with each other and with other components of the protein quality control system in a complex network. Interactions between the different partners in this network and between the substrate and the chaperone are often dynamic processes, which are especially difficult to study using standard structural biology techniques. Consequently, structural data on chaperone/substrate complexes are sparse, and the mechanisms of chaperone action are poorly understood. In this thesis I present investigations of the structure, dynamics and substrate-interactions of two molecular chaperones, using various biophysical and *in vivo* methods.

In the first part I show that the mitochondrial membrane protein chaperone TIM910 binds its substrates in a highly dynamic manner. Not only is the TIM910 complex in constant exchange between monomeric and hexameric species, but also the bound substrate samples multiple conformations on a millisecond timescale. Based on nuclear magnetic resonance (NMR), small-angle X-ray scattering (SAXS), analytical ultracentrifugation (AUC) and *in vivo* mutational experiments I propose a structural model of the chaperone/membrane protein interaction. TIM910 binds its substrates in a hydrophobic pocket on the exterior of the chaperone in a modular fashion, where the number of TIM910 complexes bound depends on the length of the substrate.

In the second part I studied the behavior of the N-terminal receptor domain of the ClpC1 unfoldase from *M.tuberculosis* in the presence of different antibiotics and ligands. The N-terminal domain of ClpC1 is the binding site for various new antibiotics against *M.tuberculosis*. The antibiotic cyclomarin completely abolishes dynamics induced by the ligand arginine-phosphate. We propose that this suppression of dynamics is the underlying principle for the mechanism of action of this antibiotic.

In both cases X-ray structures of the apo or antibiotic bound form were available, but not sufficient to explain the mechanism of action. The X-ray structure of TIM910 provided no evidence on where or how substrates are bound. Likewise, X-ray structures of the apo and cyclomarin-bound N-terminal domain of ClpC1 show only minor differences in structure.

Both examples show that static structural data is often not enough to

explain how a molecular system works, and only the combination of different techniques, including newly developed methods enable the atomic-level understanding of chaperone/substrate complexes.

0.2 Resume

Les chaperons moléculaires, une famille de protéines diverses en structure et taille, sont dédiés à accompagner, replier et protéger d'autres protéines afin qu'elles atteignent leur conformation finale et leur emplacement dans la cellule. Dans ce but, les chaperons moléculaires doivent être hautement spécialisés dans l'exécution de tâches spécifiques, telles que le repliement, le transport ou la désagrégation, et polyvalents dans leur motifs de reconnaissance, afin de pouvoir interagir avec un grand nombre de protéines différentes. Différents chaperons moléculaires collaborent au sein de la cellule, formant ainsi un réseau complexe qui assure le contrôle de la qualité du protéome. Les interactions entre les différents partenaires de ce réseau et entre les chaperones et leurs substrats sont souvent dynamiques, ce qui rend leur observation structurale particulièrement difficile pour les techniques de biologie structurale. Par conséquent, il y a à ce jour peu d'information sur les structures et mécanismes d'interaction au sein des complexes chaperon-substrate. Dans cette thèse, je présente des études sur la structure, la dynamique et les interactions entre les substrats de deux chaperons moléculaires, en utilisant diverses méthodes biophysiques et *in vivo*.

Dans la première partie, je montre que la chaperone TIM910, située dans l'espace inter-membranaire des mitochondries, lie ses substrats, des protéines membranaires destinées aux deux membranes mitochondriales, d'une manière très dynamique. Non seulement le complexe TIM910 est en échange constant entre les espèces monomèriques et hexameriques, mais aussi le substrat lié échange entre mulitples conformations à une échelle de millisecondes. Sur la base de la résonance magnétique nucléaire (RMN), de small-angle X-ray scattering (SAXS), de l'ultracentrifugation analytique (AUC) et des expériences mutationnelles *in vivo* et des tests fonctionnels d'import dans les mitochondries, je propose un modèle structurale de l'interaction entre le chaperon et la protéine membranaire. TIM910 lie ses substrats dans une poche hydrophobe à l'extérieur du chaperon. Cette interaction est modulaire et se fait avec un ou deux hexamères de TIM910, en fonction de la longueur du substrat.

Dans la deuxième partie, nous avons étudié le comportement du récepteur N-terminal du unfoldase ClpC1 de *M. tuberculosis* en présence d'antibiotiques et de ligands différents. Le domaine N-terminal de ClpC1 est le site de liaison de divers antibiotiques nouveaux contre *M. tuberculosis*. L'antibiotique Cyclomarin A supprime complètement la dynamique induite par le ligand arginine-phosphate. Nous proposons que cette suppression de la dynamique soit le principe fondamental du mécanisme d'action de cet antibiotique.

Dans les deux cas, les structures X-ray des chaperons dans leur état apo et la structure de ClpC-NTD liée à des antibiotiques étaient disponibles,

mais ces structures statiques ne suffisent pas pour expliquer le mécanisme d'action. La structure X-ray de TIM910 n'a pas fourni d' indication sur l'endroit ou la façon dont les substrats sont liés. De même, les structures X-ray du domaine N-terminal de apo et de Cyclomarine A de ClpC1 ne présentent que des différences de structure mineures. Les deux exemples montrent que les données structurelles statiques souvent ne permettent pas d'expliquer le fonctionnement d'un système moléculaire, donc la combinaison de différentes techniques et le développement de nouvelles méthodes pour étudier les complexes chaperon-substrat sont primordiaux pour comprendre leur fonction.

0.3 Collaborations

Collaborations and the use of different techniques play an important role in this PhD thesis. Many new findings that have been made during this work would not have been possible without the expertise and support from several people. Therefore, I would like to acknowledge here the work of our collaborators.

All SAXS data analysis has been performed by Martha Brennich from the EMBL Grenoble, who has been also very supportive in designing and setting up the SAXS measurements.

In vivo yeast mutation experiments and mitochondrial import assays have been conducted by Nils Wiedemann from the University of Freiburg and Caroline Lindau, his PhD student.

The VDAC peptides for binding assays of outer membrane proteins have been provided by the group of Doron Rapaport from the University of Tübingen. Additional cross-linking experiments between TIM910 and the VDAC peptides to support the NMR data were done by Tobias Jores.

Native mass spectrometry of the hexameric TIM91012 complex has been performed by Elisabetta Boeri-Erba from the IBS in Grenoble.

Furthermore, we used the in house analytical ultracentrifugation platform and the cell-free platform. For their help and support I would like to thank the managers of these platforms Christine Ebel, Aline Roy and Lionel Imbert.

For the ClpC1 NTD project I would like to recognize the work of Uli Kazmaier, who synthesized the cyclomarin antibiotic.

At last and most importantly I would like to acknowledge the work of Audrey Hessel throughout the whole project, especially for advice in cloning strategies and making point mutations in the binding site of TIM910.

Contents

List of Figures

Chapter 1

Introduction

1.1 Chaperones

If one searches in the Oxford dictionary for the word chaperone, the following description is given: "A person who accompanies and looks after another person or group of people". Originating from the french word "chaperon", meaning "hood" and being used as a symbol for giving protection.

In biology the term chaperone has been very fittingly given to a family of proteins that protects and guides their client proteins from an immature state to their final active form. Per definition a chaperone can be "any protein that interacts with, stabilizes, or helps another protein to acquire its functionally active conformation."[43] This defininition is of course very broad and already implies that the family of chaperones is extremely diverse in their function, structures and mechanisms of action. Chaperones can be found in any part of the cell and most proteins depend on chaperones to achieve their final conformation.[43] Indeed, chaperones are not only responsible for guiding proteins through a productive folding pathway, but also for the removal of folding incompetent intermediates or aggregates.[17,21] Thus, the implication of molecular chaperones in protein quality control mechanisms is vital for cell survival. Proper function of these machineries is supposed to be a major factor in cellular health and ageing.[5,49,127]

Most small proteins of up to around 100 amino acids are able to fold by themselves without the help of a chaperone. Thus, indicating that the information about the tertiary structure is already included in their amino acid sequence, as has been shown by Anfinsen in 1973.[4] Nevertheless, many large proteins or proteins that are part of multimolecular complexes need assistance in folding.[43] The need for chaperones is even more pronounced in the *in vivo* cellular environment where proteins have to be folded in a

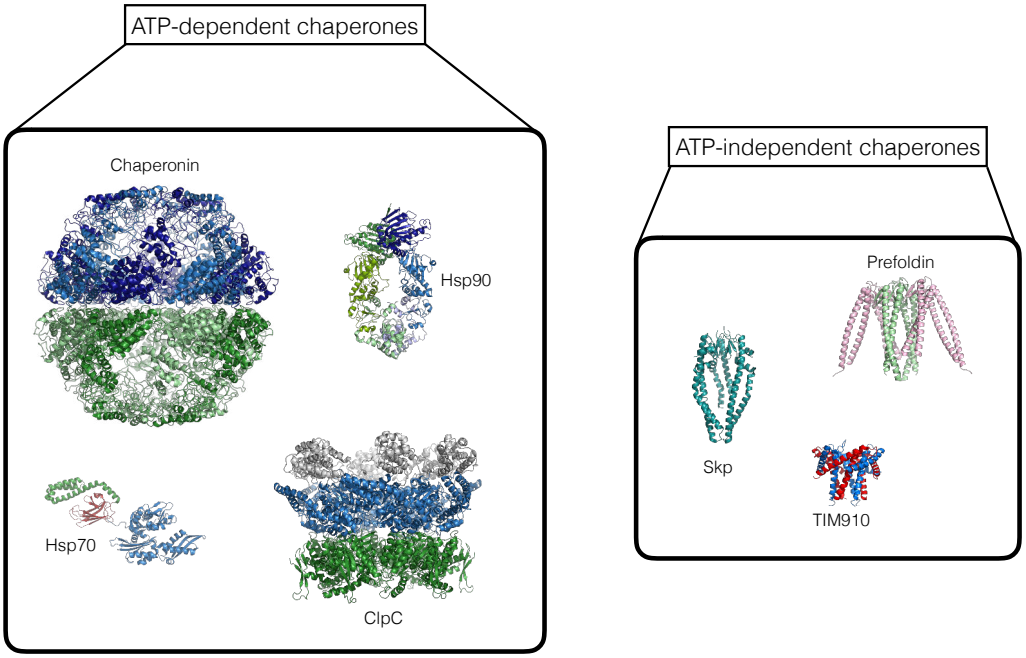


Figure 1.1: **Classification of chaperones according to the use of ATP.** One possible strategy to group chaperones is the use of ATP. ATP-dependent chaperones often use large nucleotide induced domain movements to interact and fold their client proteins, while ATP-independent chaperones often take over more passive tasks like client protein transport.

crowded environment (300-400 g protein L^{-1}).[32] Also membrane proteins need continuous guarding by chaperones from the moment of their translation up to their insertion into the membrane, due to their intrinsically hydrophobic and aggregation prone character. Exposure of hydrophobic patches is the main driving force for protein aggregation. Since protein aggregates can have deleterious effects on cells, proteins evolved to minimize strongly aggregating hydrophobic sequences and protect those sequences by surrounding them with arginine, lysine and proline residues. At the same time, chaperones evolved to specifically target those sequences.[101]

Chaperones can be categorized according to different criteria. One such criterium is their dependence on adenosine triphosphate (ATP) to fulfill chaperone function. While the group of ATP-dependent chaperones is able to perform a multitude of "active" functions, like protein folding or disaggregation, ATP-independent chaperones mostly adopt a "passive" role by either transporting substrates between different compartments and/or keeping them in a folding competent state while bound to the chaperone. For this reason ATP

independent-chaperones are often referred to as “holdases”.

The chaperone network in the cell is very versatile. A given chaperone is not just responsible for a multitude of different client proteins, but the different chaperone complexes also work together to ensure correct folding and localization of their substrates.[126]

Most proteins need chaperones from their translation until they reach their final conformation and location. The first chaperone a freshly translated protein encounters is the nascent-polypeptide binding complex (or trigger factor in bacteria). This ATP-independent chaperone is associated with the exit tunnel of the ribosomal complex and keeps nascent proteins from aggregating.[34] The protein can then either fold by itself or be handed over to the Hsp70/Hsp40 (DnaK/DnaJ in bacteria) complex for further cycles of ATP assisted binding and release.[156] If the protein cannot be folded by the Hsp70 system it is further translocated to the Hsp90 system[79] or the Tric/CCT (GroEL/GroES in bacteria) system.[114] The chaperone cascade can be still continued by a vast number of specialized chaperones, that can help client proteins to insert into a membrane (like the insertase BamA) or localize to a certain cellular compartment (like the holdases TIM910 or Skp which transport membrane proteins in the mitochondrial intermembrane space or in the bacterial periplasm). If the client protein fails to fold, it can be targeted for degradation at several steps of the folding cascade.[17,21]

The following list of chaperone families is by far not complete and should only serve as an introduction for the reader into the structural diversity and different mechanisms chaperones use to handle their clients. I selected chaperones according to the availability of information (i.e. publications on PubMed) thereby focusing especially on the existence of structural data, dynamic measurements or interaction with client proteins. Furthermore, I selected examples that might serve to discuss the functionality of the two chaperones I worked with (TIM910 and ClpC1) and to introduce those systems.

1.1.1 Examples of ATP-dependent chaperones

The use of ATP and their implication in protein quality control are the only common features for this diverse class of chaperones. This family of proteins comprises, among others, the classical chaperones Hsp70, Hsp90 and Tric/CCT that assist protein folding, unfoldases like ClpX or disaggregases like ClpB. Many of these components can interact with each other and are connected in an intricate network to prevent protein misfolding and aggregation. Frequently ATP triggers large-scale rearrangements, like cavity openings or domain motions. A description of the most prominent members and

some of their binding partners is given below.

Hsp70 and Hsp40

Hsp70 (DnaK in *E.coli*) is the most abundant cellular chaperone and can be found in all domains of life.[7] Hsp70 together with its co-chaperones fulfills several functions in the cell: It can bind misfolded proteins and promote their folding by repeated ATP-driven binding and release cycles, it can bind unfolded proteins and deliver them to different cellular compartments[150] or it can target terminally misfolded or aggregated proteins for degradation, either by keeping the protein in a degradation competent state or by interaction between the chaperone and the degradation network.[3]

Hsp70 consists of a nucleotide binding domain (NBD) and a C-terminal substrate binding domain (SBD). Additionally, the SBD is divided into an α-helical lid domain (LID) and a β-sandwich domain for substrate binding.[11] Hsp70 has to interact with at least two co-chaperones to perform its function. One is a nucleotide exchange factor (NEF), the other Hsp40 (also called J-protein or DnaJ in *E.coli*).[109] In the canonical folding pathway Hsp40 binds to substrate proteins, thus keeping them from aggregation, delivers them to Hsp70 and stimulates Hsp70 ATPase activity, thereby enhancing the ability of Hsp70 to interact with the substrate.[120] The substrate is then released upon NEF stimulated adenosine diphosphate (ADP) dissociation and free to fold.[80] If the folding fails and hydrophobic patches are still exposed, the substrate can rebind and go through another folding cycle.

There are two distinct states in the folding cycle of Hsp70. One is the open (ATP-bound) state, where the SBD and LID domain are docked to the NBD domain. In this way the binding site of the SBD is open and free to bind exposed hydrophobic patches of client proteins. In the closed (ADP-bound) state the LID domain covers the SBD binding site and thus traps the bound substrate.[63] Therefore, client proteins in the open state have a high on/off rate and thus a low affinity, but a low on/off rate and thus a high affinity in the closed state. It has been shown by Schlecht et al. that Hsp70 is in an equilibrium between open and closed state when it is bound to ATP, ADP or in a nucleotide free state. Binding of the nucleotide only shifts the population to either mainly a closed or an open states.[107]

Although there are several X-ray structures of the closed state of Hsp70,[11,116] determination of the structure of open Hsp70 could only be achieved by cross-linking NBD and SBD according to a model based on a homologous protein that is in a locked open conformation. The cause for this inability to crystallize are most likely protein dynamics in the open state.[63] A detailed investigation of conformational dynamics in the SBD has been conducted by

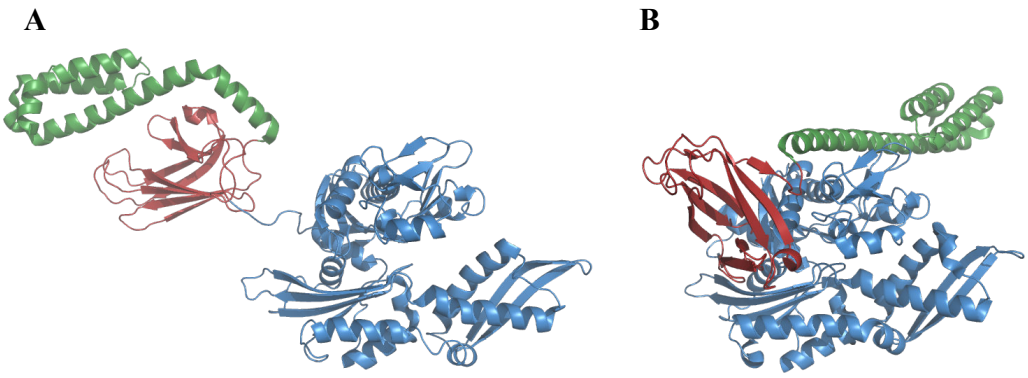


Figure 1.2: **X-ray strutures of the Hsp70 open and closed state.** The Hsp70 chaperone in its ATP bound open (**B**, PDB: 4B9Q) and ADP/substrate bound closed (**A**, PDB: 2KHO) conformation. The LID domain is shown in green, the SBD in red and the NBD in blue.

Zhuraleva et al.[154,155] They found that ATP binding to the NBD and the resulting allosteric change of the SBD cause enhanced conformational exchange in the substrate binding site and that this change in allostery is propagated by a network of hydrophobic residues. This conformational flexibility of the substrate binding site can be observed as millisecond exchange dynamics by NMR and might provide Hsp70 the ability to bind to a wide range of different substrates.[154]

Hsp70 can act on misfolded, unfolded or aggregated substrates. To accomodate these client proteins in its binding site, the LID can adopt different conformations to adapt to the size of the substrate. Although it is not necessary that the LID closes over the bound substrate, contact between the LID and the substrate is indispensable for chaperone function.[107]

The mechanism of action of Hsp70 seems to be based on the removal of tertiary contacts in the client protein, while local secondary elements are allowed to fold. To perform their chaperone function, one client protein can also bind more than one Hsp70 molecule.[110]

The versatility of the Hsp70 machinery is most likely based on the modular architecture including NEFs and Hsp40s. While there are only 11 Hsp70s in humans, there are 41 different Hsp40s, thus allowing specialization on different targets.[57] These Hsp40s differ significantly in their sequence and structure. The only common feature is the so called J-domain. This 70 amino acid long sequence contains a highly conserved histidine, proline, aspartic acid (HPD) loop motif that is indespensable for its ATPase stimulating function.[120,124] The position of the J-domain and the presence of other

sequence motifs, like substrate binding domains or motifs for degradation targeting, depends on the Hsp40 class and is also within these classes highly variable.[57] The most minimalistic version of Hsp40 contains only a J-domain and no substrate binding domain. In this case, Hsp40 does not deliver the substrate to Hsp70, but only stimulates ATPase activity and thus substrate binding of Hsp70.[47] Furthermore, Hsp40 can be free, tethered to a membrane or enriched in certain cellular compartments. This membrane binding can occur for example at sites with high client protein concentration to stimulate Hsp70 interaction.[86]

Hsp90

Hsp90 is a ubiquitous chaperone found in the cytosol, the nucleus and different organelles and is tightly connected to the Hsp70 system and binds substrates in a "late" folding stage.[59] Hsp90 forms a homodimer, with each monomer containing an ATP binding N-terminal domain (NTD), a flexible linker, a substrate binding middle domain (MD) and a C-terminal domain (CTD) for homodimerization. Additionally, Hsp90 contains several interaction sites for its various co-chaperones that regulate ATPase activity, post-translational modifications, binding of specific client proteins or interactions with other cellular networks.[71]

Nucleotide-free Hsp90 homodimers form a highly dynamic V-shaped structure, connected by the two CTDs. Upon ATP binding (and interaction with co-chaperones) the two NTDs intertwine and form a closed state. This final conformation is preceded by at least two intermediate states. ATP hydrolysis results in an even more compact conformation and finally ADP/Pi release returns Hsp90 to the inital state.[111] As has been shown for Hsp70, Hsp90 is most likely continuously sampling open and closed conformations in solution, which can be selected by the client protein.[118]

In the first step of its chaperone cycle Hsp90 has to interact with Hsp70. This is mediated by the co-chaperone Hop. A low resolution electron microscopy (EM) reconstruction of Hsp90 in complex with Hsp70, Hop and a substrate reveal the different steps of this reaction.[61] During this interaction the client protein is bound simultaneouly to Hsp90 and Hsp70 and the SBD from Hsp70 with the bound substrate is positioned at the MD/CTD substrate binding site of Hsp90. Furthermore, the two ATPase domains of Hsp90 and Hsp70 are found to be in close contact, suggesting that the two hydrolysis events are coupled. ATP hydrolysis of Hsp90 is then used for release of Hsp70 and transfer of the substrate to Hsp90. The chaperone function (i.e. the ability to keep client proteins from aggregation) of Hsp90 itself is not ATP-dependent. Interestingly, the complex is not symmetric as

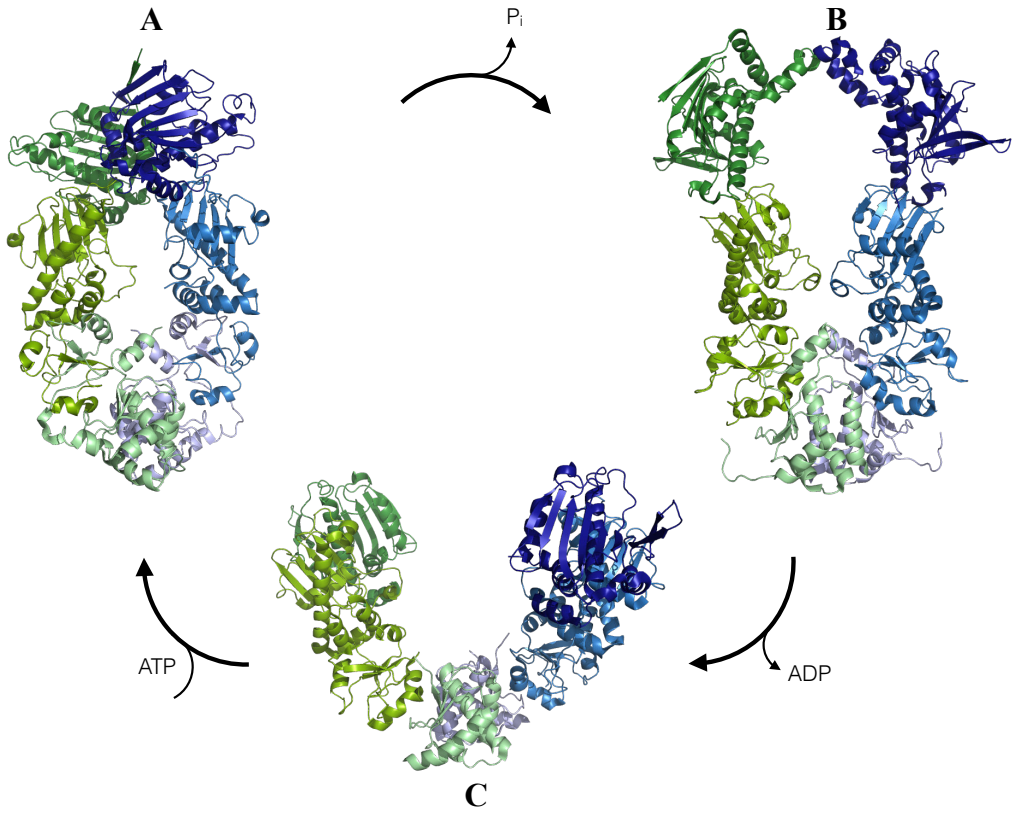


Figure 1.3: **X-ray structures of the Hsp90 ATP, ADP and nucleotide free states.** Hsp90 in its ATP bound closed state (**A**, PDB: 2CG9), in the compact ADP state (**B**, PDB: 2IOP) and the open nucleotide free state (**C**, PDB: 2IOQ). The NTDs are shown in dark blue/green, the MDs in middle blue/green and the CTDs in light blue/green.

one would expect of a homodimer, rather it seems that first only one of the NTDs adopts the closed conformation.[61] Over 200 proteins have been identified as Hsp90 client proteins,[31] among them many chaperones, transcription factors and kinases. Several of those proteins are implicated in tumor genesis or progression making Hsp90 an interesting target for the pharmaceutical industry. While Hsp90 can act as a classical chaperone by assisting folding or preventing misfolding, it plays a specific role in keeping transcription factors or kinases in a nearly natively folded state to provide a fast reaction upon activation. Thus, Hsp90 could offer a storage mechanism for cellular signalling molecules.[125]

How Hsp90 binds its substrate proteins is still not fully understood. Hsp90 preferentially binds folding intermediates, a specific amino acid sequence or structural motif seems not to be a prerequisite. A recent cryo-

EM structure of Hsp90 in complex with its kinase co-chaperone Cdc37 and the kinase Cdk4 suggests that Hsp90 binds a folding intermediate of Cdk4, whereby the chaperone bound middle part stays unfolded and the N- and C-terminal parts remain folded.[134] However, Street et al. found that Hsp90 preferentially binds elements that are structured in the apo substrate.[118] Furthermore, Hsp90 also interacts with the intrinsically disordered protein (IDP) Tau, binding to aggregation-prone repeats of the protein. Comparing the binding site of Hsp90 and Hsp70 revealed that the binding site of Hsp90 is much broader and less hydrophobic than the Hsp70 site. This could explain how Hsp90 recognizes client proteins in a later folding stage, that will exhibit more scattered and less hydrophobic patches than Hsp70 substrates. Thus, Hsp90 would be able to equally bind folding intermediates and IDPs that both exhibit scattered hydrophobic regions.[59] The conformation of Hsp90 bound tumor suppressor protein p53 is controversial. While Hagn et al. propose that p53 binds in a folded state, Park et al. detected a molten globule state[89] for p53 when bound to Hsp90.[42]

Chaperonins

Chaperonins are a group of chaperones that exist in all domains of life from bacteria to eukaryotes. They promote folding by trapping their substrates in a large folding chamber. Opening and closing of this chamber is ATP regulated. All chaperonins share a similar architecture: they form two back to back rings with an equatorial, intermediate and apical domain. Additionally chaperonins contain a lid for closing of the folding chamber. In group I chaperonins the lid is formed by a separate co-chaperone, while in group II chaperonins the lid is built into the apical domain. Group I chaperonins can be found in bacteria (GroEL/GroES) and organelles (Hsp60/Hsp10) derived from endosymbiosis, while group II chaperonins are found in archea (thermosomes) and the eukaryotic cytosol (TriC also named CCT). I will introduce the reaction cycles of one member of each of these groups, GroEL-GroES for group I and TriC/CCT for group II.[114]

GroEL is formed by two times seven identical subunits, with ATP binding pockets in each subunit. Apo GroEL exposes a continuous ring of hydrophobic patches, that could be used for binding of unfolded or misfolded substrate proteins.[144] The back to back rings are oriented in a staggered conformation, meaning that one subunit always contacts two subunits from the opposite ring. Upon ATP binding, the intermediate and apical domains tilt and move outwards via several intermediate states until they reach a completely open state. During the intermediate states the subunits expose different binding surfaces, likely to serve the binding of a variety of substrate proteins. In the

open state the interdomain apical contacts are broken and thus also the hydrophobic ring formed by these domains. The opening process could force a bound substrate protein to stretch and thus unfold. Opposing to this theory is the fact that ATP binding proceeds faster than substrate binding, making it more likely that a substrate encounters a chaperonin that is already in an intermediate/open state.[19] The movement to the open state positions the GroEL ring in a way that it can receive and bind the mobile loops of its co-chaperone GroES. Binding of GroES enlargens and caps the folding chamber and induces a clockwise twist that projects the substrate into the negatively charged folding chamber where it is free to fold.[19] Interestingly also apo GroEL shows intrinsic unfoldase/foldase activity.[73] The mechanism of release and initiation of a new cycle is still controversial. It is not known if either ATP hydrolysis in one ring causes binding of ATP to the second ring and thereby release of ADP and substrate in the first ring and inititation of a new cycle in the second ring or if folding cycles can take place simultaneously in both rings. Although historically the first theory was favoured, the observation of particles with two closed chambers contradicts this view.[76]

In contrast to the archeal thermosome that is formed mostly of eight identical subunits per ring, the eukaryotic TriC complex is made of eight similar but distinct subunits (CCCT1-CCT8). Due to the similarity of these subunits their arrangement has been unknown for a long time, despite the existence of cryo-EM and X-ray structures. Leitner et al. solved this problem by using a combination of cross-linking, mass spectometry and combinatorial modelling. The resulting arrangement revealed a surprising assymetry of the complex and led to the discovery of several new features of this chaperone.[70] While thermosome and GroEL-GroES form symmetric complexes the folding chamber of TriC is segregated into a positively charged part formed by CCT5/2/4 and a negatively charged part formed by CCT3/6/8. Whereby substrate binding occurs only in the negatively charged part.[70] Although all subunits of TriC are able to bind nucleotides, only CCT5/2/4 can hydrolyse ATP. This separation of TriC in a substrate binding and nucleotide hydrolysing part shows a much higher specialization than the other chaperonins and might have developped hand in hand with an increasingly larger proteome.[56] Contrary to thermosome and GroEL-GroES that bind rather unspecifically to hydrophobic patches, TriC is specialized in chaperoning certain client proteins. Nevertheless, TriC does not recognize distinct patterns in its substrates but binds to proteins with diverse sequences and folds. Among them cell-cycle proteins, signalling proteins, human immunodeficiency virus (HIV) proteins, aggregation prone proteins and many more.[148] Although the binding sites in the different subunits all contain a mixture of hydrophobic and polar residues and substrate binding is multivalent, certain substrates

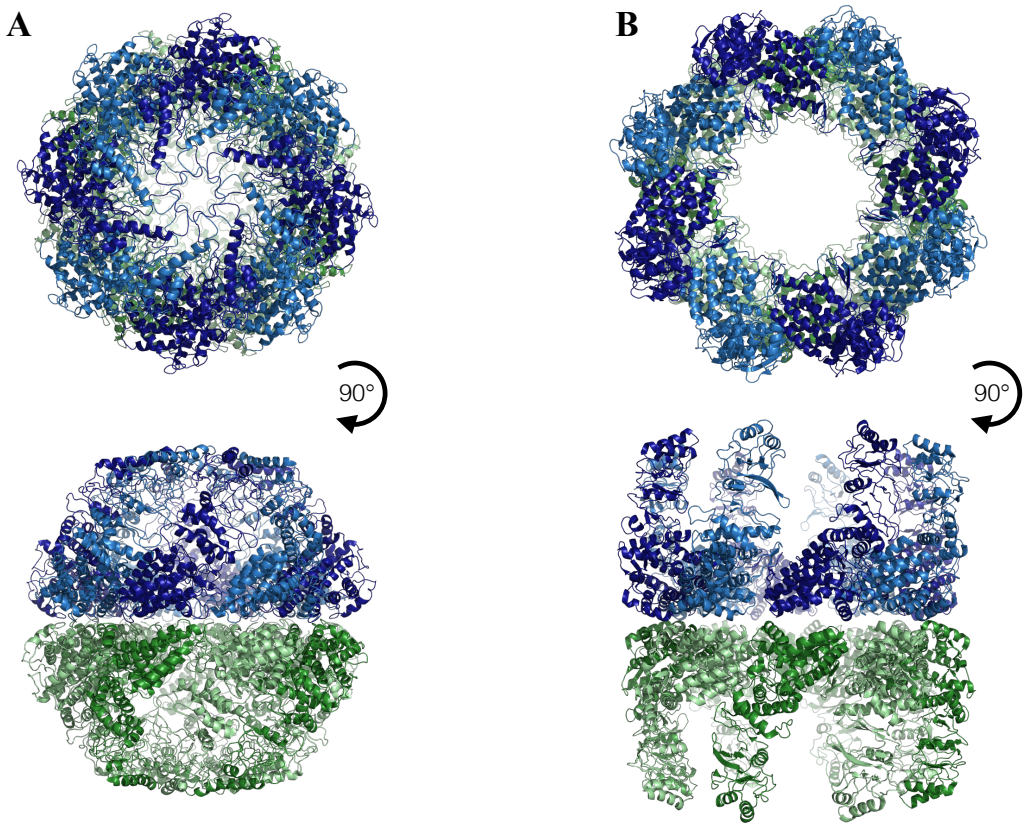


Figure 1.4: **An archeal chaperonin in open and closed conformation.** The Mm-cpn group II chaperonin of *Methanococcus maripaludis* in nucleotide-free open (**B**, PDB: 3IYF) and nucleotide-induced closed state (**A**, PDB: 3LOS). The chaperonin forms a homo 16-mer, different colours should only help to differentiate the two chambers and the individual chains.

have a preference for certain subunits. The HIV protein p6 for example has a higher affinity for CCT3 than for all the other subunits.[54] ATP binding and subsequent closure of the ring causes the release of the substrate into the interior of the folding chamber. The assymetric hydrolysis efficiency of the different subunits may lead to a sequential release of the substrate and thus lead it along a predetermined folding pathway.[56] Thus, assymetry might be an important feature in the function of this chaperone.

The ATP-independent chaperone Prefoldin works in close contact with TriC. Prefoldin has a "jellyfish" like structure, with a β-barrel core and α-helical coiled coil "tentacles". With the tips of its "tentacles" it interacts with nascent polypeptides and delivers them to TriC.[113] A more detailed description of how ATP-independent chaperones interact with their substrates can

be found in the section "Examples of ATP-independent chaperones".

AAA+ proteins

Members of the AAA+ (ATPases Associated with diverse cellular Activities) family are important players in protein homeostasis. These proteins share a common domain architecture with typically one or two ATPase domains (termed D1 and D2) and an additional N-terminal domain (NTD) important for substrate recognition. ClpA, ClpB, ClpC and ClpX are well-known representatives of this family. Their purpose in protein quality control is the removal of protein aggregates or the targeting of proteins for degradation by the serine protease ClpP. While ClpB serves as a disaggregase, ClpA, ClpC and ClpX unfold and translocate substrates into the pore of the ClpP complex for degradation.[29]

Since part of my experimental work was devoted to ClpC1 from *M. tuberculosis*, I will focus here mainly on describing the structure and function of this particular AAA+ unfoldase.

The only known structure of a ClpC complex is from *B. subtilis* and in complex with its adaptor protein MecA.[142] The intrinsic flexibility and exchange dynamics between different oligomeric states impeded structural studies of most of the above mentioned Clp proteins. Crystallisation of ClpC was only possible upon removal of flexible loop regions, rendering the protein non-functional and at the same time underlining the importance of dynamics for the function of these complexes. Although ClpC is a ubiquitous and conserved protein in bacteria, slight differences between bacterial families exist, for example in the use of accessory proteins like MecA. The functional state of ClpC is a homohexamer. In *B. subtilis* ClpC, only binding of the adaptor protein MecA and ATP induces the formation of a functional hexamer, while in *M. tuberculosis* ATP alone is sufficient for hexamer formation and no homologuous adaptor protein has so far been described for *M. tuberculosis* ClpC1.[62] In the eukaryotic proteasome system, proteins destined for degradation are ubiquitinylated and thus recognized by the proteasome. In bacteria the targeting signal for ClpCP degradation has been unknown for a long time and only been elucidated by Trentini et al. in 2015.[129] The kinase McsB phosphorylates target proteines on their arginine residues. These phosphorylation sites are often found in secondary structure elements and thus only accessible when the protein is unfolded or misfolded. The binding site of arginine phosphate is located in the NTD of ClpC. A crystal structure of *B. subtilis* ClpC NTD shows two arginine phosphates bound to the protein, while the 3D structure of the protein stays the same as in apo NTD. Interestingly, binding of arginine phosphorylated substrate alone can

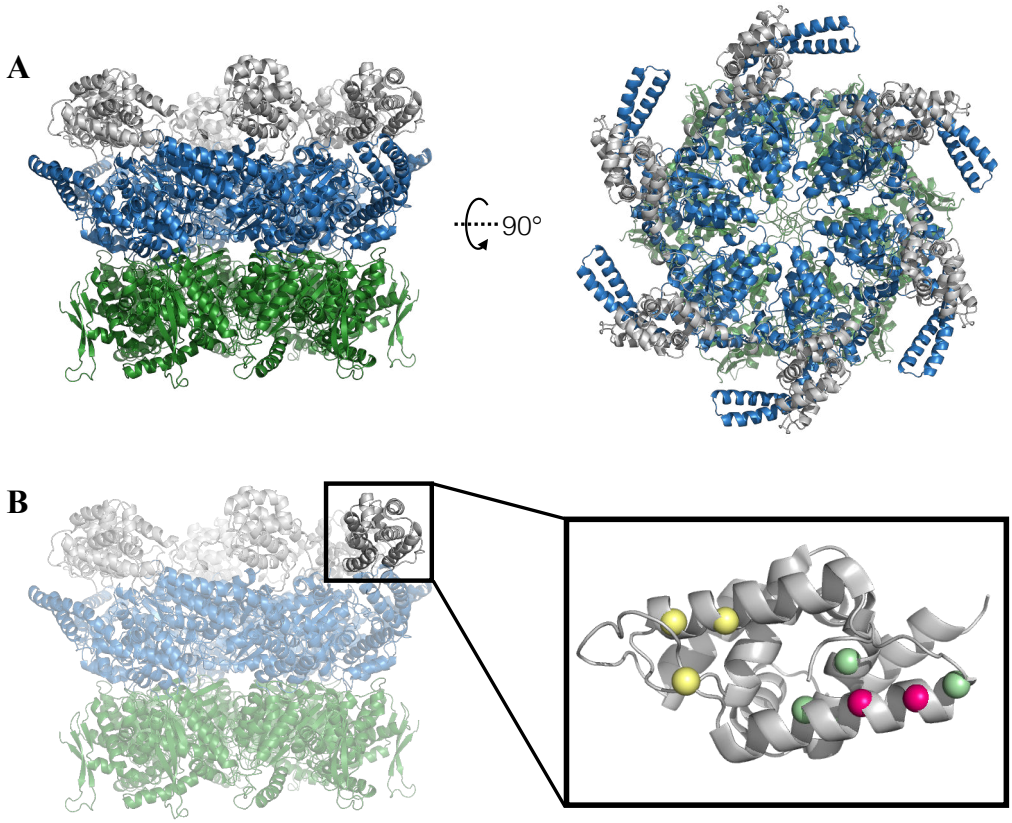


Figure 1.5: **A structural model of *Mycobacterium tuberculosis* ClpC1. A:** The ClpC1 unfoldase of *Mycobacterium tuberculosis* (model built based on the X-ray structure of *Bacillus subtilis* ClpC, PDB: 3PXI). In grey the N-terminal domain, in blue the D1 ATPase domain and in green the D2 ATPase domain. **B:** The NTD is the binding site of several antibiotics. Resistance mutations towards these antibiotics are shown as spheres: lassomycin (yellow),[37] ecumicin (pink)[36] and pyrazinamide (green).[153]

induce ATPase activity in *B. subtilis* ClpC without the help of the adaptor protein MecA, meaning that MecA is no longer necessary for hexamer formation. Moreover, arginine phosphate competes with MecA for ClpC binding. These effects can be produced by substrates phosphorylated on arginines, but also by the phosphorylated amino acid itself. How the binding of arginine-phosphorylated protein can induce ATPase activity and hexamer formation is currently unknown.[129]

It was proposed that ClpC recognizes substrate proteins with its NTDs that subsequently feed the substrate into the pore formed by the two ATPase domains D1 and D2. During the transfer through the pore the previously misfolded substrate gets unfolded and is handed over to the ClpP protease

for degradation. Threading of the substrate through the D1 and D2 pore is performed in a cyclic manner through an allosterically regulated "pulling" and "holding" of the substrate.[75]

ClpC1 from *M. tuberculosis* has attracted special attention in the last years as a possible drug target against tuberculosis. The target for all three recently discovered bactericidal compounds cyclomarin, lassomycin and ecumicin is ClpC1. While lassomycin and ecumicin both increase ATPase activity but not proteolysis,[36,37] cyclomarin enhances proteolysis.[108] All three antibiotics have been found to bind to the NTD of ClpC1, either by mutational studies or, in the case of cyclomarin, also by an X-ray structure and ITC measurements.[133] These antibiotics could be interesting lead compounds for antituberculosis drug development since they are selective for mycobacteria and kill growing, as well as non-replicating cells.[36,37,108] A crystal structure of *M. tuberculosis* ClpC NTD in complex with cyclomarin has been solved by Vasudevan et al.[133] Surprisingly, the structure of the NTD remains unchanged, except a slight increase of structuredness upon drug binding. Vasudevan et al. propose that binding of cyclomarin changes the dynamics of the linker connecting the ATPase domains and the NTD thereby regulating access to the ClpC pore.[133]

1.1.2 Examples of ATP-independent chaperones

This PhD project has been started to study the ATP-independent TIM910 chaperone and its substrate binding. TIM910 is an essential mitochondrial chaperone that transports membrane proteins in the intermembrane space (IMS). I dedicated an own section to TIM910, where I will discuss in more detail the structure, substrate binding and biological function of this chaperone. Before, I would like to discuss some other members of the ATP-independent chaperone family. I selected examples, not only because of their physiological relevance, but to introduce the large variety of binding concepts and interaction mechanisms present in these systems.

Small heat shock proteins

The diverse family of small heat shock protein (sHSP)s serves the cell by protecting it from misfolded proteins like aggregates and fibrils. How sHSPs perform this task remains poorly understood. A common feature of sHSPs is the organisation into large oligomeric complexes that undergo constant subunit exchange and that are thus quite heterogeneous. A mass spectrometry study by Stengel et al. showed that Hsp18.1 exists mainly as a dodecamer at room temperature, but forms smaller and higher order complexes at heat shock

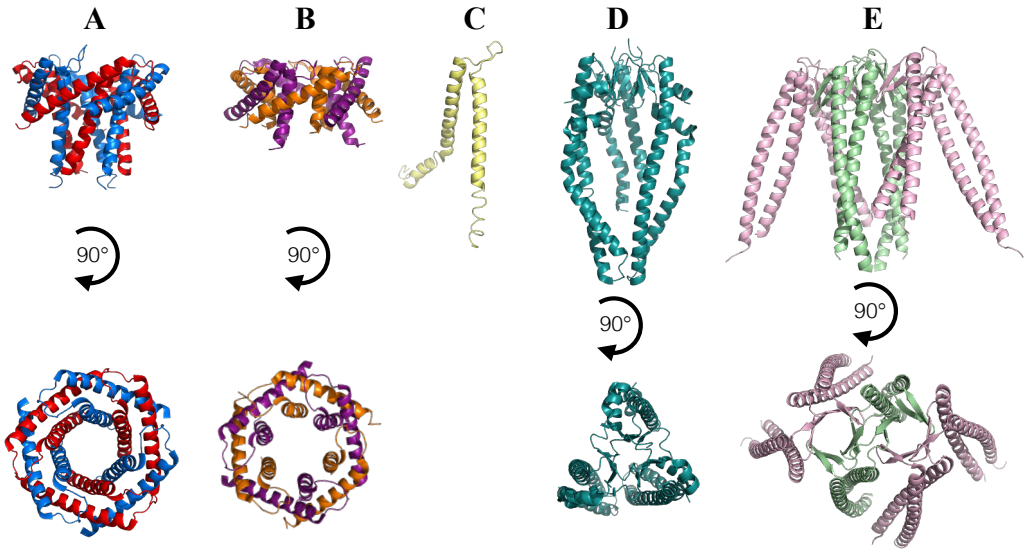


Figure 1.6: **Different examples of ATP-independent chaperones. A:** The essential intermembrane space chaperone TIM910, forms a hexamer of three alternating Tim9 (red) and Tim10 (blue) subunits (PDB: 3DXR). **B:** The non-essential intermembrane space chaperone TIM813, forms a hexamer of three alternating Tim8 (orange) and Tim13 (purple) subunits (PDB: 3CJH). **C:** A model of the mitochondrial inner membrane associated small Tim protein Tim12. In contrast to the other small Tim proteins it contains a prolonged C-terminus, that is predicted to form a helix in the folded state. **D:** The bacterial periplasmic Skp chaperones forms a homotrimer (PDB: 1SG2). **E:** The cytoplasmic chaperone Prefoldin forms a hexamer of α (green) and β (pink) subunits (PDB: 1FXK).

temperatures.[117] The addition of substrate then leads to a rapid formation of large polydisperse complexes with more than 300 different stoichiometries. The authors speculate that the initially observed dodecamer only serves as a reservoir to build higher order complexes upon heat shock conditions.[117]

Structural insight comes mainly from solid-state NMR in combination with SAXS and EM. The structure of the sHSP αB crystalin can be separated in three domains: a N-terminal domain (NTD), a central α-crystallin domain (ACD) and a C-terminal domain (CTD) with a common IXI motif.[53] An ACD dimer is the general building block of all oligomers, explaining the preference for even numbered oligomers. While the ACD stays structurally unchanged upon substrate binding, the NTD and CTD can adopt different conformations.[78] Furthermore, amorphous aggregates and fibrils do not seem to have the same binding site on αB crystalin. Thus, the inherent structural plasticity of sHSPs might be the underlying principle for recognizing a wide range of substrates.[78] A better understanding of the mechanism how sHPSs

work in the cell has been provided by Zwirowski et al. 2017. The authors could show that sHSPs and aggregates organize in a way favourable for later substrate recovery, but harmless for the cell. To this end, sHSPs assemble in a static core particle with the aggregate, that is surrounded by an outer shell of sHSPs that is in dynamic exchange with the environment. sHSPs of the outer dynamic shell can be displaced by Hsp70, making the aggregates accessible to the disaggregase Hsp100. Thus, under stress conditions sensitive proteins can be sequestered and stored away, but later on reactivated with the help of the Hsp70/Hsp100 system.[157]

Trigger factor

As mentioned before, newly synthesized proteins bind specialized chaperones that are associated with the ribosome exit tunnel to protect them and promote co-translational folding. In *E.coli* a protein called trigger factor is dedicated to this task. Trigger factor consists of a N-terminal ribosome binding domain (RBD), a Peptidyl-prolyl isomerase domain (PPI) domain and a C-terminal domain (CTD). The RBD and PPI domain are located on opposite sides of the molecule, while the CTD is in the middle and forms a cradle like structure with the other two domains. When trigger factor is associated with the ribosome, the hydrophobic interior of the cradle is presented towards the exit tunnel to receive and protect nascent polypeptide chains.[34]

One substrate protein can be bound by several trigger factor molecules, depending on the length of the substrate. The model substrate alkaline phosphatase is bound by three trigger factor molecules. Rearrangement of side chains in the binding site provide the needed placticity to interact with different sites of the substrate protein. In contrast to other holdase chaperones (for example Skp) the substrate is bound in one preferred conformation and does not dynamically interchange between binding sites.[103]

Skp

The bacterial ATP-independent chaperone Skp escorts outer membrane protein (OMP) precursors through the periplasm. Skp forms a trimer consisting of a β-barrel core with extending coiled α-helices. The cavity formed between the β-barrel core and the helices contains hydrophobic patches and can accomodate a folded protein of up to 25 kDa of size.[141] Skp only binds the hydrophobic transmembrane part of OMPs, external domains like the periplasmic domain of OmpA stay outside of the cavity in a folded conformation.[140]

Substrates inside of the cavity of Skp are generally unfolded and sample a multitude of conformations in the submillisecond timescale. In comparison to an unfolded protein in solution, the substrate inside of Skp is significantly compacted due to sterical hindrances. The interaction between the substrate and the chaperone is dominated by an avidity effect consisting of a multitude of weak interactions. The chaperone itself undergoes only minor changes upon substrate binding, among them a slight increase of structuredness in the helical "tentacles". In contrast to Prefoldin the substrate does not interact with the tip of the "tentacles", but is buried deeper inside of the cavity.[18] To accomodate substrates of increasing sizes Skp could either expand its cavity or use several Skp trimers for substrate binding. Burmann et al. showed that the "tentacles" contain a hinge region that experiences increased dynamics when the substrate is bound.[18] Indeed, Schiffrin et al. could show, using molecular dynamics, that both mechanisms are possible and that Skp can do both, expand its cavity and bind substrates using several Skp complexes.[106]

To insert the OMP precursors as β-barrels into the bacterial outer membrane, Skp works in concert with another ATP-independent periplasmic chaperone SurA. In contrast to Skp SurA does not contain a cavity for substrate binding but a crevice formed by its N-terminal domain that could serve as a binding site for extended polypeptides.[13] While Skp prevents misfolding by keeping the substrate in an unfolded state, SurA increases the folding efficiency of the substrate and its insertion as β-hairpins into the membrane. Both chaperones bind their substrates in an unfolded dynamic conformational ensemble. Keeping the substrate in this dynamic state and allowing it to sample different conformations might be the key to enable insertion into the membrane.[128]

cpSRP43

cpSRP43 is a chaperone that protects light harvesting chlorophyll a/b binding proteins (LHCP)s and guides them to their place of insertion in the membrane. cpSRP43 can exist in three states: an open state, a closed state and a semi-open state. NMR spectra of cpSRP43 show two conformations for 12 residues. Upon addition of the co-factor cpSRP54 the population is shifted to only one conformation, the closed state, which renders the chaperone active and able to bind its substrate. Subsequent binding of the translocase Alb3 promotes the semi-open state which leads to a release of the substrate.[72]

This system nicely shows how conformational flexibility and interaction with co-factors can regulate substrate binding and release in ATP-independent chaperones.

Spy

The bacterial chaperone Spy binds its substrate Im7 in a hydrophobic groove. The first encounter between Spy and Im7 is driven by electroctatic forces, only the subsequent stability of the complex is determined by hydrophobic interactions.[65] Spy binds preferably unfolded clients in a compacted conformational ensemble, partially folded clients are only recognized by their unfolded parts and Spy has a lower affinity towards these substrates. Upon binding to Spy, the client shows increased dynamics at the binding interface.[45] Koldewey et al. proposed that folding and release of the client are "self controlled". While the substrate is folded inside of the chaperone, the hydrophobic surface decreases, making the complex less stable and leading to release of the folded substrate.[65]

In this system chaperone binding and release are only regulated by the different affinity of the chaperone towards unfolded (hydrophobic patches exposed) and folded (hydrophobic patches shielded) substrates, no ATP or other co-factors are needed for this folding cycle.

1.2 Mitochondrial protein import and the TIM910 chaperone

The mitochondrial genome has been highly reduced and encodes only for about 1% of all mitochondrial proteins. The remaining proteins are translated in the cytosol and need to be imported into mitochondria. To this end, mitochondria developped an intricate import machinery, which will be introduced in this chapter.

1.2.1 Different import machineries

To date five different import pathways into mitochondria are known: 1) the presequence pathway, 2) the mitochondrial carrier pathway, 3) the Mia40 pathway, 4) the β-barrel pathway and 5) the MIM pathway. Pathways 2 and 4 depend on the TIM910 chaperone and will be described in more detail later in this chapter.[145]

The general entry gate for mitochondrial protein import is the translocase of the outer membrane (TOM) complex. Pathways 1), 2), 3) and 4) use this complex for protein translocation through the outer membrane and even a subset of proteins of pathway 5) need a subunit of the TOM complex for correct assembly. The TOM complex is constituted of the channel-forming protein Tom40, the presequence receptor Tom20, the central receptor Tom22,

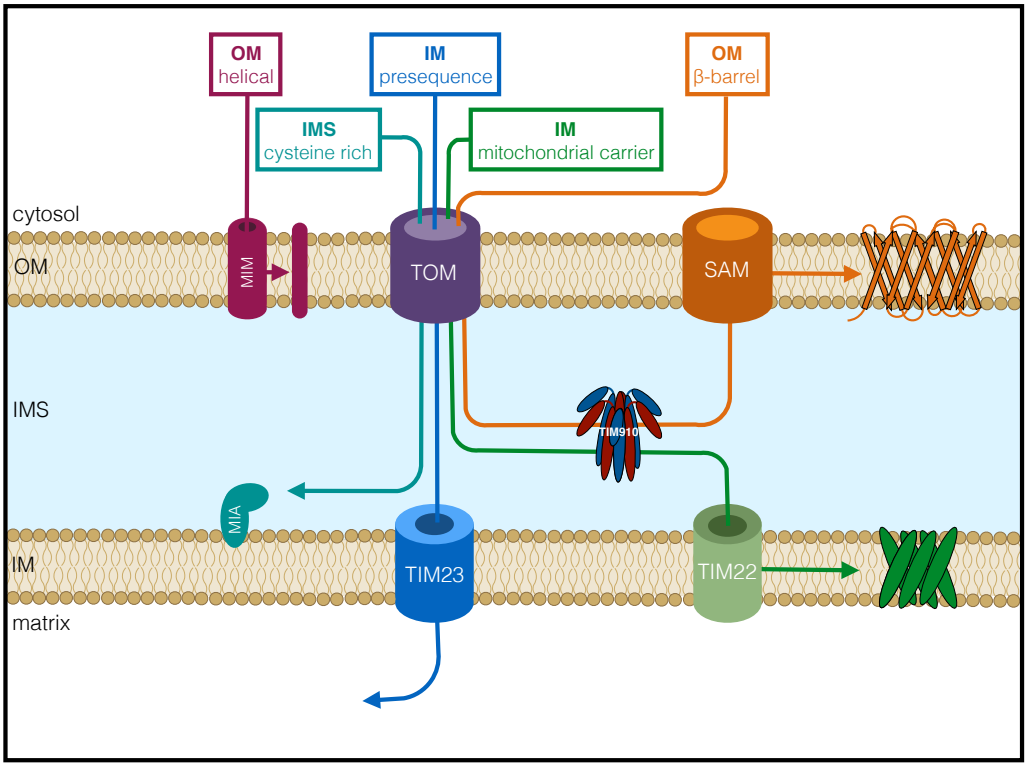


Figure 1.7: **The five import pathways of proteins into mitochondria.** Four out of the five import pathways use the TOM complex to enter mitochondria. Only outer membrane α-helical proteins can be directly inserted by the MIM complex (pink). The majority of imported proteins carries a N-terminal targeting sequence and is imported by the TOM complex and the inner membrane TIM23 complex (blue). Cysteine-rich intermembrane space proteins are recognized by the MIA40 complex (teal). Mitochondrial carriers of the inner membrane and β-barrel proteins of the outer membrane depend on the IMS chaperone TIM910 and are membrane inserted by either the TIM22 complex (inner membrane, green) or the SAM complex (outer membrane, orange). Import scheme adapted from Wiedemann et al.2017.[145]

the mitochondrial carriers (MC) receptor Tom70 and three small Tom proteins necessary for stability and assembly of the complex.[84,112]

1) The presequence pathway

60% of all proteins imported into mitochondria carry a N-terminal presequence and are thus destined for insertion by the TIM23 complex into the mitochondrial inner membrane or the matrix.[137] Presequences are typically 15-50 amino acids long and form an amphipathic α-helix whereby one side is positively charged and the other hydrophobic.[99] The presequence is recognized by the Tom20 receptor of the TOM complex and then handed over to the central Tom22 receptor for subsequent insertion into the Tom40 pore.[149] Presequence proteins do not require intermembrane space chaperones for the transfer across the IMS, but are directly handed over to the TIM23 complex from TOM. The TIM23 complex consists of the presequence receptor Tim50, the regulatory protein Tim21, which links the TIM23 complex to the respiratory chain, and the channnel forming protein Tim23.[130] An additional protein, Tim17, can be found in the same complex. Although structurally similar to Tim23, the function of Tim17 has not been determined yet.[26]

The entering presequence is recogniced by the Tim50 receptor, which also interacts with the unfolded N-terminal domain of Tim23 to activate and open the channel.[96] Translocation of the presequence protein requires a membrane potential and the ATP-dependent protein import motor (PAM) located on the matrix side. PAM is a mitochondrial homologue of Hsp70 and like Hsp70 interacts with a series of co-chaperones to regulate its activity.[58] Proteins destined for the inner mitochondrial membrane can be either inserted directly by the TIM23 complex by lateral release with the help of the "gatekeeper" protein Mgr2 or translocated into the matrix and then inserted via the OXA pathway. Insertion by the OXA machinery is also termed "conservative sorting" because it resembles the sorting pathway used by bacteria, which is still conserved for proteins encoded by mitochondrial DNA.[52]

2) The mitochondrial carrier pathway

The carrier pathway will be discussed in more detail in the next section. Briefly, the MC precursor is delivered to the TOM complex by the cytosolic chaperones Hsp70 and Hsp90 and recognized by the Tom70 receptor. After translocation through Tom40, the MC precursor is bound to the IMS chaperone TIM910 and transported to the TIM22 complex for insertion into the inner membrane.[64]

3) The Mia40 pathway

Mia40 is specialized in the import of cysteine containing proteins. Several IMS proteins contain either CX_3C or CX_9C motifs, which form intramolecular disulfide bonds in the mature protein. Mia40 works together with the sulfhydryl oxidase Erv1, by transmitting electrons from the substrate via Mia40 to Erv1 and then further to cytochrome c or molecular oxygen.[81] The full pathway is explained in the next section with Tim9 and Tim10 as example proteins.

3.1) Tim9 and Tim10 as an example for Mia40 import

All small Tim proteins (Tim8, Tim9, Tim10, Tim12 and Tim13) have cysteine motifs and therefore are imported via the Mia40 import machinery. The small Tims have to be kept in a reduced state in the cytosol for import through the TOM complex. Oxidized subunits are not import competent. The previously widely accepted role of zinc binding to keep the small Tims in a reduced state in the cytosol, seems obsolete now that additional experiments showed that zinc inhibits the import of small Tims. A more convincing hypothesis suggests the cytosolic thioredoxin system as a major factor for keeping the small Tims in a reduced state.[30] So far no chaperone has been identified that would transport the small Tims to the TOM complex in mitochondria.

Once the small Tims are translocated through the Tom40 pore they are bound to Mia40 via their first (N-terminal) cysteine residue. Interestingly, also yeast Tim12 binds Mia40 with its first cysteine residue, although contrary to the other small Tims, this cysteine residue is not part of the CX_3C motif.[83] In contrast to other mitochondrial proteins, small Tims only need a reduced TOM complex without receptors for import.[40] Milenkovic et al. identified the minimal sequence necessary to bind to Mia40.[82] Nine amino acids around the first cysteine residue are required for binding, whereby two of these residues must be hydrophobic and one of them a leucine, all other residues of the binding signal can be substituted by alanines.[83]

4) The β-barrel pathway

β-barrel proteins are, like mitochondrial carriers, substrates of the TIM910 chaperone and will thus be discussed in one of the following sections. Briefly, β-barrel proteins are targeted to the TOM complex by a β-hairpin signal in the last C-terminal β-strand.[68] After translocation through Tom40 the β-barrel precursor is transferred with the help of TIM910 and TIM813 to the

sorting and assembly machinery (SAM) complex for insertion into the outer mitochondrial membrane.[41,147]

5) The MIM pathway

This last pathway serves the insertion of α-helical proteins into the mitochondrial outer membrane and is somewhat less studied than the other import pathways. Unlike the β-barrel proteins, α-helical proteins do not have to be imported into the IMS to be inserted into the outer membrane, but can be directly inserted from the cystosol. The machinery responsible for the insertion is called MIM complex and consists of several copies of Mim1 and one copy of Mim2.[27,94] Three different types of α-helical proteins can be discriminated: signal-anchored proteins with a N-terminal transmembrane helix, tail-anchored proteins with a C-terminal transmembrane helix and polytopic α-helical outer membrane proteins. The MIM complex is essential for the insertion of signal-anchored proteins.[10] Polytopic proteins need the Tom70 receptor of the TOM complex and the MIM machinery, although the MIM machinery might be dispensable.[88] Tail-anchored proteins might insert without any insertion machinery or the import system has not been identified yet.[60]

1.2.2 TIM910 dependent import pathways

The IMS chaperone TIM910 is responsible for the transport of MCs to the TIM22 complex for insertion into the inner membrane and for the transport of β-barrel proteins to the SAM complex for insertion into the outer membrane. In contrast to the presequence pathway TIM910 does not require a presequence for substrate recognition, but binds to internal targeting sequences in the precursor protein.[152] The presumed role of TIM910 is to protect these membrane protein precursors during their transport in the hydrophilic environment and keep them from aggregation or misfolding.

A first model of how mitochondrial carrier precursors are imported into mitochondria has been proposed by Köhler et al.[64] This import scheme is divided into five stages: In the first stage the mitochondrial carrier precursor binds to the TOM receptor in the cytoplasm, in the second stage the MC precursor crosses the Tom40 pore, which depends on the presence of Tim9 and Tim10, in the third stage the MC precursor is fully imported into the IMS, depending on Tim12, in the fourth stage the mitochondrial carrier precursor is inserted into the membrane by the TIM22 complex, whereby an electric potential is necessary for insertion, and in a final fifth stage the MC reaches its final mature state.[64]

This model has been expanded due to novel findings and depicts now an even broader and more versatile function of the TIM910 chaperone in this complex import network.

Membrane protein precursors are transported through the cytosol by Hsp70 and Hsp90 chaperones and then imported via the pore of the TOM complex.[112,151] While presequence containing proteins are guided by acidic patches in the channel, membrane proteins are guided by patches of hydrophobic residues inside of Tom40.[112] MCs pass the Tom40 channel in a loop conformation, with the two termini protruding into the cytosol.[146] Membrane protein precursors may even already be in a partially folded state during translocation, as has been shown for Tom40.[97] Cross-linking experiments showed that an approximately 60 residue long N-terminal segment of Tom40 can protrude from the cytosol into the mitochondrial IMS through the Tom40 pore and recrute the TIM910 chaperone.[112]

Little is known about the implication of TIM910 in the import of outer membrane β-barrel proteins. Import and assembly of Tom40[147] as well as of Sam50[41] depend on TIM910 and TIM813. While MCs are only imported with the help of TIM910, outer membrane proteins need both IMS chaperone complexes (TIM910 and TIM813) for efficient import.[41,64,147]

TIM910 recognizes hydrophobic transmembrane segments together with hydrophilic flanking regions. Hydrophobicity alone is not sufficient for TIM910 binding.[132]

Outer membrane proteins carry a β-signal responsible for sorting of β-barrel proteins to the SAM complex by binding to the Sam35 subunit. The signal is situated in the C-terminal part of the last β-strand and consists of a large polar residue (K or N), a conserved glycine and two large hydrophobic residues. Upon binding of the β-signal to Sam35, the SAM complex rearranges to be able to accomodate several transmembrane β-strands.[68] Subsequent release of the β-barrel protein into the membrane requires a conserved glycine in the β-signal[68] and the polypeptide-transport-associated (POTRA) domain of Sam50.[119] Recently, Jores et al.[55] discovered the minimal sequence required for mitochondrial targeting. Thus, the β-signal alone found by Kutik et al.[68] is not sufficient. The complete targeting signal must include the last two C-terminal β-strands forming a β-hairpin structure. Structuredness of this signal seems to be a prerequisite, since only a cyclic version of the targeting sequence, which is more prone to form a β-hairpin than the linear version of this peptide, was able to bind to the Tom20 receptor *in vitro* and hamper mitochondrial import *in vivo*. Interestingly, the β-hairpin signal interacts with the same binding site as presequences on the Tom20 receptor.[55]

SAM can also form a supercomplex with TOM by interacting with the

cytosolic domain of Tom22, the central receptor of TOM. Formation of this supercomplex and transfer of the membrane protein from TOM to SAM depends also on the small Tim proteins.[95]

TIM910 can also play an important role in complex assembly. During the import of Tim54, a component of the TIM22 complex, TIM910 is not needed for membrane insertion. Formation of a supercomplex between TOM and TIM22 allows crossing of the IMS by the Tim54 precursor. However, TIM910 is essential for the final assembly of Tim54 in the TIM22 complex.[139]

1.2.3 TIM910 structure

Tim9 and Tim10 are members of the small Tim family, comprising the essential proteins Tim9, Tim10 and Tim12 and the non-essential proteins Tim8 and Tim13. The structure of the small Tims is highly conserved, they all share a common twin CX_3C motif that stabilizes a helix-loop-helix fold.[12,143] All five small Tim members are exclusively found in the mitochondrial IMS. While the fold of the small Tims is strictly conserved, the number of small Tims is variable.[39] Some eukaryotes have been found with only a single small Tim, leading to the conclusion that all small Tims were derived from a single progenitor via gene duplication and modification. No small Tim ancestor has been found in prokaryotes, thus the small Tim progenitor must have been put in place by the eukaryotic host cell to facilitate membrane protein import into mitochondria. According to genetic analysis the most recent duplication event resulted in the generation of Tim12, which due to its localisation at the membrane offers properties distinct from the other small Tims. If Tim12 was derived from Tim10 or from Tim9 is not clear. While yeast Tim10 is more closely related in sequence to yeast Tim12, human Tim12 (Tim10b) is more closely related to human Tim9.[39]

The main purpose of the small Tims is to transport hydrophobic membrane protein precursors across the aqueous IMS. Although very similar in their structure, each of the small Tim complexes is specialized in performing different tasks.

Tim9 and Tim10 can be found in three different complexes within the cell: the soluble TIM910 complex, the membrane/TIM22 attached TIM91012 complex and a soluble form of the TIM91012 complex. The most abundant form is the soluble TIM910, it is about three to four times more abundant than the membrane attached TIM91012 complex.[1] The soluble TIM91012 complex is presumably only a transport form, where Tim12 is shuttled from Mia40 to the TIM22 complex by TIM910.[38] While it has been shown by several groups that Tim9 and Tim10 can spontaneously self assemble into a hexameric TIM910 complex with a 3:3 stoichiometry,[77,136,143] the composi-

tion of the TIM91012 complex is less clear. A phosphorimaging experiment of Adam et al.[1] suggested a stoichiometry of the TIM91012 complex of 3:2:1. Nevertheless, Gebert et al.[38] suggested a different stoichiometry of TIM91012 of 3:1:2.

The rather dynamic nature of the TIM910 complex with its supposedly flexible N- and C-termini, which can be assumed because of high B-factors in this region, hampered crystallisation of the complex for a long time. Thus, the first structural insights have been obtained by SAXS, NMR and circular dichroism (CD).

CD spectra of Tim9 and Tim10 with oxidized or reduced CX_3C motif revealed that reduced Tim9 and Tim10 are rather unfolded, while oxidized proteins show a more helical behaviour. Only upon hexamer formation the two subunits become structured. This was further corroborated by NMR analysis of the two subunits. Tim9 and Tim10 seemed to behave as molten globules in their oxidized form, whereby Tim9 is supposedly more structured than Tim10.[77] The quite heterogeneous peak intensities and the number of visible glycine residues in NMR spectra led the authors to the conclusion that the N- and C-termini are unfolded while the middle part forms a molten globule. Analysis of 1D NMR spectra of the TIM910 hexamer showed that Tim9 is more structured in the complex, while Tim10 remains a molten globule. A SAXS structure of TIM910 showed, that the complex had no cavity and that one subunit, presumably Tim9, was more elongated.[77]

The first crystal structure of a TIM910 (human) complex has been solved in 2006 by Webb et al.[143] It revealed a hexamer with alternating Tim9 and Tim10 subunits comprising an α-propeller fold with helical tentacles. Each subunit contains two disulfide bridges in the CX_3C motif forming an inner and an outer disulfide bridge. The fold of both subunits is a helix-loop-helix motif with the last 10-15 residues of the N- and C-terminus presumably unfolded. The helices form an inner (N-terminus) and outer (C-terminus) ring around an approximately 15 Å cavity. The cavity is lined by hydrophilic sidechains and the helices of the two subunits are held together by hydrophobic interactions and two salt bridges. No apparent binding site can be detected in this complex, leading to the question if a stuctural reorganization is necessary for substrate binding.[143] Three years later the structure of yeast TIM910 has been solved, proving to be almost identical to the human TIM910 structure.[6]

1.2.4 TIM910 substrates

As discussed in the previous sections, three main types of proteins are imported via the TOM import complex, (i) proteins carrying a presequence

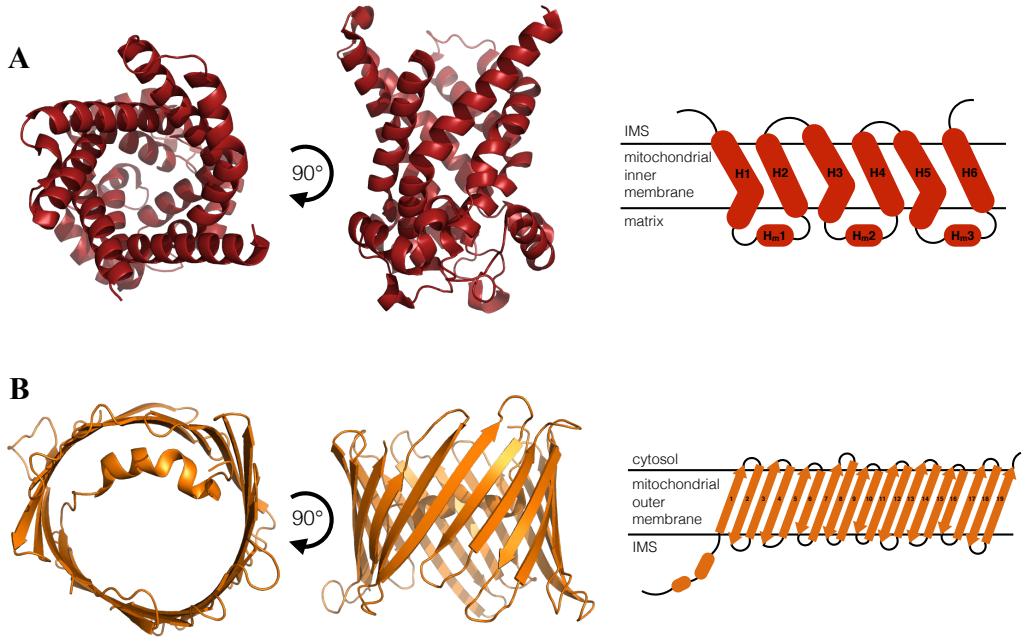


Figure 1.8: **Two classes of TIM910 substrates.** TIM910 transports two different types of substrates. **A:** Mitochondrial carriers of the inner membrane share a common fold with six transmembrane helices, whereby every second helix is kinked due to a proline residue, and three short matrix helices (PDB: 1OKC). **B:** Substrates of the outer membrane fold into β-barrels with 19 transmembrane β-strands and an N-terminal helix (PDB: 3EMN).

destined for the mitochondrial matrix,[137] (ii) proteins carrying internal targeting sequences destined for the inner mitochondrial membrane and (iii) β-signal carrying proteins destined for the outer mitochondrial membrane. While presequence-carrying proteins are directly transferred to the TIM23 import machinery in the inner membrane, proteins with internal targeting sequences or β-signal are strictly dependent on the small Tims.[152] Inner membrane substrates, including primarily the large group of mitochondrial carriers, all form α-helical structures in their membrane-inserted form,[91] while outer membrane substrates are β-barrel proteins in their functional membrane inserted form.[9,48] An exception from this rule is the cardiolipin acyltransferase tafazzin, which only contains one transmembrane element and is first integrated into the outer membrane with the help of TIM910 and then redistributed to the inner membrane.[20,46]

Inner-membrane substrates

The main substrates of the TIM910 chaperone are MCs, a large family of 53 members in humans and 35 in yeast, dedicated to the transport of metabolites like amino acids, nucleotides or keto acids across the inner mitochondrial membrane. Selective transport of metabolites in and out of mitochondria is a crucial process for every cell to ensure supply with energy and components for amino acid and nucleotide synthesis. The importance of this metabolite exchange is underlined by the existance of 14 different genetic disorders all related to mutations in MCs.[87]

All MCs show common sequence motifs, which will be explained with regard to the X-ray structure of the ADP/ATP carrier (AAC). In addition to the basic MC motifs, MCs may contain specialized domains for example for ligand binding.[69] The first structure of a MC, the bovine AAC, was solved in 2003. It comprises six transmembrane helices, whereby every second helix is kinked by a proline residue in a conserved motif. The transmembrane helices are connected by loops, which in the matrix side contain three extra helices. The gating between two putative states might be controlled by two salt bridge networks on the top and bottom of the protein.[91] So far only the inhibitor bound cytosolic state could be crystallized. A major challenge for structural studies of these proteins seems to be their inherent flexibility,[16] thus only three X-ray structures of AAC and no other mitochondrial carrier have been solved since then.[91,102]

Outer-membrane substrates

Mitochondrial β-barrel proteins are the most common proteins in the outer membrane and are derived from bacterial β-barrel proteins. Three types of mitochondrial β-barrel proteins exist, Tom40, the channel forming component of the TOM complex, Sam50, the pore of the SAM insertion machinery and VDAC, the most abundant protein of the mitochondrial outer membrane and channel for small molecules and ions.[28]

The cryo-EM structure of the *Neurospora crassa* TOM core complex has only been solved recently.[8] In this complex Tom40 forms a dimer consisting of two 19 β-stranded β-barrels. The central receptor Tom22 connects the two Tom40 pores, while the α-helical small Tom proteins Tom5, Tom6 and Tom7 surround them. The N- and C-termini of Tom40 form α-helices, whereby part of the N-terminus is inserted into the Tom40 pore.[8] TOM complexes can contain a variable number of Tom40 proteins, while the *Neurospora crassa* TOM complex contains two Tom40 molecules, the *Saccharomyces cerevisiae* complex formed assemblies with three Tom40 pores, as has been shown by

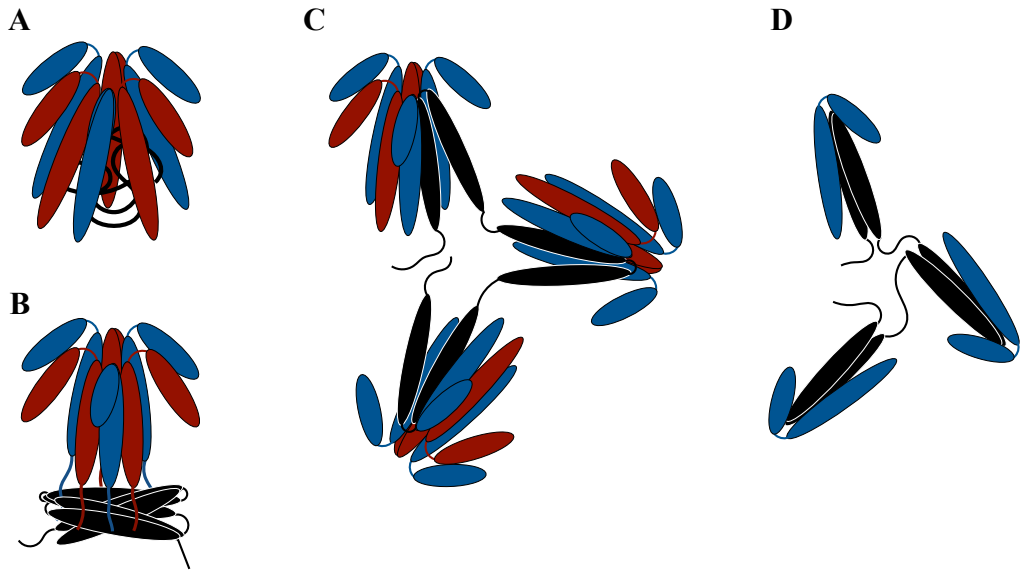


Figure 1.9: **Possible substrate binding models of TIM910.** Based on the structure of TIM910 and similarities to other holdase chaperones several binding models have been proposed in the literature. **A:** TIM910 could bind its client proteins in its cavity in an unfolded dynamic state similar to the bacterial Skp chaperone.[18] **B:** The unfolded N-terminal "tentacles" of TIM910 could serve as lipid surrogates and bind to the transmembrane regions of its substrates. A similar binding mechanism has been suggested for prefoldin.[113] Based on the fact that the cavity of TIM910 is hydrophilic Webb et al.[143] suggested two binding scenarios. Either one of the Tim subunts could be exchanged for two helices of the substrate protein (**C**) or the substrate transmembrane helices could dimerize with the helices of TIM910 or one of the TIM910 subunits (**D**). If dissociation of the TIM910 hexamer is necessary for binding or if only one of the two subunits of TIM910 is involved in the interaction is not known.

cross-linking experiments.[112]

X-ray and NMR structures of human VDAC-1 reveal a 19 β-stranded β-barrel with a N-terminal α-helix inserted in the pore formed by the barrel. The β-barrel is not symmetrically organized and the first four β-strands seem to be less stable than the rest of the protein.[9,48] The same conserved fold is shared by Tom40.[8]

1.2.5 Substrate binding models

When Webb et al. solved the first crystal structure of a TIM910 complex in 2006 they suggested three possible binding scenarios based on the fact that the complex has to disassemble because the TIM910 cavity is hydrophilic

and too small to host a substrate protein. Furthermore, no apparent binding site could be found in the TIM910 structure. In the first model helices of the substrate protein and the TIM910 complex dimerize during transport. In the second model Tim9 subunits are displaced by substrate helices. In the third model the "tentacles" (N-termini of Tim9 and Tim10) serve as lipid surrogates to transport the substrate.[143]

Peptide binding scans showed that TIM910 binds preferentially to the hydrophobic transmembrane segments of its substrates. In AAC2 it interacts most strongly with predicted helices 3 and 4 followed by helices 5 and 2.[22] Similarly, in Tim23 it binds only the hydrophobic C-terminal transmembrane part, but not the hydrophylic N-terminus.[23] Whether TIM910 can also function as a chaperone for Tim23, a classical TIM813 substrate, is up until now controversial. Cross-linking experiments by Paschen et al.[90] and Davis et al.[23] showed that Tim23 can be specifically cross-linked to Tim9 and Tim10, while peptide scans by Curran et al.[22] detected no binding of TIM910 to Tim23. A small molecule inhibitor of TIM910, MitoBloCK-1, found in a chemical-genetic screen, binds specifically to TIM910 and only acts on typical TIM910 substrates like mitochondrial carriers, Tom40 or Tim22 but not on Tim23.[44] Whether the described cross-links are caused by the interaction of free TIM910 or membrane/TIM22 associated TIM91012 is not clear from the published data.

Several studies investigated the different functions of N- and C-terminus of the small Tim proteins. All experiments converge on the idea that the C-terminus is important for complex stability and the N-terminus for substrate binding. This has been shown for Tim9, Tim10[6,135] and Tim12.[74] It should be mentioned though, that the effect of a truncation of the N- or C-terminus can only be observed when a rather large part of the terminus is removed (about one third of the full-length protein). The role of the individual subunits in substrate binding is not entirely understood. While an assembled hexamer is necessary for its function *in vivo*, it seems that Tim10 is the only subunit interacting with the substrate in a peptide scan. Moreover, 1D NMR spectra of Tim9 and Tim10 with an AAC peptide only provided chemical shift perturbations of Tim10.[135] The authors suggest therefore that Tim9 only plays a structural role in the TIM910 complex while Tim10 binds the substrate.

Owing to the common origin of bacteria and mitochondria, similarities and overlapping functions between mitochondrial IMS chaperones and bacterial periplasmic chaperones are expected. Although TIM910 and periplasmic chaperones show a similar preference for binding sequences with a high content of aromatic amino acids with a slight positive charge, the bacterial chaperone SurA is not able to substitute the function of TIM910. However,

this inability could be due to a defect in transfer of the substrate to the TIM22 complex and not due to substrate recognition and binding.[2]

1.2.6 TIM813

TIM813 is the second soluble chaperone complex present in the mitochondrial intermembrane space. In contrast to TIM910 it specifically transports only a small subset of membrane proteins. The best studied substrate of TIM813 is Tim23 the main component of the TIM23 import machinery. While TIM910 is an essential protein, deletion of TIM813 is not lethal but leads to a 75% decrease in Tim23 import, meaning that part of the function of TIM813 can be taken over by another machinery.[98] Indeed Tim10 has been shown to cross-link to Tim23, but Tim8 and Tim13 do not cross-link to the canonical TIM910 substrate AAC.[23]

Mutations or deletions in the human homologue of Tim8 DDP1 lead to the deafness-dystonia syndrome, a genetic progressive neurodegenerative disorder.[64] Why mutations in DDP1 only affect neuronal cells and leave yeast cells for example more or less unharmed is currently unknown. A plausible cause might be that neuronal cells are more susceptive to reduced levels of Tim23, which hamper import of components for oxidative phosphorylation and energy generation.[98]

The TIM813 structure is similar to TIM910, with an identical helix-loop-helix fold held together by a twin CX_3C motif. A major difference is the length of the disorered N- and C-termini. While in TIM910 only the last 10-15 N-terminal residues are assumed to be disordered, up to 45 residues in TIM813 seem to be flexible. The length of the tentacles could be an important factor in substrate specificity. Furthermore, it was proposed that substrate binding in TIM910 and TIM813 is markedly different. While TIM910 is believed to dissociate and reassemble upon substrate binding, TIM813 might stay intact and bind its substrates solely with its disordered tentacles and the help of six hydrophobic pockets on the protein surface.[12] TIM813 and TIM910 also seem to target different binding sites on Tim23. While TIM813 preferentially binds the hydrophilic N-terminus of Tim23, which is presumably unfolded in the final inserted form, TIM910 binds only to the hydrophobic C-terminal part.[23,90]

Davis et al. investigated binding specificity of the different small Tims towards Tim23 by using cross-linking experiments and found that each Tim only binds to a specific place on Tim23. They could even identify a stretch of five residues where each residue bound preferentially to another small Tim: residue 153 to Tim9, 155 to Tim13, 156 to Tim8 and 158 to Tim10. Moreover, the position of the cross-links suggests that Tim23 must be in a folded (α-

helical) state when bound to the small Tims.[23] Additionally, Paschen et al. found that TIM813 can only bind to a Tim23 precursor but not to the membrane-inserted form.[90]

1.2.7 Tim12

Despite the essential role of Tim12 in mitochondrial membrane protein import, little is known about this small Tim protein that stands out due to its exclusive localization at the TIM22 import machinery at the inner membrane.

In vivo Tim12 can be found in three complexes: a 180 kDa assembly with the Mia40 complex for disulfide bond formation, a 70 kDa intermediate soluble complex with Tim9 and Tim10 and a membrane bound 300 kDa complex with Tim22, Tim54, Tim18, Tim9 and Tim10. The soluble form of Tim12 with Tim9 and Tim10 is supposed to be a pure transport form from the Mia40 complex to its final destination in the TIM22 complex, since Tim12 without Tim10 is not capable to associate with TIM22. The stoichiometry of the complex is unknown. While Gebert et al. found a 3:1:2 ratio for TIM91012,[38] Adam et al determined it to be 3:2:1.[1] However, contrary to those previous experiments, Lionaki et al. purified Tim12 from yeast mitochondria as homooligomeric complexes.[24]

The presumed function of Tim12 is to tether substrate loaded TIM910 complexes to the membrane for insertion using the TIM22 complex. Analogous to the other small Tims, Tim12 contains a twin CX_3C motif and likely adopts the same helix-loop-helix fold. A first biochemical study of yeast Tim12 showed that depending on the purification protocol Tim12 can either form disulfide-bonded dimers, with two additional cysteines present in yeast Tim12, or monomers. The authors propose that only dimeric Tim12 is active and able to bind Tim9, while monomeric Tim12 can neither bind Tim9 nor Tim10.[24] This hypothesis is rather unlikely since these two additional cysteines are not conserved in other species (see figure A.6). Given the high level of conservation of the small Tims, it is not expected that their mechanism of action differs significantly between species.

Like the other small Tims, the N- and C-termini of Tim12 seem to have distinct functions. Similar to Tim9 and Tim10, the N-terminus of Tim12 is important for substrate binding while the C-terminus is important for complex formation with Tim9 and Tim10. A special feature of the C-terminus is its ability to bind lipids. Tim12 is able to bind mitoplasts as well as liposomes, especially when they are enriched in cardiolipin, a lipid typical for the mitochondrial inner membrane. Upon deletion of the C-terminus of Tim12 lipid binding is completely inhibited. Interestingly, these specific effects of the N- and C-terminus are only visible when a rather large portion of the

termini is deleted (30-40 residues), deletion of only 20 residues shows no effect.[74]

If localization to the TIM22 complex is only mediated by lipid binding, or also by interaction with the TIM22 complex is not entirely understood.

1.3 Dynamics in chaperone/substrate complexes

If we assume that dynamics is everything but a static state, we encountered it during this introduction in very different forms and timescales. We saw that sHSPs exist in oligomeric assemblies that are in constant exchange of subunits and that this exchange is crucial for their function. We saw that ATP-dependent chaperones often undergo large domain movements and that these movements are the basis for their function. We saw that some chaperones can fluctuate between different conformations and that the population of these conformations can be influenced by other proteins that can turn these chaperones "on" or "off". We saw that chaperones can adapt their binding sites to their substrates to make them efficient binders of a large amount of substrates. We saw that the interactions between chaperones and substrates are often dynamic and transient to allow the substrates to sample a large conformational space to promote folding. However, the most dynamic part in these complexes, is likely the substrate, which can be folded, unfolded or kept in a heterogeneous dynamic folding competent state.

All these processes pose a major challenge for structural biology, which is rather specialized on the study of static states or at least a discrete number of states. Nevertheless, many new findings about chaperone/substrate complexes have been done in the last years and several techniques exist that are well suited for the study of these elusive complexes. While X-ray and cryo EM are still the methods of choice to determine the overall structure of a complex (or parts of the complex), especially if the complex can be trapped in distinct stable conformations, NMR has emerged as a technique that can give valuable information on the dynamics of binding sites or the bound substrate. NMR is ideally suited to study different conformations, local dynamics, folding or exchange processes. And with the development of specific methyl labelling techniques in combination with deuteration, even larger systems of up to several hundred kDa are accessible for NMR.

Chapter 2

Aim of this project

Aim of this project

The aim of this research project was to elucidate the binding mechanism of the TIM910 chaperone and its membrane protein substrates. As described above, structural information and biochemical data on these complexes are scarce, owing to the difficulty to work with hydrophobic aggregation-prone substrates that most likely interact in a highly dynamic manner. Up to date no information about any full-length substrate bound to TIM910 is available, binding models were derived solely form peptide spot assays using short substrate sequences and assumptions derived from the apo X-ray structure of TIM910. Having stated this, already the first step to any biochemical or structural study, the formation of the complex, seemed to pose a considerable challenge.

Nevertheless, we ventured to study the behaviour of this elusive system. Our goals were: 1) to describe the dynamics of the apo hexameric chaperone and its isolated subunits by NMR 2) to form a stable complex between chaperone and substrate 3) to derive residue-resolved information about structure and dynamics of the chaperone/substrate complex by NMR, ideally on both, the chaperone and the bound substrate 4) to determine the substrate binding site on TIM910 5) to obtain information about the stoichiometry and the overall structure of the complex using lower resolution techniques like SAXS or AUC 6) to compare outer and inner membrane protein substrates of TIM910, that differ significantly in their membrane inserted form 7) to elucidate the mechanism of substrate handover at the inner mitochondrial membrane by studying the membrane associated Tim12 subunit 8) to confirm the biological relevance of our data by *in vivo* experiments.

To address all the above mentioned points we relied primarily on NMR, as a technique ideally suited to study dynamic systems. Nevertheless, we combined NMR with a series of different biochemical and biophysical techniques to validate our findings and to obtain additional information not accessible to NMR.

In the Appendix of this thesis you will also find the manuscript for a side project I have been working on. The goal of this side project was to study the effect of several new antibiotics that target the ClpC1 chaperone from *M.tuberculosis* and are thereby able to kill the organism. I investigated the N-terminal domain (NTD) of ClpC1 by NMR and the influence of different drugs on its structure and dynamics. We chose the NTD for this NMR study, since mutational studies have demonstrated that the NTD is the primary interaction site with these antibiotics. Furthermore, full length ClpC1 proved to be insoluble when expressed in a deuterated environment, a strategy necessary for the study of proteins this size.

Chapter 3

Results

Studying dynamics and substrate binding of the mitochondrial intermembrane space chaperone TIM910

As I already mentioned before, the structure of the TIM910 chaperone has been solved more than ten years ago.[143] What has been unclear since then is how TIM910 binds its substrates, where it binds and how it interacts with the different components of the mitochondrial import machinery to direct substrates to their final location. The apparent difficulty to work with aggregation-prone membrane proteins and a highly flexible chaperone complex have so far impeded any structural information on the mechanism of action of TIM910.

I will present here a series of experiments that we performed on the isolated subunits of the TIM910 complex, the hexameric TIM910 complex and substrate-bound TIM910. To obtain information about these systems we had to use a variety of different techniques in order to to generate a coherent view of the chaperone system.

3.1 The apo TIM910 complex

3.1.1 The individual subunits of TIM910 are highly dynamic

Protein expression: The hexameric TIM910 complex can be formed in two ways. Either Tim9 and Tim10 are co-expressed and then directly purified as a hexamer, or Tim9 and Tim10 are expressed and purified separately and then mixed in a 1:1 ratio to spontaneaously form a hexameric complex (described elsewhere[136]). For our experiments we decided for co-expression using the pETDuet system, where we could reach yields more than ten times higher (60 mg/L culture medium) than with separate expression (3 mg/L culture medium per subunit). We expressed all Tim proteins in SHuffle *E.coli* cells. All small Tim proteins contain a twin CX_3C motif that needs to be correctly disulfide-bonded to allow folding and complex formation. SHuffle cells contain a chromosomal copy of the disulfide isomerase DsbC, that aids disulfide formation and rendered most of the expressed protein soluble. Several experiments in our project required the use of free Tim9 and Tim10 subunits. To reach higher protein yields, we used co-expression of TIM910 and unfolded the purified complex with guanidinium-HCl to obtain free subunits. Subsequently, we used the N-terminal 6xHisTag of Tim10 to separate Tim9

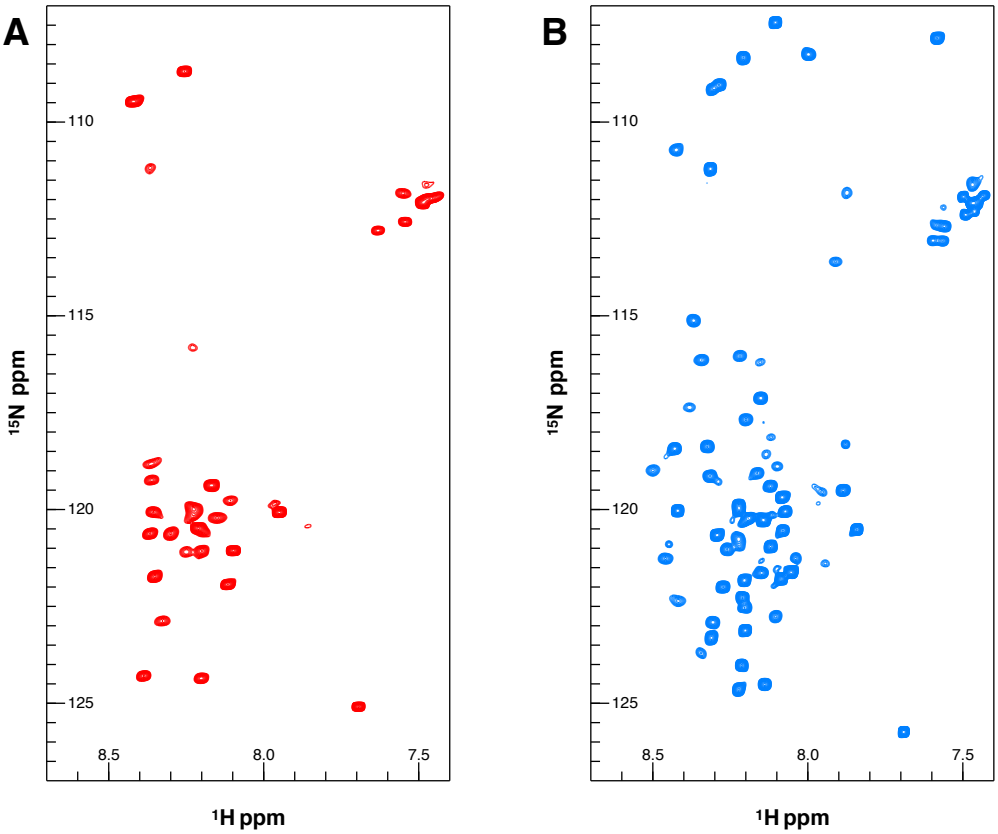


Figure 3.1: **NMR spectra of Tim9 and Tim10.** In Tim9 ^{1}H-^{15}N correlated NMR spectra only about one third of the peaks (**A**) are visible, while three quarters of Tim10 peaks (**B**) can be detected.

from Tim10 on a NiNTA column. For the following NMR experiments we removed the HisTag with TEV protease. Tim subunits with disulfide bonded CX_3C motif easily refold after removal of the denaturant, either by flash dilution or dialysis. Purified subunits behaved in size-exclusion chromatography (SEC) as described in the literature, whereby Tim9 formed dimers and Tim10 mainly monomers with subpopulations of dimers and trimers.[143] This oligomerization behaviour of Tim10 was concentration dependent, meaning that the population of higher oligomers increased with increasing protein concentration.

Structure of reduced Tim subunits: To study the process of oxidative folding and to compare the structure of Tim subunits with reduced and oxidized CX_3C motif, we performed NMR experiments in the presence and

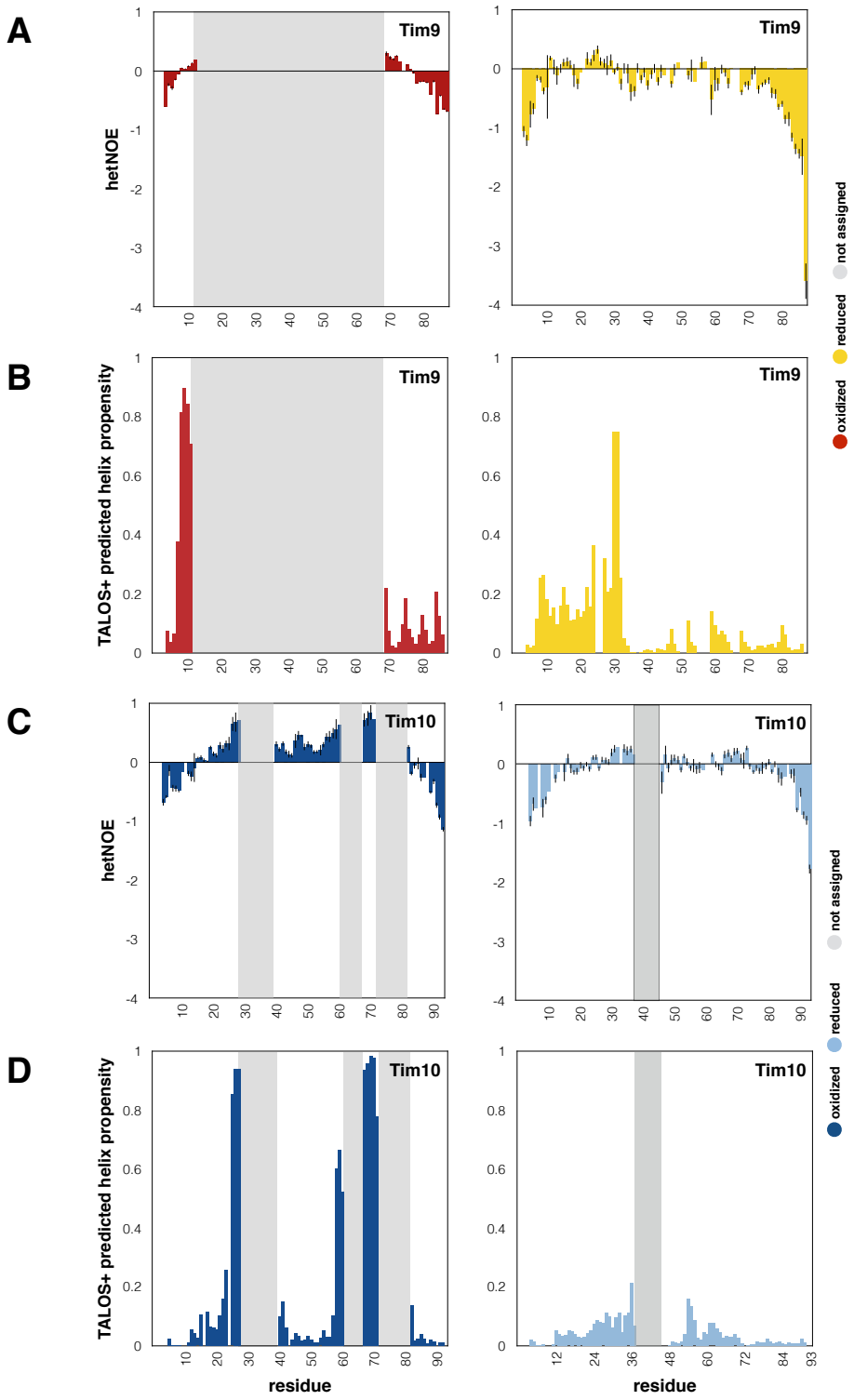


Figure 3.2: **Secondary structure of oxidized and reduced small Tim proteins.** hetNOE of Tim9 (**A**) and Tim10 (**C**) with oxidized and reduced twin CX_3C motif. Shown in grey is the part of the sequence that cannot be seen in NMR spectra. TALOS+ predicted helix propensity of Tim9 (**B**) and Tim10 (**D**) with oxidized and reduced twin CX_3C motif. In contrast to the reduced form, many residues are not visible in the oxidized proteins.

absence of the reducing agent tris(2-carboxyethyl)phosphine (TCEP). We assigned reduced Tim9 and Tim10 subunits and found that NMR spectra of both subunits show characteristics of an unfolded protein, i.e. a narrow ^{1}H chemical shift dispersion (see figure 3.1). We assessed secondary structure propensity and foldedness by using on the one hand assigned chemical shifts and the TALOS+ software to derive secondary structure, and on the other hand ^{15}N heteronuclear Overhauser effect (hetNOE) relaxation experiments (see figure 3.2). Typically, residues in folded areas show hetNOE values of around 0.8, while residues in flexible parts have values of around or below 0. If we look at the right-hand side panels of figure 3.2, we see that the helix propensity in reduced Tim9 and Tim10 subunits is rather low and that the hetNOE values are always around or below zero, leading us to the conclusion that Tim proteins without formed CX_3C motif do not have any secondary structure. Furthermore, reduced Tim9 and Tim10 subunits are unable to form a hexameric complex (SEC data, not shown).

Oxidative refolding of Tim9: Upon removal of the reducing agent the small Tims can reform the disulfide bridges and are again able to form a hexameric complex between Tim9 and Tim10. At pH 6.5 this oxidative refolding is sufficiently slow to be observed by NMR. To this end we recorded a series of short ^{1}H-^{15}N correlated NMR spectra. The rate of conversion of a peak from its position in the reduced to the oxidized form can be fitted and the regions where folding occurs faster or slower evaluated. Figure 3.3A shows a residue-wise plot of folding rate constants of Tim9. Several residues that are located around the two CX_3C motifs already disappear before the first NMR spectrum could be recorded (dead time 15 min). Data of these residues cannot be analyzed (marked grey in the plot). Interestingly, residues between 15 and 40 show faster rate constants, probably indicating that the N-terminal helix of Tim9 folds faster than the C-terminal helix.

Structure and dynamics of oxidized Tim subunits: In contrast to the reduced Tim subunits, oxidized subunits carry short well folded helical regions that are located around the two CX_3C motifs (see figure 3.2 left-hand side panels). Additionally, oxidized Tim subunits have large unfolded regions located at the N- and C-termini and the connecting loop. A striking difference between oxidized and reduced Tim NMR spectra is the disappearance of a considerable amount of peaks in spectra of the oxidized proteins. This phenomenon is especially pronounced in the Tim9 subunit, where 70% of peaks disappear (35% in Tim10). The missing region is located around the two helices and leaves only the flexible termini visible. Peak broadening in

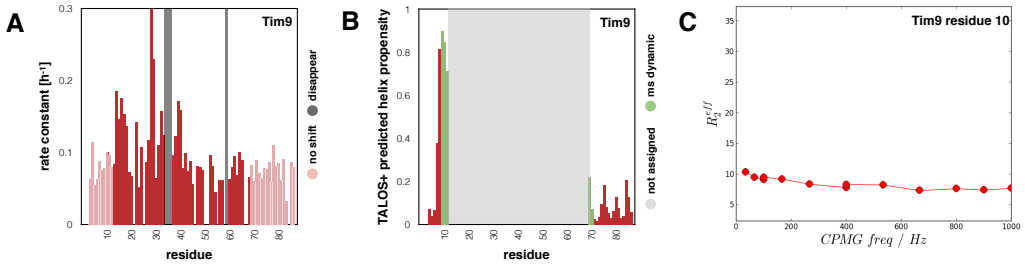


Figure 3.3: **Oxidative folding and millisecond exchange dynamics in Tim9 A:** Per residue rate constant of oxidative folding of Tim9. In light red residues that do not shift upon folding. Residues of the N-terminal helix have higher rate-constants than residues of the C-terminal helix or loop. Four residues (marked in dark grey) close to the CX_3C motif disappear already after 15 min, the dead time of the experiment. **B:** TALOS+ predicted helix propensity of oxidized Tim9. Residues that do not appear in NMR spectra are marked in grey. In green residues that experience millisecond dynamics as seen by CPMG RD NMR experiments. **C:** Residue 10 of Tim9 is the last residue which can be still seen in an NMR spectrum before the disappearing region and shows a non-flat CPMG RD curve.

NMR is often caused by an exchange of the chemical environment of a spin in the µs-ms timescale. A possible way to detect these exchange dynamics is to measure Carr-Purcell-Meiboom-Gill sequence (CPMG) relaxation-dispersion (RD) NMR experiments, where the effective spin relaxation rate constant ($R_{2,eff}$) is measured as a function of a variable repetition rate (ν_{CPMG}) of refocusing pulses applied during a relaxation delay. The presence of non-flat CPMG RD curves indicates exchange dynamics at the µs-ms timescale. In Tim9 residues flanking the disappearing region show non-flat RD CPMG curves (marked in green colour in figure 3.3B). Regarding the underlying cause of these dynamics we can only speculate. A possible scenario could be unfolding/folding of the helical parts or opening and closing of disulfide bonds in the CX_3C motifs.

3.1.2 The TIM910 hexamer has a rigid core and flexible termini

Tim9 and Tim10 form a stable hexamer with three Tim9 and three Tim10 subunits.[143] The size exclusion profile of co-expressed TIM910 after purification on a NiNTA column shows a homogeneous profile with a slight shoulder to the right (see figure 3.4A). AUC of this complex gives a sedimentation profile corresponding to a 61 kDa hexamer. The main peak accounts for 95% of the sample, the minor peak is in agreement with either Tim9 or Tim10

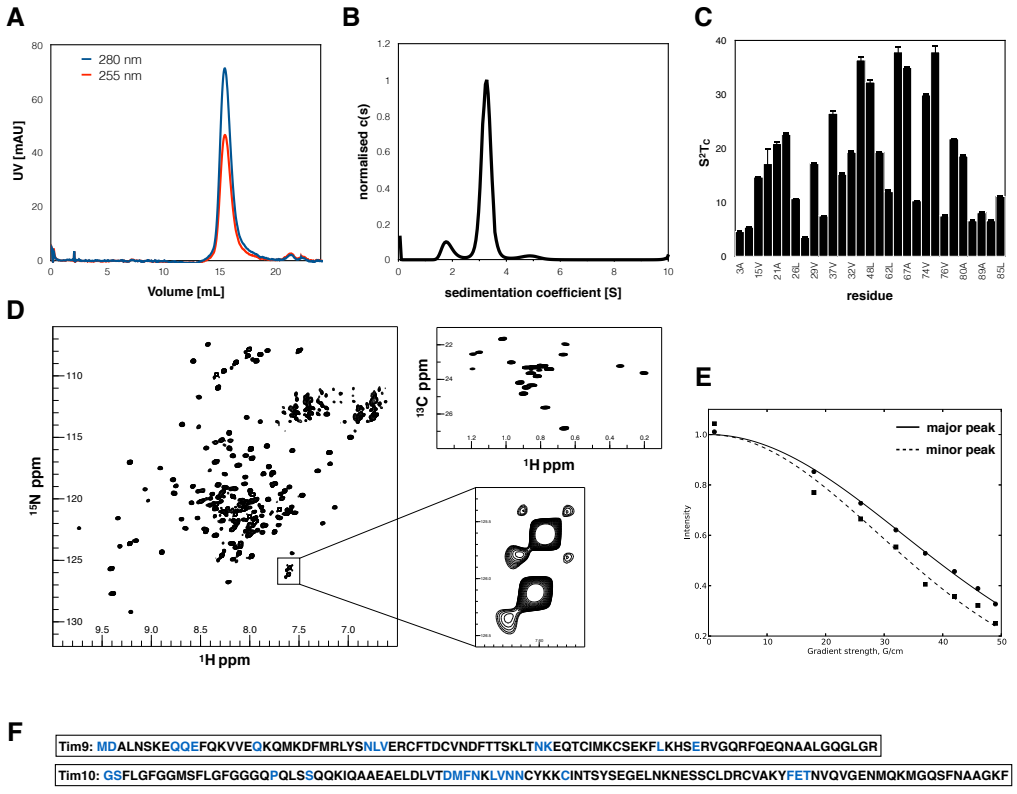


Figure 3.4: **Purification and NMR resonance assignment of hexameric TIM910** **A:** Size exclusion profile of TIM910 on a Superdex200 10/300 GL column showing a homogeneous sample, with a slight shoulder towards smaller species indicating the presence of free subunits. **B:** Sedimentation curve of apo TIM910. The main peak at 3.3 S corresponds to a hexamer of three Tim9 and three Tim10 subunits with a theoretical molecular weight of 61 kDa. The main peak makes up 95 % of the total sample. The minor peak contribution at 1.9 S could correspond to a Tim9 or Tim10 trimer with 27 +/- 3 kDa, making up 2 % of the total sample. **C:** $S^2\tau_c$ values of alanine, leucine and valine residues of apo TIM910. Residues in the center of the sequence are generally more structured than at the ends. A structural representation of the order parameters in the bound form can be found in figure 3.6C. **D:** TIM910 NMR spectra of backbone amides and Leu, Val sidechains. The C-termini of Tim9 and Tim10 subunits of the TIM910 backbone spectrum is highlighted and shows a doubling of peaks. **E:** DOSY rates of the doubled peaks are different for the minor and the major peak. **F:** NMR sequence assignment of the TIM910 hexamer. In blue residues that are either not visible in the NMR spectrum or not assigned.

free subunits (figure 3.4B).

NMR spectra of [^{13}C,^{2}H,^{15}N] (CDN) labelled TIM910 or alanine, leucine, valine (ALV) methyl labelled TIM910 show well dispersed spectra in agreement with a folded α-helical protein. The number of peaks suggests a symmetric behaviour of the TIM910 complex, where all three Tim9 or Tim10 subunits give rise to only one common set of peaks (see figure 3.4D). Using a labelling approach where each time only one of the subunits was isotopically labelled, we were able to assign 85% of all TIM910 resonances (amide, Cα, CO, Cβ) at 35°C. We confirmed this assignment by performing a second assignment at 60°C where we were able to assign 95% of all peaks. The assignment at 60°C could then be back transferred to 35°C by performing a temperature series. Additionally, we assigned 91% of all ALV methyl resonances using NOESY experiments and our previous backbone assignment. Peaks that we could not assign are shown in blue in figure 3.4F. Interestingly, peaks that are not assigned do not appear in the spectrum either due to peak overlap or because of peak broadening (at 60°C all peaks are visible). In contrast to the isolated Tim subunits we could not detect millisecond dynamics in the hexameric TIM910 complex. One exception is residue L54, which is situated in the connecting loop and shows non-flat CPMG RD curves in both, backbone and methyl, experiments. A list of all methyl CPMG RD curves can be found in supplementary figure A.1 and a selection of residues in figure 3.5D.

An additional set of peaks can be found in TIM910 NMR spectra which does not correspond to the unassigned regions but to a small subpopulation of free Tim9 and Tim10 subunits. The last 5-10 C-terminal residues show two sets of resonances. This second set of peaks does not represent a second conformation of one of the Tim9 or Tim10 subunits, but can be attributed to free subunits inside the sample. We were able to show this by performing a 2D diffusion-ordered spectroscopy (DOSY) experiment, which revealed that the major population diffuses like a hexameric protein and the minor population like a free subunit (see figure 3.4E). This peak doubling can be easily observed on the C-terminus, which is highlighted in figure 3.4D. We think that the presence of these free subunits is due to an equilibrium between hexamer and free subunits, since we are not able to remove them by SEC. A series of NMR experiments at increasing temperature shows that the equilibrium between hexamer and monomer is inverted with increasing temperature, indicating disassembly of the hexameric complex. Thus, the ratio between hexamer and monomer depends on the temperature used for the measurement.

If we compare the structuredness of Tim9 and Tim10 in their different states (reduced, oxidized and in the TIM910 complex) we can derive the following: reduced Tim9/Tim10 are unfolded, oxidized Tim9/10 have a folded

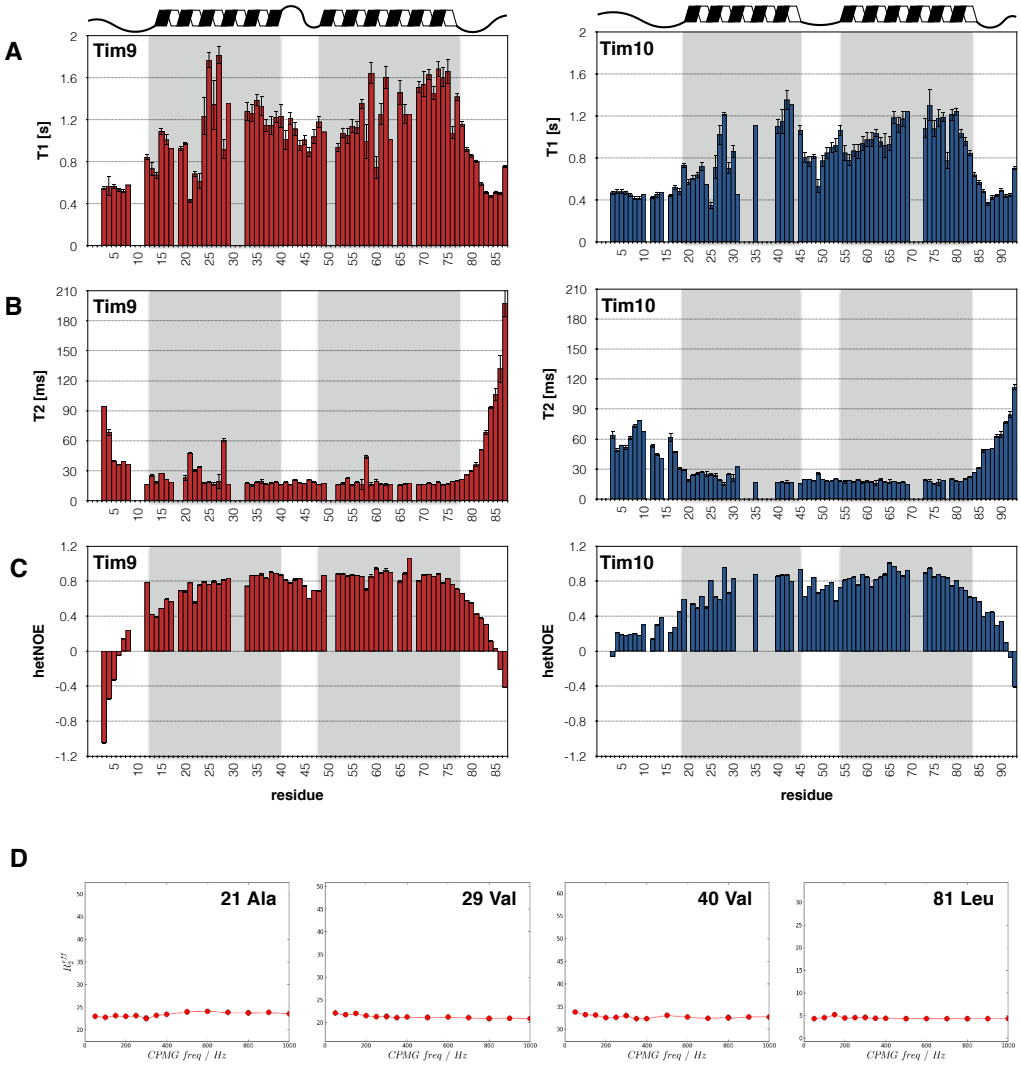


Figure 3.5: **NMR relaxation experiments on hexameric TIM910.** T_1 relaxation times (**A**), T_2 relaxation times (**B**) and hetNOE (**C**) of Tim9 (red) and Tim10 (blue) residues in the TIM910 hexamer. A cartoon on top of each picture as well as a marked grey area depicts the localisation of the two helices according to the X-ray structure (PDB: 3DXR). In both subunits the N- and C-termini that are not visible in the X-ray structure are highly flexible, while the loop region in-between the two CX_3C motifs is rather rigid. **D:** Methyl CPMG RD curves of four residues in apo TIM910. No residue in apo TIM910 except one (L54) shows non-flat CPMG RD curves in both methyl based and amide based CPMG experiments. Curves for all methyl residues are shown in supplementary figure A.1.

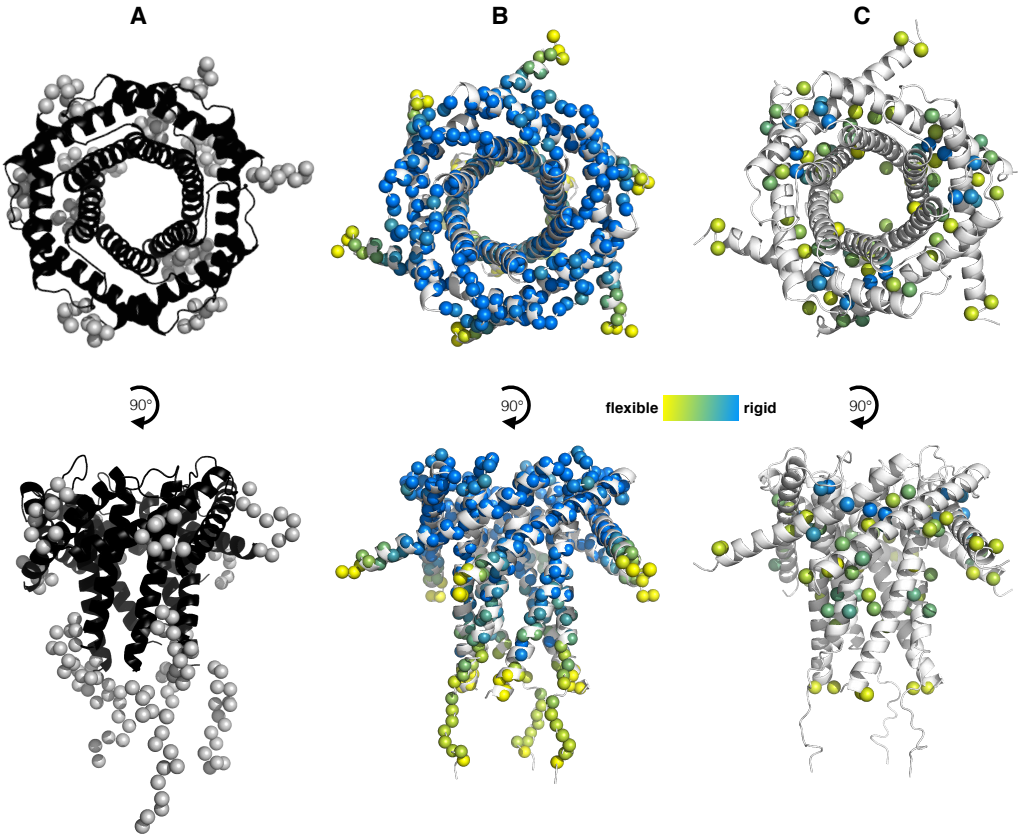


Figure 3.6: **Structuredness of the TIM910 hexamer A:** SAXS model of the TIM910 hexamer generated by the program CORAL using the TIM910 X-ray structure (3DXR) as a template. SAXS data was recorded in batch mode using 0.5 mg/mL TIM910 protein.The N-termini are extended and seemingly unfolded as observed by NMR, while the C-termini appear to be more compact. **B:** hetNOE values of protein backbone amides plotted on a TIM910 iTasser model (including parts that are missing in the X-ray structure). Yellow coloured residues are flexible, blue residues rigid. As can be seen also in figure 3.5, N- and C-termini are very dynamic while the twin CX_3C motifs and the connecting loops are rigid. **C:** Order parameters derived from relaxation-violated coherence transfer NMR experiments of ALV sidechains (raw data see figure 3.4C).

helical part around the two CX_3C motifs, but also large flexible parts, and finally hexameric TIM910 has a rigid core and only the extreme N- and C-termini are unfolded. We thus assume that Tim9 and Tim10 can reach their final fold only in the TIM910 complex. NMR relaxation experiments (T_1, T_2 and hetNOE) showing the rigid and flexible parts of TIM910 can be found in figure 3.5A-C as well as order parameters for ALV methyl groups in figure 3.4C. For a better visual representation, the hetNOE data (figure 3.6B) and the methyl order parameters (figure 3.6C) are plotted on an iTasser model of TIM910 derived from the X-ray structure (PDB 3DXR). While the N- and C-terminal helices and the connecting loop are rigid, the N- and C-termini are flexible. The flexible N-terminal part of TIM10 is larger than in TIM9.

In addition to our NMR experiments we performed a SAXS measurement of the TIM910 complex and used the software CORAL to build the parts missing in the X-ray structure (figure 3.6A). The X-ray structure is shown in black and the modeled residues in grey. While the C-termini seem to be more compact, the N-termini are elongated.

3.2 TIM910-substrate complexes

3.2.1 TIM910 binds substrates of the inner and outer mitochondrial membrane

The goal of this work was to study the interaction of TIM910 with its substrates. How TIM910 binds its client proteins is currently unknown. The lack of structural and biochemical knowledge also originates from the fact that *in vitro* reconstitution of TIM910 bound to a full length substrate protein has never been described before. *In vivo* data from yeast suggests that TIM910 is necessary for the import of mitochondrial inner and outer membrane proteins.[41,64,147] These two classes of proteins adopt significantly different folds in their membrane inserted form. While inner membrane proteins have an α-helical conformation,[91] outer membrane proteins adopt a β-barrel conformation.[9,48] The first step in this project was thus the *in vitro* reconstitution of a TIM910/substrate complex.

TIM910 can rescue membrane proteins from aggregation

An important point, before we started our *in vitro* complex reconstitution experiments, was to ensure that the TIM910 protein we use is active and thus able to rescue membrane proteins from aggregation. To this end we implemented a cell-free experiment where we expressed the GDP/GTP car-

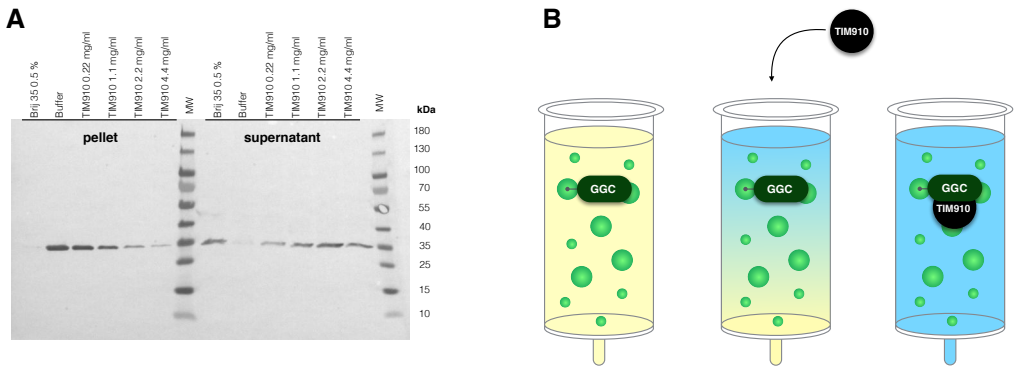


Figure 3.7: **The TIM910/substrate complex can be reconstituted *in vitro*. A:** GGC is expressed in a cell-free environment in the presence of detergent (Brij 35), buffer or different concentrations of TIM910. The amount of GGC in solution can be quantified by a western blot staining His-tagged GGC with an anti-His antibody. Without detergent or TIM910 all GGC is insoluble. Upon addition of detergent all GGC is found in the soluble part. The same effect can be achieved by adding TIM910. GGC is kept in solution in a TIM910 concentration-dependent manner. **B:** Reconstitution of TIM910 in the presence of GGC. GGC is bound to a NiNTA column under denaturing conditions (yellow), subsequently the denaturant is gradually removed in the presence of TIM910. After washing out free TIM910, TIM910-GGC complex can be eluted from the column.

rier (GGC) in the presence of increasing amounts of purified TIM910. We quantified the amount of soluble/insoluble membrane protein by performing a western blot using an anti-His antibody against His-tagged GGC. Figure 3.7A shows that in the presence of detergent (Brij35) all GGC is soluble, while addition of buffer yields no soluble membrane protein. Finally, addition of TIM910 keeps the membrane protein in solution in a concentration dependent manner. Thus, TIM910 is able to bind and rescue newly translated GGC.

TIM910 forms stable complexes with mitochondrial carriers *in vitro*

To reconstitute a TIM910/substrate complex *in vitro* we chose GGC and AAC as representative substrates of the mitochondrial carrier family. All experiments were performed with substrate proteins containing no cysteine residues, expressed as inclusion bodies and purified under denaturing conditions.

Dialysis of the denaturant in presence of different concentrations of TIM910 did not result in any soluble membrane protein. Similarly, flash dilution of the membrane protein into TIM910 containing buffer solutions at different chap-

erone concentrations was not successful. We repeated the experiment by only drop-wise addition of membrane protein to the TIM910 solution, but could not obtain soluble substrate/chaperone complexes. We conducted all the described experimental set-ups at room temperature and at 4°C and changed the dialysis times and volumes as well as the final protein concentration and ratio of membrane protein and chaperone. In none of these experiments TIM910 succeeded in keeping GGC in solution. We speculated that the aggregation process of GGC is faster upon removal of the denaturant than the complex formation of the TIM910 hexamer, which we assumed was the active conformation of TIM910. In fact, SEC profiles of TIM910 in the presence of different amounts of guanidinium-HCl showed that hexamer reassambly only occurs below a guanidinium-HCl concentration of 100-50 mM. We concluded that keeping the membrane protein from aggregation for a longer period of time would be critical for complex formation. To this end we developped a protocol whereby His-tagged GGC is attached to a NiNTA column under denaturing conditions and the denaturant removed gradually in the presence of TIM910. The complex is eluted when the guanidinium-HCl concentration is less than 50 mM (for a more detailed protocol see Materials and Methods). A schematic representation of the set-up can be found in figure 3.7B. Using this procedure we were able to rescue about one quarter of the invested membrane protein and obtain homogeneous chaperone/substrate complexes that could be concentrated to high concentrations for NMR experiments (up to 0.5 mM). Moreover, we used this protocol also successfully with two other mitochondrial carrier proteins, namely AAC3 (AAC) and a short version of AAC3 containing only two transmembrane helices (AAC_{short}, residue 50-180).

TIM910 can adapt to the length of its client proteins

Although TIM910/substrate samples are stable over weeks at 4°C and several days at 35°C for NMR measurements, running SEC on Superdex columns always resulted in a separation of chaperone and substrate, whereby only the chaperone eluted from the column. We suspect that this is due to a non-specific interaction with the gel filtration resin, since we could perform runs on Agilent Sec5 columns without destroying the complex. Unfortunately, the separation power of Sec5 columns in this molecular weight range was not sufficient to separate free TIM910 from substrate-bound TIM910. To circumvent this problem we used AUC to estimate the quality of our chaperone/substrate samples. As you can see in figure 3.8A, TIM910-GGC and TIM910-AAC_{short} samples display homogeneous sedimentation profiles. The difference in sedimentation coefficients reflects a different stoichiometry for shorter and longer substrates. While the 34 kDa-large GGC sediments ac-

cording to a complex bound to two TIM910 hexamers, 15 kDa-large AAC_{short} sedimented as a complex of one TIM910 hexamer bound to one membrane protein.

To further investigate this effect of substrate length on the stoichiometry of the complex, we performed NMR DOSY experiments. To obtain better signal/noise ratios we used methyl labelled samples to measure diffusion rates, whereby in each experiment one of the binding partners was either labelled on ALV, AI or LV. Experimental diffusion rates can be found in figure 3.8B, the labelled binding partner is always denoted with a star (*). We compared these values with theoretical diffusion rates of X-ray structures or structural models derived with the HYDRONMR software.

While full-length mitochondrial carriers (GGC, AAC) form complexes with a ratio of one membrane protein per two TIM910 hexamers, the shorter AAC_{short} forms a 1:1 complex. Note also that the diffusion rate of GGC labelled TIM910-GGC complex and TIM910 labelled TIM910-GGC complexes are nearly identical (light and dark green figure 3.8B). Slight differences between experimental and theoretical diffusion rates can be explained by mobility of certain parts of the protein in the measured complexes. For example, the iTasser model of TIM910 appears bigger than the measured TIM910 complex. We know though that the N- and C-termini of TIM910 are highly flexible and would thus not behave like a rigid object, therefore the measured TIM910 seems smaller than the structural model. The same holds true for all theoretical values of SAXS models, which do not take into account flexible parts in the protein.

In addition to the sedimentation coefficient in AUC and the translational diffusion coefficient in DOSY experiments, we also can derive the stoichiometry of TIM910/substrate complexes by evaluating the rotational diffusion of these complexes in relaxation experiments. To this end we measured methyl order parameters of apo TIM910 and TIM910-GGC complexes. The $S^{2*}\tau_c$ values of these complexes are shown in figure 3.8C on the left. Furthermore, we used the theoretical τ_c of our SAXS envelope for apo TIM910 and TIM910-GGC and assumed that the order parameter of residues in the rigid core part of TIM910 is around 0.8. This theoretical values, are in good agreement with the measured values depicted in green and black. The fact that the ratio between $S^{2*}\tau_c$ values of apo TIM910 and TIM910-GGC is approximately two for all residues and the good agreement with the theoretical SAXS based values suggest that TIM910-GGC indeed form a 2:1 complex. The ratio of these theoretical SAXS-derived values is shown in grey in figure 3.8C (right panel). Divergence of this theoretical value would indicate that these residues show either higher or lower flexibility in their substrate-bound form. The ratio of experimental S^2 values of apo TIM910 and TIM910-GGC

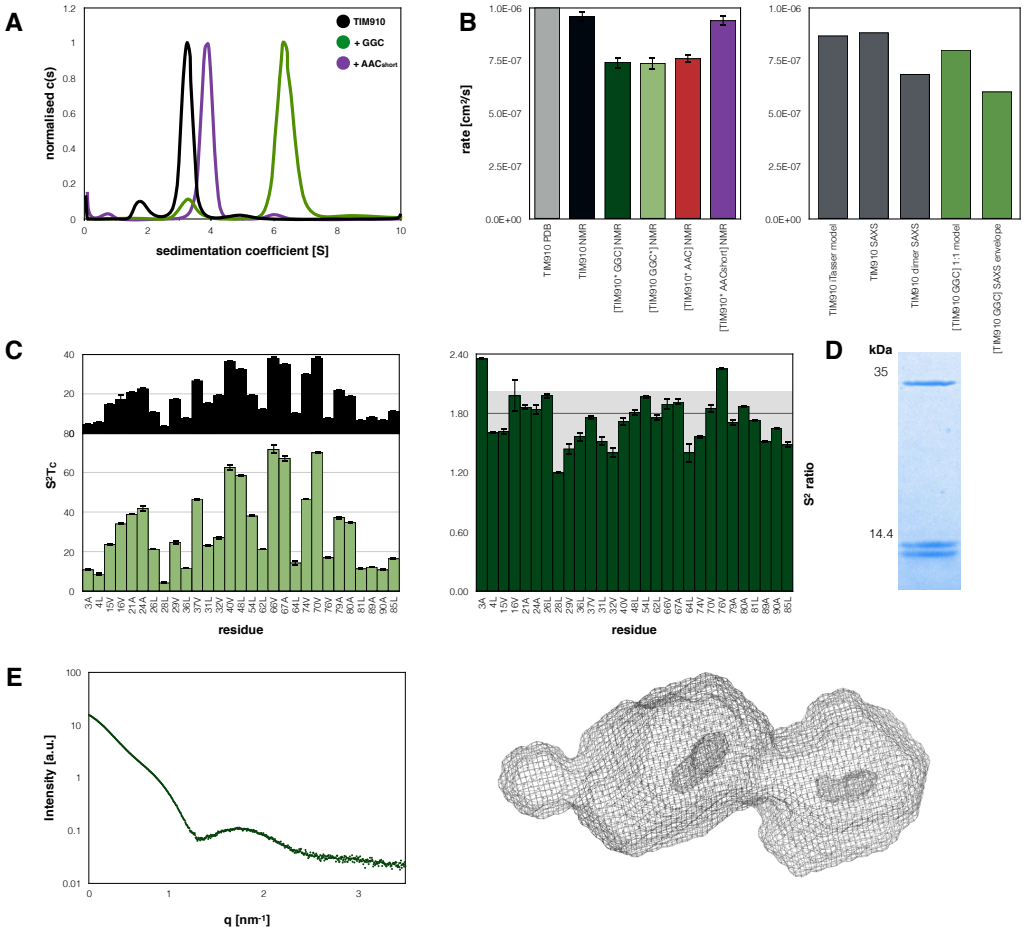


Figure 3.8: **The stoichiometry of the TIM910-substrate complex depends on the lengths of the substrate. A:** Sedimentation curve of apo TIM910 (black), TIM910-AAC_{short} (purple) and TIM910-GGC (green). TIM910-AAC_{short} shows a main peak at 3.8 S, corresponding to a complex of one TIM910 hexamer and one membrane protein with a molecular weight of 75 kDa. TIM910-GGC has a main peak at 6.4 S making up 82 +/- 4 % of the total sample, and a minor peak at 3.4 S making up 6 % of the total sample. The main peak is in agreement with either a 2:1 or a 2:2 complex of TIM910:GGC depending on the shape of the complex. The minor contribution corresponds to free TIM910. For apo TIM910 see figure 3.4A. **B:** Experimental DOSY rates of apo TIM910 (black), TIM910 labelled TIM910-GGC (dark green), GGC labelled TIM910-GGC (light green), TIM910-AAC (red) and TIM910-AAC_{short} complexes compared to the theoretical diffusion rate of the X-ray TIM910 structure (PDB: 3DXR), a TIM910 iTasser model derived from the X-ray structure, the apo TIM910 SAXS model, a dimer of the apo TIM910 SAXS model, a hypothetical TIM910-GGC 1:1 complex and the experimental SAXS envelope of the TIM910-GGC complex. **C:** $S^2\tau_c$ values of alanine, leucine and valine residues in GGC bound (light green) and free (black) TIM910. S^2 ratio of GGC bound TIM910 and free TIM910. The theoretical S^2 ratio derived using HYDRONMR is depicted in grey. Residues 28, 29, 31, 32, 36, 64, 74, 85 and 89 show a slight decrease in structuredness upon substrate binding. **D:** Example SDS PAGE of a TIM910-GGC sample. The band close to 35 kDa represents GGC, the two bands at around 10 kDa the Tim9 and Tim10 subunits of TIM910. **E:** SAXS scattering curve of the TIM910-GGC complex and the corresponding DAMMIF envelope, where two cavities are visible.

shows that residues around 30 have a higher flexibility in the substrate-bound form than in the apo form (see figure 3.8C).

We further corroborated the stoichiometry between TIM910 and GGC by estimating the intensity of SDS PAGE bands of TIM910 (two bands at 10 kDa) and GGC (one band at 34 kDa) in TIM910-GGC complex samples (see figure 3.8D). To this end we integrated the SDS PAGE bands of 27 independent complex samples using the Image Lab software and took the average of all samples. The result is in perfect agreement with a 2:1 TIM910:GGC complex (see supplementary figure A.2).

Finally, we implemented a protocol for SAXS experiment to measure aggregate-free TIM910-GGC complexes. Performing SAXS experiments with these chaperone/substrate complexes is challenging, since SAXS is a technique that is extremely sensitive towards aggregation and membrane proteins are prone to aggregation. Thus, using a batch mode set-up did not give sufficiently good data to build a structural model of the complex. To avoid aggregation we directly measured our TIM910-GGC samples after elution from an online NiNTA column. This allowed us to obtain aggregate-free scattering curves and to build a first model of the complex using the DAMMIF software. As you can see in figure 7E, the resulting SAXS envelope contains two visible cavities which correspond to the two central cavities of two TIM910 hexamers.

To summarize, TIM910 binds its substrate in a modular manner whereby the stoichiometry of the complex depends on the length of the substrate.

Substrate binding induces peak broadening in TIM910 NMR spectra

NMR spectroscopy is the ideal tool to investigate binding sites of dynamic complexes in solution. The change in the local chemical environment upon substrate/ligand interaction can be used to determine the binding site. A change in environment typically manifests itself either in a change of chemical shift of the concerned residue or a change in peak intensity. ^{1}H-^{15}N as well as ^{1}H-^{13}C correlated NMR spectra of TIM910 labelled TIM910/substrate complexes suffer from extreme peak broadening. Spectra of TIM910-GGC (A), TIM910-AAC (B) and TIM910-AAC$_{\text{short}}$ (C) of the Leu/Val region (top), the Ala region (center) and the backbone amides (bottom) are shown in figure 3.9. Apparently the symmetry of the TIM910 complex is not disturbed upon substrate binding, which would manifest itself in a doubling or tripling of peaks.

Thus, we assume that the binding entity in TIM910/substrate complexes is the TIM910 hexamer and the different conformations of the substrate lead

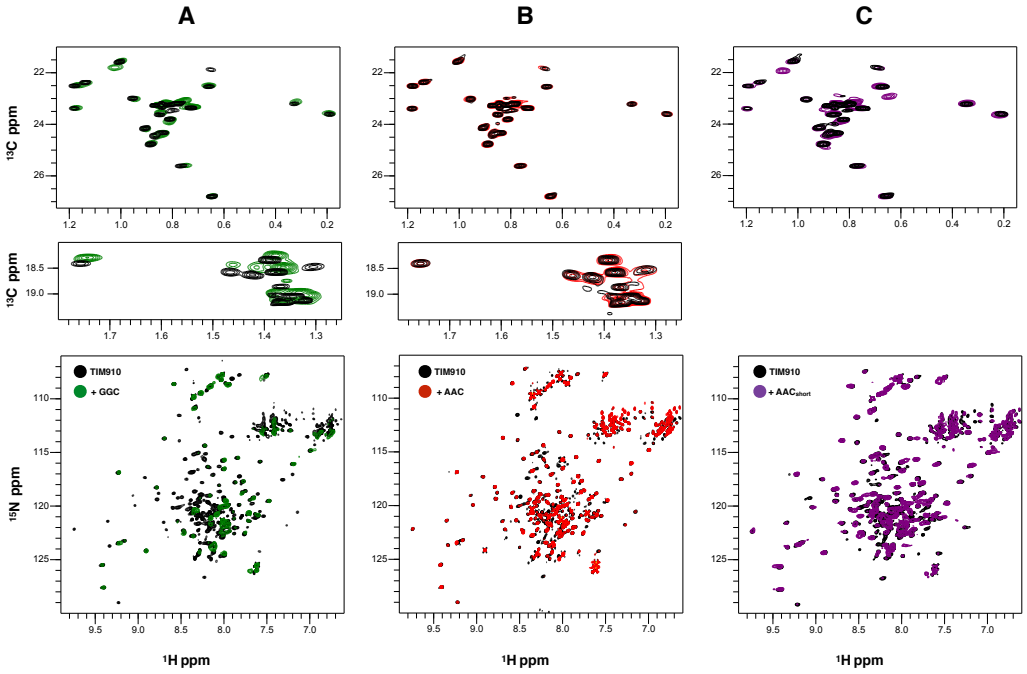


Figure 3.9: **NMR spectra of apo TIM910 and TIM910/substrate complexes.** ^{1}H-^{13}C correlated spectra of the leucine, valine region on the top, the alanine region in the middle and ^{1}H-^{15}N correlated spectra of the backbone amides at the bottom. **A:** Spectra of apo TIM910 (black) and GGC bound TIM910 (green). **B:** Spectra of apo TIM910 (black) and AAC bound TIM910 (red). **C:** Spectra of apo TIM910 (black), AAC_{short} bound TIM910 (purple).

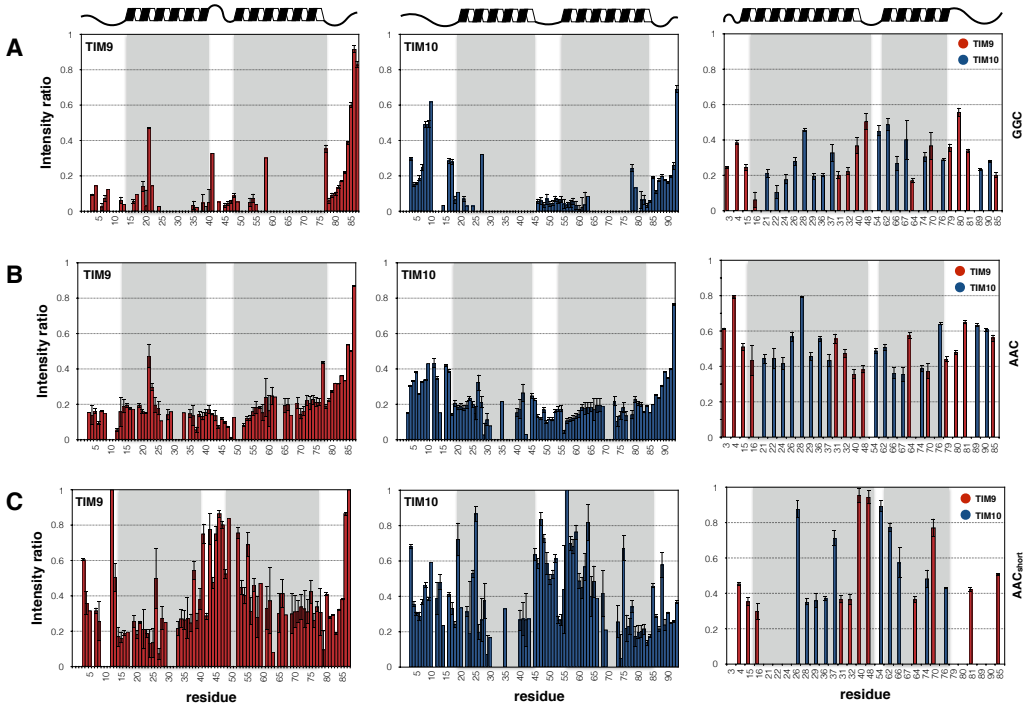


Figure 3.10: **NMR peak intensity of TIM910/substrate complexes.** Intensity ratios of the TIM910-GGC (**A**), the TIM910-AAC (**B**) and the TIM910-AAC$_{\mathrm{short}}$ (**C**) complex relative to apo TIM910. Left backbone amides of the Tim9 subunit in TIM910, middle backbone amides of the Tim10 subunit in TIM910 and right ALV sidechains of TIM910. Helical regions are depicted by a helix cartoon on top of each image and a grey area.

to an averaged effect on TIM910.

NMR intensity loss affects the whole TIM910 chaperone

Frequently the substrate binding site can be revealed by looking at those peaks that suffer from intensity loss. This loss in peak intensity can be either due to the increased size of a complex or due to local millisecond dynamics. The peak height ratios of substrate-bound and apo TIM910 can be found in figure 3.10. The left and center plots show peak ratios of backbone amide spectra, the plot on the right peak ratios of specifically labelled ALV/LV samples. Interestingly, the extent of peak height reduction are not the same for all the samples. TIM910-GGC (A) suffers more pronounced intensity loss than TIM910-AAC (B) and TIM910-AAC_{short} peaks are best visible. A possible reason for this could be the different amino acid composition of the three samples or a different dynamic behaviour of the substrate. Generally, all peaks in TIM910/substrate samples are broadened and do not allow an accurate determination of the binding site. This is even more evident when the intensity loss is plotted on the TIM910 structure (supplementary figure A.3). A vague estimation of the binding site can be derived from the peak intensity profile of the TIM910-AAC_{short} sample. It seems that the connecting loop and the N- and C-termini are less affected than the helical region (supplementary figure A.3).

NMR peak broadening is caused by millisecond exchange dynamics in the substrate binding site

Obviously we expect that the molecular weight of our TIM910/substrate samples is increased from about 60 kDa to about 150 kDa and thus the intensity of peaks reduced due to molecular tumbling and an increase in R_2 relaxation rate. Nevertheless, the degree of intensity loss lead us to suspect that the change in molecular weight was not the only factor causing peak broadening. To confirm our hypothesis we performed CPMG RD experiments as described in section 3.1.1. We used specifically methyl labelled samples for all NMR experiments involving chaperone/substrate complexes to profit from the intrinsically higher intensity of methyl groups, to obtain a sufficient signal/noise ratio. We found three different categories of residues in our samples: 1) residues that showed flat CPMG RD curves, 2) residues that showed non-flat CPMG RD curves and 3) residues that disappeared upon substrate binding. Additionally we observed the appearance of one (GGC) or two (AAC_{short}) new peaks in our spectra that we could not assign due the low quality of these spectra. These peaks were thus not considered for

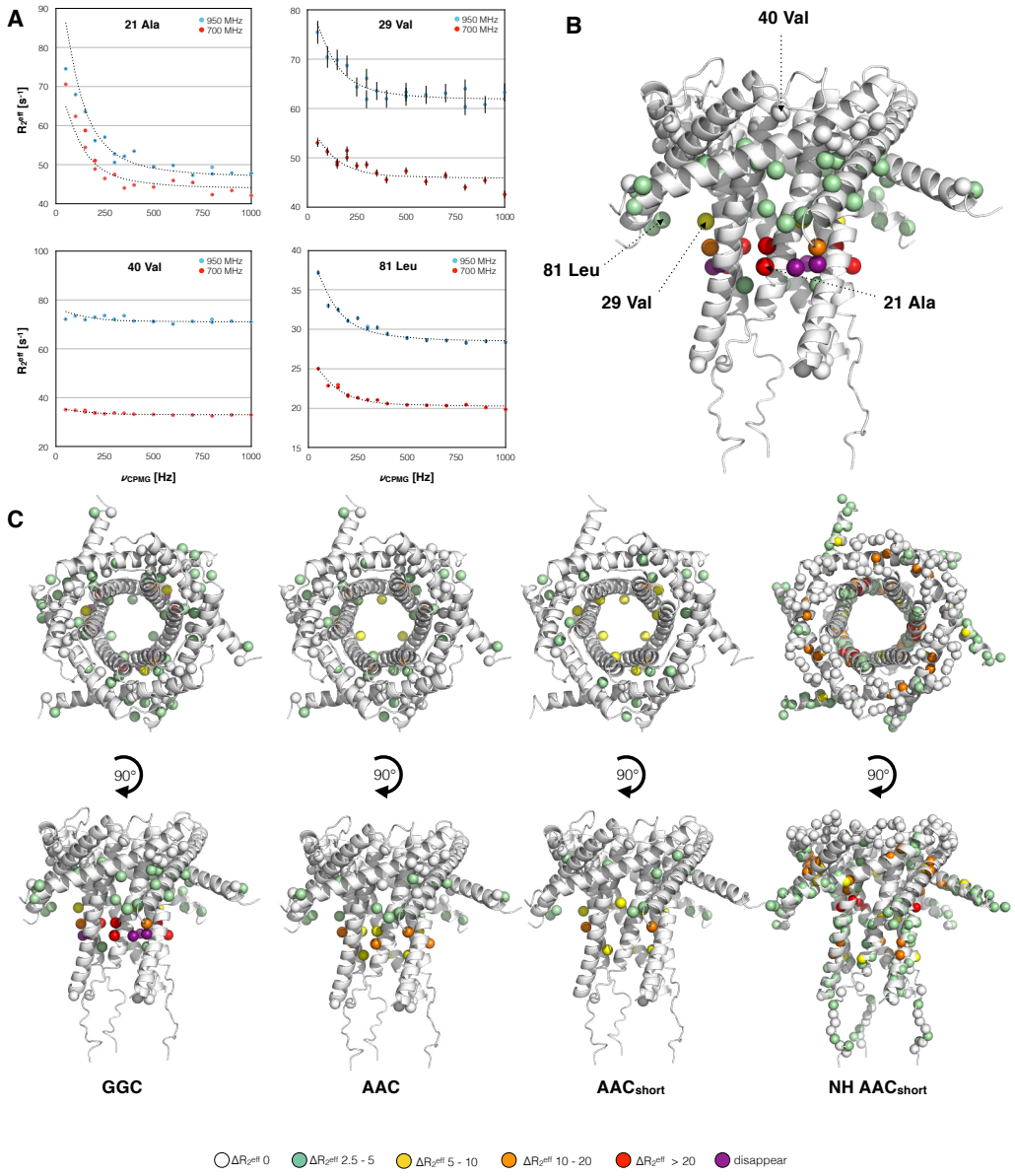


Figure 3.11: **CPMG relaxation dispersion of TIM910/substrate complexes. A:** CPMG RD curves of four example residues of TIM910 bound to GGC. **B:** Position of example residues from **A** on TIM910. **C:** $R_{2,\mathrm{eff}}$ difference of TIM910 in TIM910/substrate samples plotted on the structure. From left to right: TIM910-GGC (ALV methyl groups), TIM910-AAC (ALV methyl groups), TIM910-AAC_{short} (LV methyl groups) and TIM910-AAC_{short} (backbone amides).

further analysis. According to the difference in $R_{2,eff}$, residues with non-flat CPMG RD curves can be sorted in different categories. Figure 3.11A shows four examples of CPMG RD curves and the position of these residues on the TIM910 structure (figure 3.11B). The difference in $R_{2,eff}$ is generally higher in TIM910-GGC samples, in agreement with the more pronounced loss in peak intensity of these samples. Residues that experience these millisecond dynamics and the difference in $R_{2,eff}$ are plotted on TIM910 structures in figure 3.11C.

We conclude that the underlying reason for the low quality of NMR spectra of TIM910/substrate samples are millisecond dynamics resulting in peak broadening. Fitting parameters of the CPMG RD data are shown in figure 3.14D and will be discussed later.

TIM910 binds its substrate in a hydrophobic cleft formed by the N- and C-terminal helices

The localization of residues that experience millisecond dynamics and the fact that not all residues are affected, suggests that the substrate binding site is situated around this area. To prove this hypothesis we performed a series of NMR experiments that either directly or indirectly reveal the binding site of the substrate.

Solvent paramagnetic relaxation enhancement (sPRE): A paramagnetic probe induces higher relaxation rates in residues that are in close spatial proximity to the probe. If this probe is in solution, residues that are exposed to the solvent experience higher relaxation rates than residues that are protected from the solvent. In the case of protein-protein interactions, residues in the binding site are shielded from the solvent and experience less relaxation enhancenment in the bound state than in the apo state. The difference in relaxation behaviour in the apo and substrate-bound form thus reveals the binding site. We used Gd(DTPA-BMA) as a soluble probe and measured relaxation rates in apo TIM910 and TIM910 bound to either GGC or AAC at different Gd(DTPA-BMA) concentrations. The difference in the slope of the concentration dependent ^{1}H R_1 rate is used to determine the solvent protected area upon substrate binding and is shown in figure 3.12A/B. Plotting these values on the structure of TIM910 shows that the connecting loop is not solvent protected while the exterior of the N- and C-terminal helices and the N- and C-termini are protected from the solvent upon substrate binding.

Paramagnetic relaxation enhancement (PRE): Similar to sPRE experiments, paramagnetic probes can be also covalently attached to one of the binding partners to determine the binding site. We chose two positions in GGC where we introduced a S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL) probe covalently attached to a cysteine. First we attched the MTSL tag to the only cysteine in wild type GGC (Cys 222), this cysteine is located in a site that is predicted to be a transmembrane helix. The second probe we attached at the N-terminus of GGC, to this end we mutated residue 222 to a serine and introduced a cysteine in position 8. Residues that show a higher ^{1}H R_1 relaxation rate in the substrate-bound form must be in close proximity to the MTSL tag and thus the substrate protein. In both cases residues between the inner and outer helix experience the highest relaxation rates (see figure 3.12C/E). Thus, the position of the tag does not change the relaxation profile, leading us to the conclusion that the substrate must be moving while it is bound to the chaperone. This is further supported by the fact that we can only observe one set of peaks in NMR spectra, which can only occur if all three Tim9/Tim10 subunits experience the same chemical environment over time. A statically bound substrate would present different amino acids to the three different Tim9/10 subunits thus giving rise to three different chemical environments.

Transferred NOE: Instead of using paramagnetic tags the binding site can also be revealed by taking advantage of nuclear Overhauser effect (NOE)-based magnetization transfer experiments. The prerequisite of efficient NOE magnetization transfer between two sites is close spatial proximity, thus only residues that are close in space to the binding partner will experience an effect (for more details regarding the experimental set-up see chapter Materials and Methods). For this experiment we prepared two samples: 1) a complex between protonated GGC and ALV labelled TIM910 in deuterated buffer and 2) apo ALV labelled TIM910 in deuterated buffer. The apo TIM910 sample only served to measure the background and ensure that the difference in magnetization transfer only originates from the bound GGC. GGC resonances in a spectral region devoid of TIM910 resonances were saturated and this saturation transferred via intermolecular cross-relaxation to residues of TIM910 that are close in space. The difference in intensities between an experiment without cross-relaxation transfer and with transfer reveal regions of TIM910 that are in close proximity to GGC. The difference between the apo TIM910 experiment and the TIM910-GGC experiment can be seen in figure 3.12E. In contrast to the other two presented experiments this transferred NOE measurement shows an effect which is much more localized. The

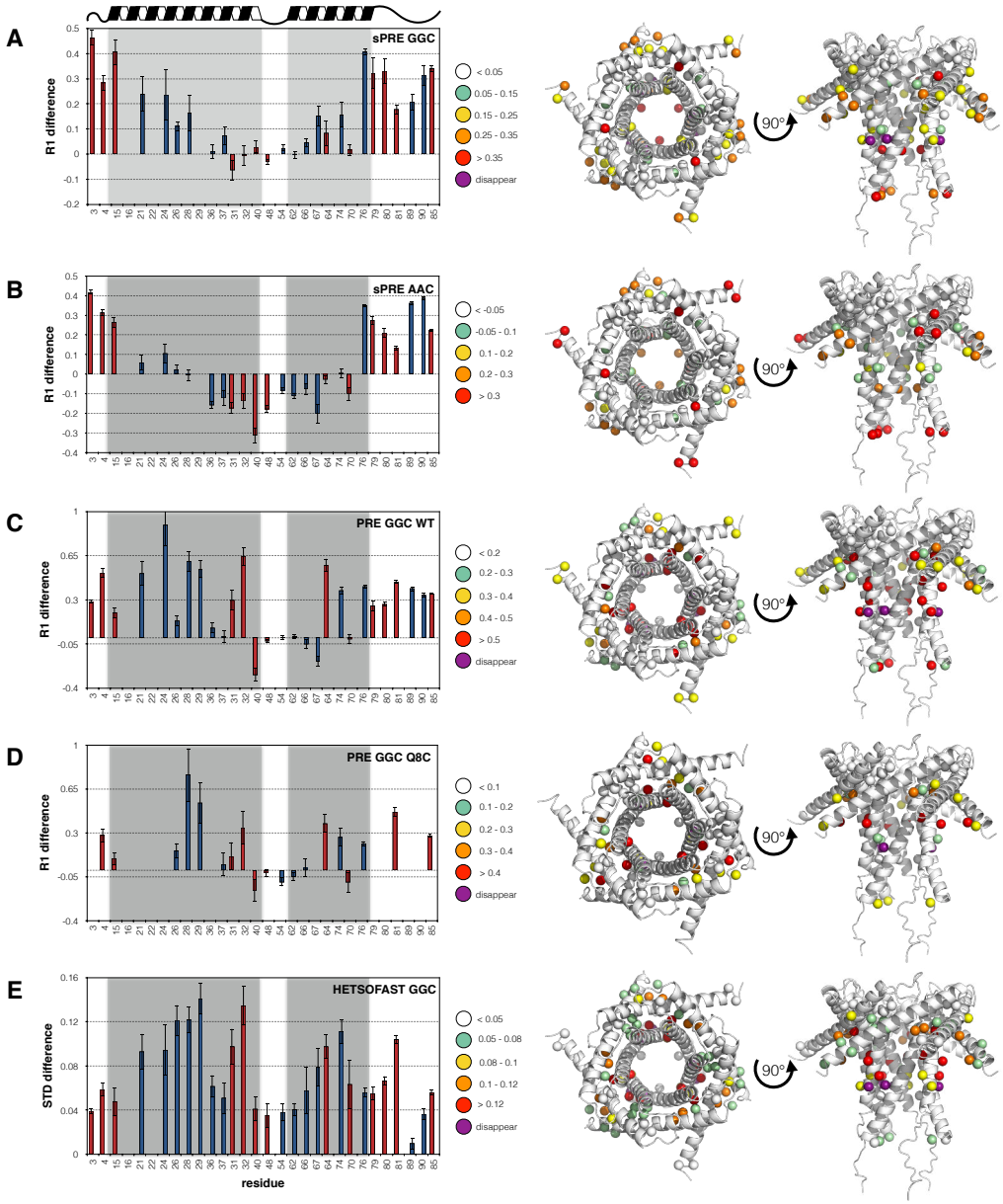


Figure 3.12: **Determining the TIM910 substrate binding site by NMR.** Shown are methyl alanine, leucine and valine sidechaines of TIM910. Residues belonging to Tim9 are coloured in red, residues belonging to Tim10 in blue. The position of the two helices is indicated by a cartoon and marked as grey area in the image. Differences in ^{1}H R_1 of solvent-PRE measurements in the free and GGC (**A**) or AAC (**B**) bound form. ^{1}H R_1 difference of TIM910 bound to paramagnetic MTSL-GGC or diamagnetic reduced MTSL-GGC. The MTSL tag has been attached to either residue 222 (**C**) or 8 (**D**). **E:** Difference in magnetization transfer in a HETSOFAST experiment between apo TIM910 and GGC bound TIM910.

substrate binding site can be easily detected between the inner and outer helix of TIM910.

Outer-membrane TIM910 substrates use the same binding site as inner-membrane substrates

All data we derived so far about client binding and complex formation is based on proteins of the mitochondrial inner membrane. Another goal of this project was to obtain information about the interaction with substrates from the outer mitochondrial membrane and describe differences and similarities in the binding behaviour. Although we tested all the aforementioned reconstitution assays with the two main outer membrane proteins, POR1 (VDAC) and Tom40, we could not obtain stable chaperone/substrate samples. Interestingly, TIM910 and Tom40 eluted together from NiNTA columns after reconstitution, but the complex was not stable enough to be concentrated. We thus decided to change our strategy and use short peptides of outer membrane proteins to detect the binding site. We received two types of peptides from our collaborators (Doron Rapaport and coworkers): 1) A linear peptide representing, in the folded protein, the last two C-terminal β-strands of human VDAC and the connecting loop (residue 257-279). 2) A cyclic version of this peptide, which has been shown to increase the propensity of β-hairpin formation between the two β-strands (see figure 3.13B).

NMR titration experiments with these two peptides showed that 1) the linear peptide does not bind to TIM910, 2) binding of the cyclic peptides causes chemical shifts and peak broadening in TIM910 NMR spectra. The location of these shifts/peak broadening perfectly coincides with the binding site previously determined for mitochondrial carriers (see figure 3.13A/C/D and 3.15C). Cross-linking experiments with the linear and cyclic VDAC peptide by our collaborators also showed that TIM910 preferentially interacts with the cyclic peptide (see figure A.4 in the appendix, performed by Tobias Jores). Peak broadening is, like in the case of mitochondrial carriers caused by millisecond exchange dynamics. Due to the low binding affinity of the peptide, we cannot exclude that in this case the exchange is caused by on/off dynamics of the peptide.

We hypothesize that the interaction of TIM910 with its substrate consists of a multitude of weak interaction that in total provide a stable binding that can be easily perturbed when the substrate needs to be handed over to a downstream element (insertion complex), similar to other chaperones like Skp.[18] We therefore assume that shorter substrates, like the VDAC paptides, have a lower affinity than longer substrates, due to a smaller number of interactions. The most important observation in this experiment was the

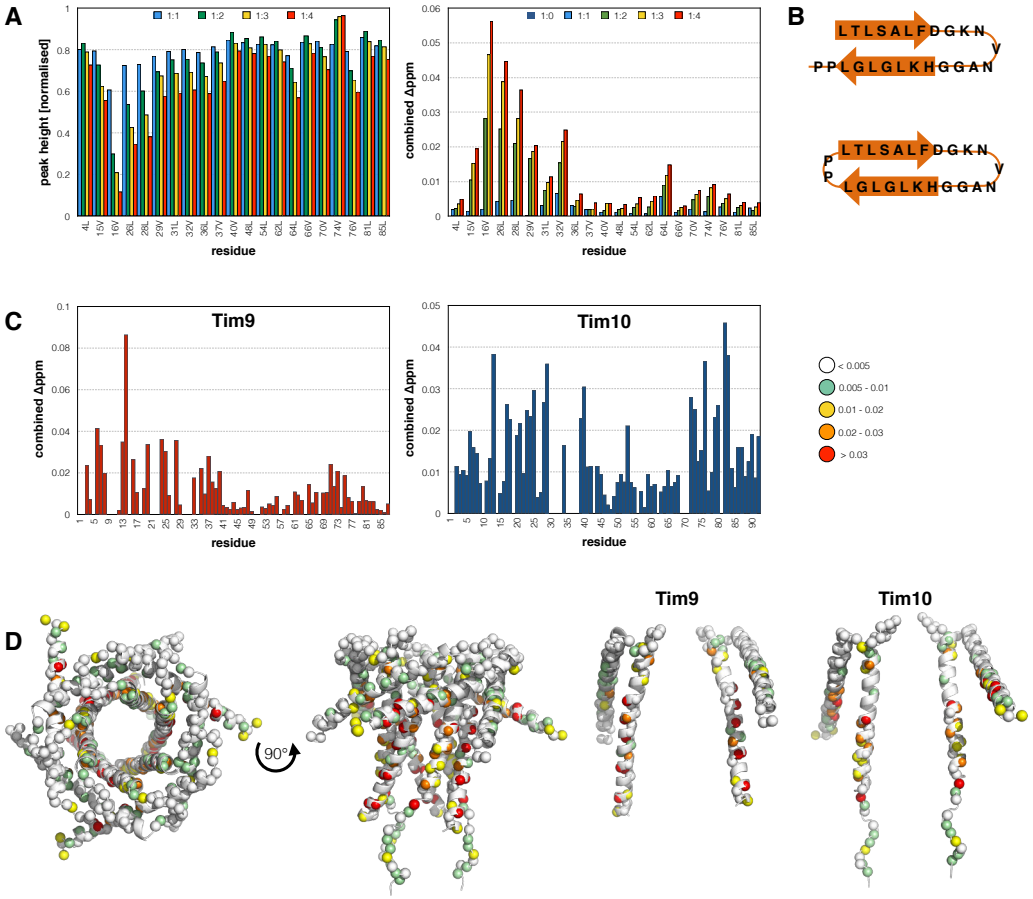


Figure 3.13: **Chemical shift/peak height changes of TIM910 spectra upon binding of a cyclic VDAC peptide. A:** Peak intensity and chemical shift changes of LV methyl sidechains in TIM910 upon addition of cyclic VDAC peptide. 0.15 mM of methyl labelled TIM910 were mixed with cyclic VDAC and NMR spectra recorded at ratios of TIM910:VDAC 1:1 (blue), 1:2 (green), 1:3 (yellow) and 1:4 (red). **B:** Sequence of the linear and cyclic VDAC peptide. An arrow indicates the presence of a β-strand in the membrane inserted form. **C:** Chemical shifts in a ^{1}H-^{15}N correlated TIM910 NMR spectrum upon addition of 1:6 ratio of cyclic VDAC. **D:** Chemical shift values of **C** plotted on a TIM910 model.

fact that TIM910 is able to bind the cyclic peptide (which most likely forms a β-hairpin), but not the linear peptide which is less folded. Hence, TIM910 seems to prefer at least partially folded substrates over unfolded substrates.

The TIM910-bound substrate is not completely unfolded

One interesting point which we did not address so far, is the conformation of the bound substrate protein. We know from NMR experiments that the substrate must be very mobile when it is bound to TIM910 (peak symmetry of substrate-bound TIM910 NMR spectra and effect of the position of the MTSL tag on GGC - see previous sections) and that TIM910 preferentially binds folded substrates (preference of cyclic VDAC peptide over linear). When we recorded CD spectra of apo TIM910, TIM910-GGC and TIM910-AAC samples we could at first not see any difference in secondary structure between the samples at 20°C (figure 3.14A). At 35°C (figure 3.14B) apo TIM910 loses its helical propensity, which is most likely due to a change in populations between monomers and hexamers (which we could also observe by NMR). Interestingly, TIM910/substrate samples do not change their CD profile with increasing temperature, which can either mean that it inhibits the dissociation of TIM910 into monomers or that substrate binding induces a more helical conformation in TIM910. Moreover, the substrate itself might adopt a helical conformation. Since we always see a mixture of the secondary structure of TIM910 and the substrate, we can not determine the isolated conformation of the substrate. Nevertheless, we can observe that CD spectra of TIM910/substrate complexes mostly show α-helical properties and not those of an unfolded protein. Additionally, we tried to perform infrared measurements on ^{13}C labelled GGC and ^{12}C labelled TIM910 complex samples, which could give us separate information about the secondary structure of the bound substrate and TIM910. Unfortunately, the ^{13}C infrared spectra of GGC were extremely broadened and could not be used for analysis. NMR spectra of CDN labelled GGC bound to TIM910 at 35°C only contained about four visible peaks, which might be part of a flexible N- or C-terminus of GGC. Increasing the temperature ameliorated the quality of the GGC spectrum and surprisingly did not lead to dissociation of the complex, even at temperatures way above the melting temperature of TIM910 (T_m∼40°C, temperatures used up to 60°C). After cooling down of the sample the same spectrum could be observed as before heating. Methyl spectra of specifically alanine-labelled GGC bound to TIM910 at 35°C, also gave only extremely broad peaks. A comparison of this spectrum with GGC spectra of unfolded GGC in 8 M Urea and folded GGC in dodecylphosphocholine (DPC) micelles is shown in figure 3.14C. The spectrum of TIM910-bound GGC does neither

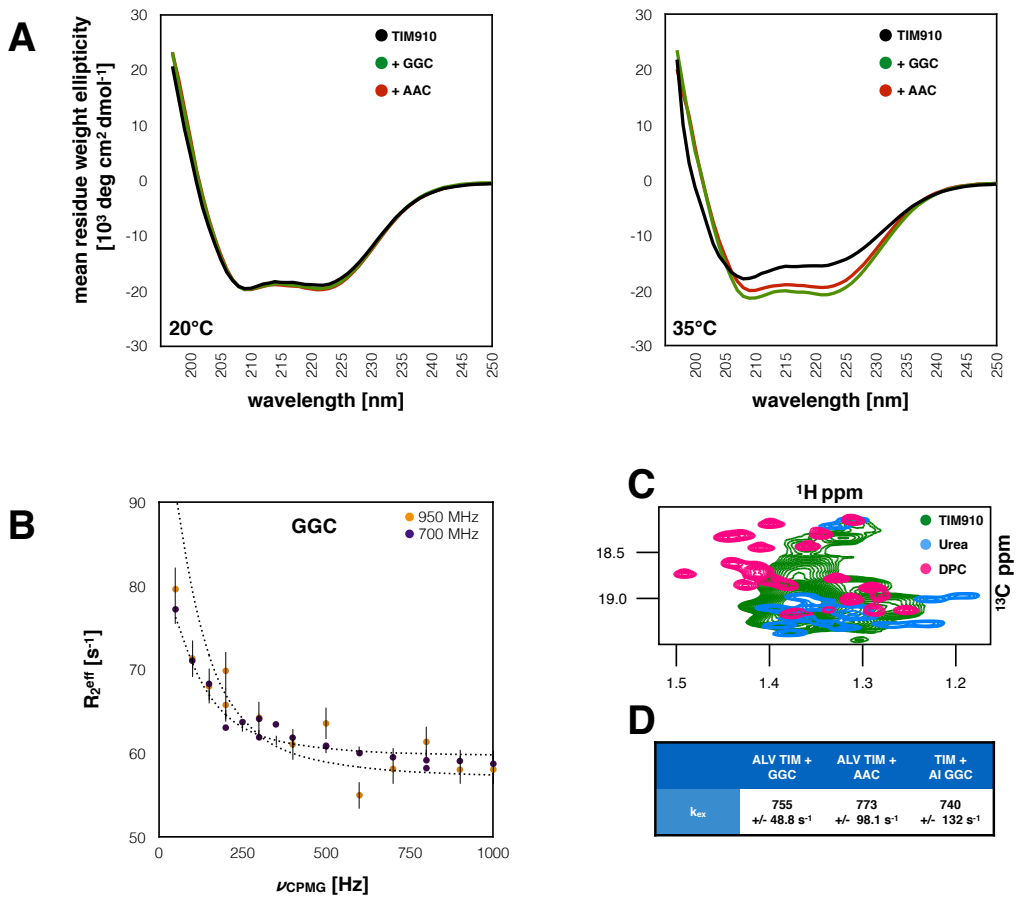


Figure 3.14: **Conformation and dynamics of TIM910 bound substrates. A:** CD spectra of apo TIM910 (black), TIM910-GGC (green) and TIM910-AAC (red) at 20°C and 35°C. All spectra show a pattern characteristic for helical proteins with a dip at 208 nm and 222 nm. While the curves are nearly identical at 20°C, the apo TIM910 curve shows less helical content at 35°C compared to 20°C wheras curves of TIM910-GGC and TIM910-AAC stay the same. **B:** CPMG RD curves of AI methyl labelled GGC bound to TIM910 show non-flat profiles. **C:** NMR spectrum of the alanine region of specifically AI labelled GGC bound to TIM910 (green), in DPC micelles (pink) and in 8 M urea (blue). **D:** Chemex fits of CPMG data of specifically ALV labelled TIM910 bound to GGC and AAC, and specifically labelled AI GGC bound to TIM910. Populations were fixed to 50 % in all fits.

resemble unfolded nor folded GGC and seems to be a mixture of different conformations. Although we were not able to assign any GGC methyl peaks, we performed CPMG RD experiments on unassigned peaks to determine if this peak broadening is due to millisecond exchange dynamics. Indeed, also TIM910-bound GGC experiences millisecond exchange which leads to a poor quality of NMR spectra and is most likely due to the dynamics of the bound substrate. Fitting the CPMG RD data of TIM910 labelled TIM910-GGC and TIM910-AAC as well as GGC labelled TIM910-GGC revealed that the exchange process in all three samples occurs on the same timescale (see figure 3.14D, for the fitting parameters refer to Materials and Methods). This observation leads us to the conclusion that this process is caused by a commonly observed phenomenon, which we tentatively ascribe to a movement of GGC in the binding site. The fact that we only observe one set of peaks in substrate-bound TIM910 supports this hypothesis.

3.2.2 Mutations in the binding site lead to growth and import defects *in vivo*

If we compare the substrate binding sites of TIM910 determined by different techniques and with different substrates, we find that all data converges on a binding site in a cleft between the outer and inner helix on the exterior of the protein. The most affected methyl residues lie in the area of residues 20 and 30 in the N-terminal helix and between 60 and 70 in the C-terminal helix (see figure 3.15). Based on the X-ray structure of TIM910 we selected residues that would confer hydrophobicity to this binding site, necessary for hydrophobic interactions with the membrane protein substrate, and that would not perturb the stability of the complex if mutated. We chose seven residues of Tim9 and Tim10 each to be mutated and tested in an *in vivo* experiment. *In vivo* yeast viability and import assays were performed by our collaborators of the Wiedemann group in Freiburg. Growth defects caused by these mutations are shown in figure 3.16D. We chose four Tim10 mutants that were not viable to test 1) if the lethality of these mutations is caused by disassembly of the TIM910 hexamer and 2) if these mutants are still able to bind substrates in an *in vitro* assay. We introduced point mutations in our TIM910 pETDuet construct and expressed and purified these mutants according to our established protocol. SEC profiles and SDS PAGE analysis show that all mutants are still able to form a hexameric complex (figure 3.16A). To check the ability of these mutants to bind substrates *in vitro* we took advantage of our TIM910/substrate complex reconstitution protocol (explained in section 3.2.1 and in the chapter Materials and Meth-

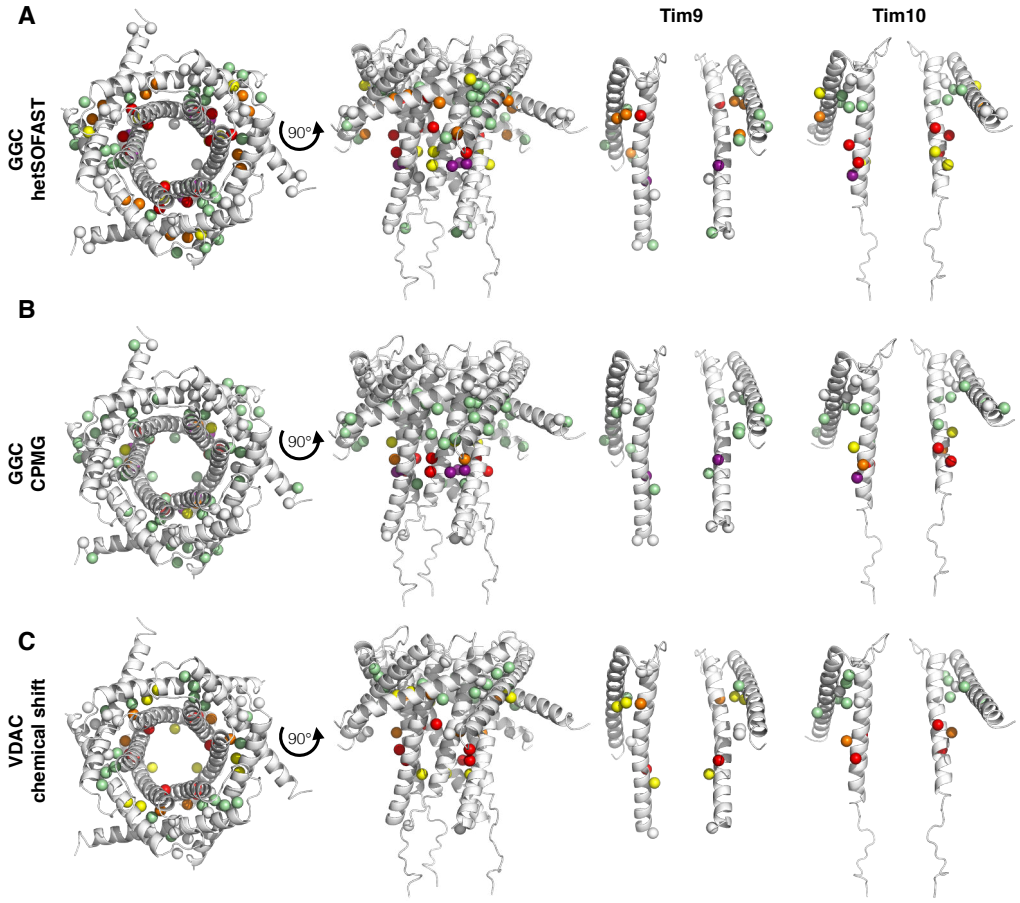
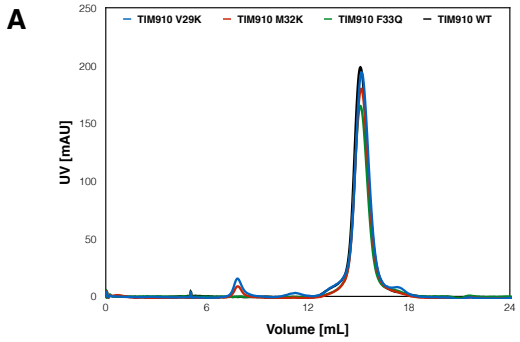


Figure 3.15: **Summary of the TIM910 substrate binding site.** The substrate binding site is conserved for outer (**C**) and inner (**A/B**) membrane substrates. In all three measurements residues in the same area of TIM910 are affected. This region is located between the N- and C-terminal helix on the outside of TIM910.



D

Subunit	mutation	lethal	growth defect	normal growth
Tim9	M21S	x		
	F24K		x	
	L27K		x	
	Y28Q	x		
	V70S			x
	F74K	x		
control	L48S			x
Tim10	L26K	x		
	V29K	x		
	M32K	x		
	F33Q		x	
	V76S			x
	M80K	x		
control	L54S			x

Figure 3.16: **Single point mutations in the TIM910 substrate binding site cause lethality in yeast.** **A:** Superdex200 10/300 GL size exclusion profiles of TIM910 wildtype and binding site mutants. Wildtype TIM910 is shown in black, $TIM910_{V29K}$ in blue, $TIM910_{M32K}$ in red and $TIM910_{F33Q}$ in green. All TIM910 complexes eluted as homogeneous hexamers, thus mutation of the binding site does not hamper hexamer formation. **B/C:** SDS PAGE of elution (E), wash (W) and aggregates (A) fractions of GGC refolding experiments with wildtype TIM910, TIM910textsubscriptV29K and TIM910textsubscriptM32K (**B**) or TIM910textsubscriptF33Q and TIM910textsubscriptF33Q/F70S (**C**). For better visual clarity the amount of loaded sample is adjusted (the elution fraction is twice as concentrated as the wash and aggregates fraction). In WT TIM910 samples GGC (34 kDa) can be found in the elution and aggregated fraction. In all TIM910 binding site mutants no GGC can be seen in the elution fraction, instead all GGC remains as aggregates on the column. The presence of TIM910 (10 kDa) alone in the elution fraction also indicates that mutant TIM910 is retained on the column but the interaction is not strong enough to permit elution of a stable TIM910-GGC complex. **D:** Table of tested mutation sites in Tim9 and Tim10 subunits and their *in vivo* effect on yeast cells lacking TIM813 and the respective subunit. Lethal muations are coloured black, mutations that lead to a growth defect in yellow and mutations with no growth phenotype in green. Control mutations are located in the loop between the two CX_3C motifs.

ods). Simultaneously, we performed reconstitution experiments with wild type TIM910 as a reference. As you can see in figure 3.16B/C, only wild type TIM910 is able to rescue GGC from aggregating on the column. The presence of TIM910 in the elution fraction of mutant samples, could indicate that mutant TIM910 is able to weakly interact with GGC, thus is retained on the column, but the interaction is not strong enough to allow elution of a stable chaperone/substrate complex.

3.2.3 SAXS provides a structural model of a TIM910/ substrate complex

Now that we confirmed the NMR-derived substrate binding site with *in vivo* experiments we wanted to know how a complex between TIM910 and a substrate protein could look like. To this end we relied on SAXS data we measured as described in section 3.2.1. As already mentioned before the DAMMIF SAXS envelope of a TIM910-GGC complex shows a 2:1 stoichiometry between chaperone and substrate. The cavities of the two TIM910 hexamers are clearly visible and thus in agreement with our hypothesis that the substrate is bound on the exterior of the chaperone and not in the central cavity (see figure 3.8E). We took advantage of the two TIM910 cavities to position two TIM910 hexamer models (including the X-ray-invisible N- and C-termini). Then we fitted the GGC amino acid chain (split in eight pieces, due to technical requirements of the program) with the CORAL software onto the two TIM910 hexamers to fulfill the measured scattering curve of the TIM910-GGC complex. Additionally, we rotated the two TIM910 hexamers with respect to each other to find their optimal position. Surprisingly, rotation of the TIM910 hexamers does not substantially alter the fit quality. Only a "head to head" or "tail to tail" orientation of the two TIM910 hexamers, where either the connecting loops or the N-termini face each other, does not fit the experimental parameters.

For each orientation of the TIM910 hexamers five GGC models are calculated. In all models that fit the experimental data, GGC is positioned on the exterior of TIM910 and wrapped around the two hexamers (see figure 3.17A-F, fitting and analysis performed by Martha Brennich). This structural model, which does not use any prior knowledge about the TIM910 substrate binding site, is hence in perfect agreement with the rest of our data.

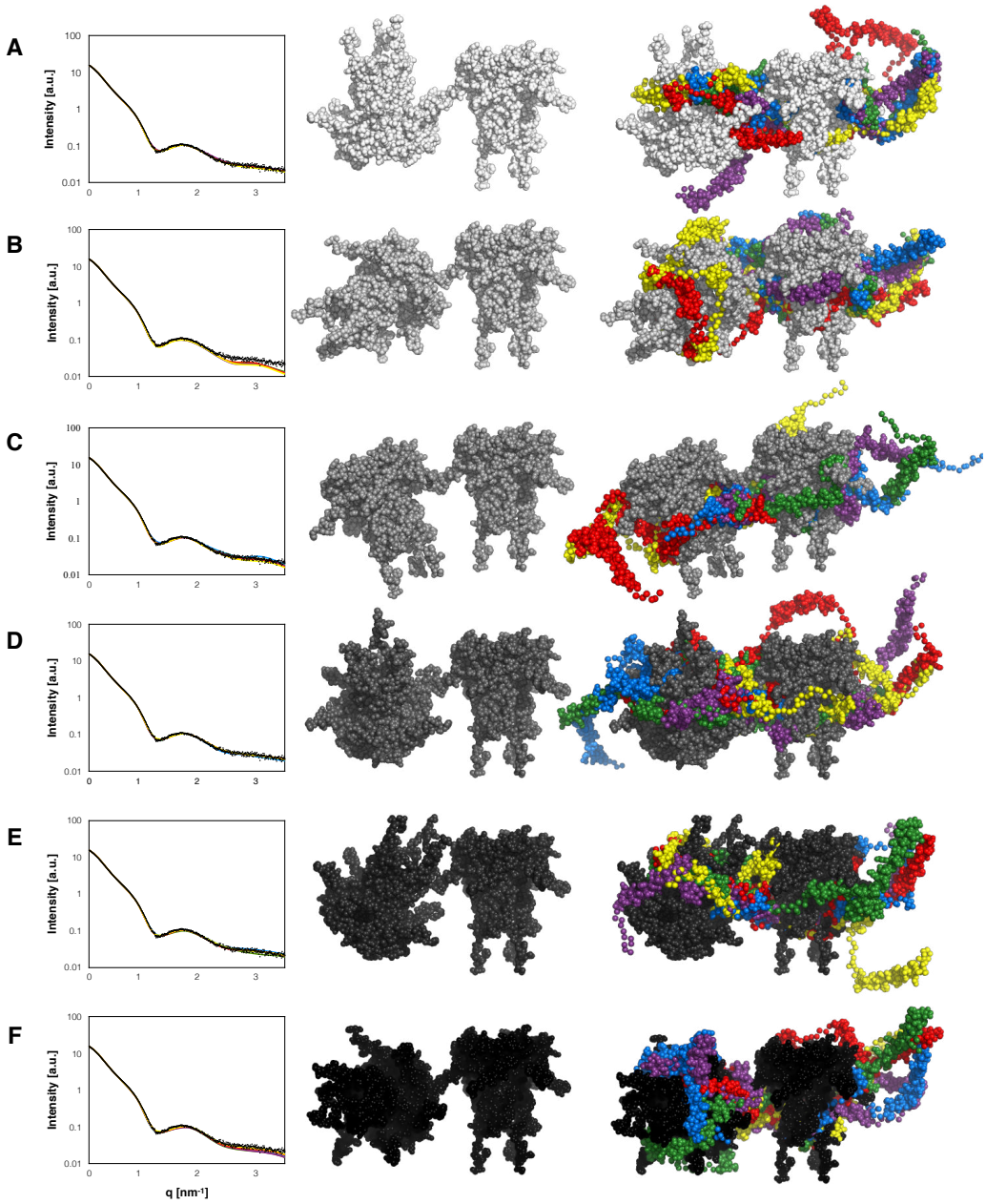


Figure 3.17: **A structural SAXS model of the TIM910-GGC complex.** SAXS scattering curve, TIM910 relative orientation and GGC models on fixed TIM910 molecules are shown for six different TIM910 orientations (figure **A**, **B**, **C**, **D**, **E** and **F**). On the left: scattering curves of TIM910-GGC complexes. Experimental data is shown as black dots, calculated TIM910-GGC models as coloured lines. In the middle: the relative orientation of the two TIM910 molecules in the envelope that are kept fixed during the model calculation. On the right: five different models of TIM910-GGC, whereby GGC is shown in colour and TIM910 in greyscale. In all models GGC is wrapped around the two TIM910 molecules and the two cavities of the TIM910 complex are being kept intact.

3.3 Tim12

3.3.1 Tim12 is an important partner for substrate handover at the inner membrane

One intriguing question in TIM910-dependent mitochondrial import is how the substrate is handed over at the insertion machinery of the mitochondrial membrane. Accumulation of substrate in the IMS in Tim12 defective mitochondria, suggests that Tim12 plays an important role in membrane protein insertion into the inner membrane. Tim12 is always found in a complex with TIM910 when associated to the TIM22 import complex.[38] Our aim was thus to study the TIM91012 complex and to obtain insights about how TIM91012 interacts with substrate proteins.

Tim12 can only fold in the presence of TIM910 *in vitro*

Tim12, like the other small Tim proteins, contains a twin CX_3C motif and according to prediction softwares adopts a similar fold as Tim9 or Tim10. Additionally, Tim12 contains a longer C-terminus, which is predicted to be helical. Lionaki et al showed that this additional helix is important for TIM91012 complex formation and lipid binding.[74] The authors suggest that Tim12 is tethered to the inner mitochondrial membrane by interacting with lipids and in particular cardiolipin, a lipid preferentially found in the inner membrane. In contrast to Tim9 and Tim10, yeast Tim12 contains two additional cysteines which are not conserved in other species (see supplementary figure A.6). To simplify expression and purification we thus decided to mutate these cysteines to alanines in one of our constructs.

Isolated Tim12 has a distinct dynamic behaviour from hexameric TIM91012

When we expressed wild type Tim12 and Tim12 without the additional cysteines, we only obtained protein in the insoluble fraction. Also expression in SHuffle cells, that should favor correct disulfide bond formation, did not render Tim12 soluble. We purified these two constructs under denaturing conditions and performed SDS PAGE without reducing agents with the purified protein samples. For both constructs the majority of protein formed disulfide oligomers, that were sensitive to reducing agent. This behaviour was even more pronounced in the wild type protein. We decided therefore to continue all our experiments using the Tim12 construct without the additional cysteines. Henceforth, I use the name Tim12 for this construct.

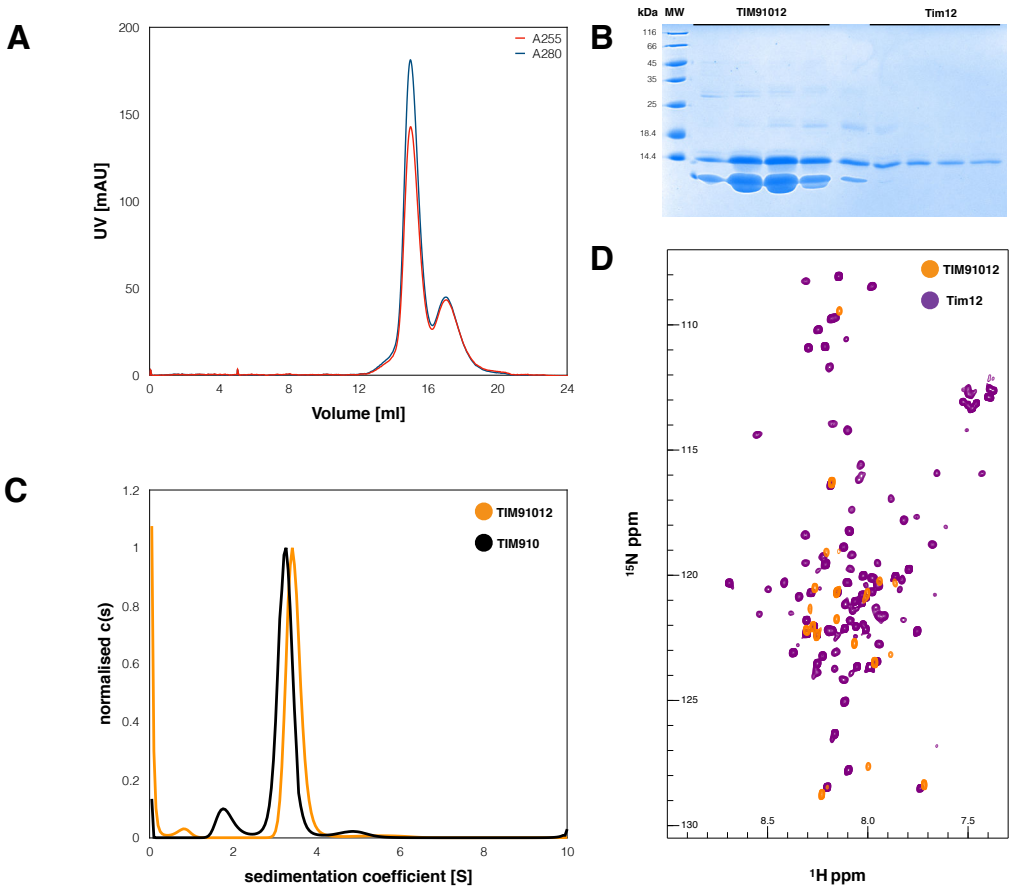


Figure 3.18: **Purification of the TIM91012 complex. A:** Gel filtration profile of TIM91012 on a Superdex200 10/300 GL column. The first peak at 16 mL corresponds to a hexamer of Tim9, Tim10 and Tim12 subunits. The second peak at 18 mL to free Tim12. **B:** Non-reducing SDS PAGE of the same gel filtration experiment. In the TIM91012 peak bands for Tim9 (10 kDa), Tim10 (10 kDa) and Tim12 (12 kDa) can be found. In the Tim12 monomer peak only one band corresponding to the molecular weight of Tim12 is present. **C:** AUC sedimentation profile of TIM910 (black) and TIM91012 (orange). The sedimentation profile of TIM910 is explained in figure 3.4A. The sedimentation profile of TIM91012 shows a main peak at 3.4 S, contributing to 91% of the total signal intensity and a minor peak at 0.9 S making up 4% of the total signal intensity. The estimated molecular weight of the species in the main peak is 65 kDa corresponding to a hexamer of Tim9, Tim10 and Tim12 subunits. The stoichiometry of the individual subunits cannot be determined with this method. The minor contribution in the TIM91012 profile could be a Tim9, Tim10 or Tim12 monomer with a molecular weight of about 8.5 +/-2 kDa. **D:** ^{1}H-^{15}N correlated NMR spectra of monomeric Tim12 (purple) and Tim12 in a TIM91012 hexamer (orange). In the hexamer only Tim12 is isotope labelled. While nearly all peaks are NMR visible in monomeric Tim12, in TIM91012 only flexible regions can be seen. A representation of Tim12 secondary structure can be found in figure 3.19.

Dialysis of denatured Tim12 against buffer (Tris-HCl pH 7.4, see chapter Materials and Methods) did not yield any soluble protein. We obtained the same result for flash dilution, titration of the protein into buffer or refolding attached to a NiNTA column. Additionally, we tried oxidative refolding by opening the cysteine bonds with reducing agent in guanidinium-HCl and subsequent exchange of buffer. Although we could employ this method successfully to Tim9 and Tim10 we were not able to obtain soluble protein for Tim12. Considering the use of TIM910 to transfer Tim12 *in vivo* from the MIA40 complex to the TIM22 complex, we tried to refold Tim12 in the presence of TIM910 by dialysis. Surprisingly we were able to obtain a small amount of soluble Tim12. We could optimize this amount by using a NiNTA based refolding protocol, whereby yields were best when Tim12 was incubated after removal of guanidinium-HCl with TIM910 for 30-60 min. The resulting sample contained two species 1) a hexameric TIM91012 complex and 2) a monomeric/dimeric Tim12 complex. A gel filtration profile after elution from the NiNTA column and cleavage of the HisTag of Tim12 is shown in figure 3.18A. The corresponding non reducing SDS PAGE in figure 3.18B. When TIM910 was added to the refolding process before the denaturant was removed the sample contained 1) a hexameric TIM91012 complex and 2) a dimer of Tim12 and Tim9. The reason for this behaviour might be dissociation of the TIM910 complex in the presence of denaturant and a preferential interaction of Tim9 with Tim12 over Tim10 for Tim12. Since we were interested in obtaining TIM91012 hexameric samples and Tim12 alone we opted for the first purification protocol. Any attempts of reconstituting hexameric TIM91012 from dimeric TIM912 via addition of Tim10 failed and resulted in the formation of hexameric TIM910 and free Tim12. A higher affinity of Tim9 for Tim10 than for Tim12 might be the reason for this behaviour. We used AUC to confirm that TIM91012 indeed forms a stable and homogeneous hexameric complex (see figure 3.18C).

NMR spectra of Tim12 in complex with TIM910 and free Tim12 have a completely different appearance. While free Tim12 gives well resolved ^{1}H-^{15}N correlated NMR spectra with nearly all 109 expected peaks visible, Tim12 in complex with TIM910 suffers from strong peak broadening (figure 3.18D). Free Tim12 seems to behave similarily to Tim9 and Tim10, whereby oxidized Tim12 (figure 3.19A) contains two α-helices and flexible N- and C-termini and reduced Tim12 (figure 3.19B/D) has no secondary structure elements. In contrast to Tim9 and Tim10, Tim12 has no exchange dynamics, thus all regions of Tim12 are NMR visible. The only visible regions of Tim12 in the TIM91012 complex are the N- and C-termini (figure 3.19C). Different from secondary structure predictions we found that the long C-terminus of Tim12 is entirely unfolded in the TIM91012 complex. Furthermore, we could not

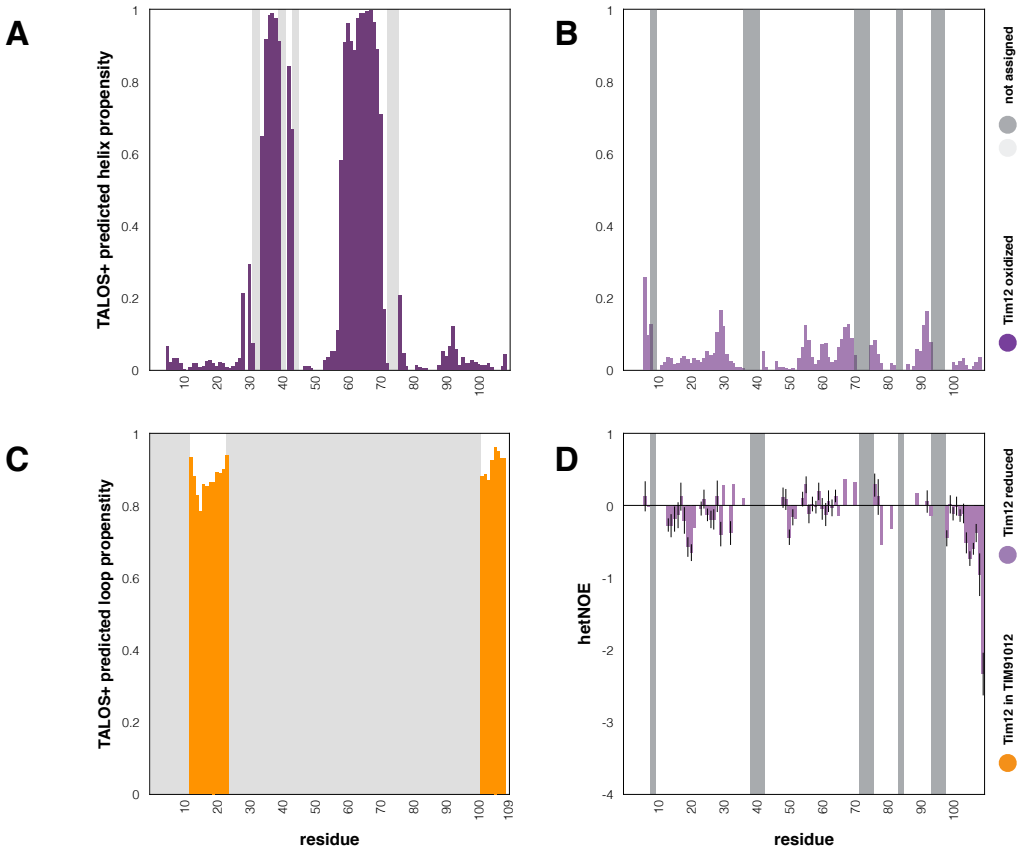


Figure 3.19: **Secondary structure of isolated Tim12 and Tim12 in complex with TIM910. A:** TALOS+ predicted helix propensity of Tim12 with oxidized twin CX_3C motif. In light grey part of the sequence that cannot be seen in NMR spectra or residues that were not assigned. **B:** TALOS+ predicted helix propensity of Tim12 with reduced twin CX_3C motif. **C:** TALOS+ predicted loop propensity of assigned residues of Tim12 in a TIM91012 complex. NMR invisible or unassigned parts are shown in grey. As can be seen in the graph, only flexible loop residues of the N- and C-terminus can be detected by NMR. **D:** hetNOE of Tim12 with reduced twin CX_3C motif. NMR invisible or unassigned residues are shown in light grey.

detect any lipid binding of Tim12 to liposomes or nanodiscs, neither in the free form nor in the TIM91012 complex, even when the lipid mixture was enriched in cardiolipin (data from Audrey Hessel, not shown).

Structural and biochemical data on Tim12 in the literature is rather limited, thus several important questions remain to be answered. One of them is to determine the stoichiometry of the TIM91012 complex. Western blot based estimations of the ratio between Tim9, Tim10 and Tim12 are not always coherent between experiments and the true stoichiometry of the complex is so far unknown. We tried to answer this question by using an NMR based approach. ^{1}H-^{15}N correlated NMR spectra of TIM91012 labelled on Tim9 and Tim10 show peaks that are either doubled or tripled, which indicates a breaking of the symmetry of the TIM910 complex. We know from native mass spectrometry experiments that the TIM91012 complex contains only one Tim12 subunit. The number of Tim9 and Tim10 subunits could not be determined due to the similarity in molecular weight of the two proteins (data not shown, performed by Elisabetta Boeri-Erba). We hypothesized that if Tim12 substitutes either one Tim9 or one Tim10 subunit, the neighbouring subunits must experience a change in chemical environment and thus be shifted in the NMR spectrum. We assume therefore that if Tim9 is substituted, Tim10 must be perturbed and the affected peaks would lose their symmetry or the other way around (see figure 4.3 for a schematic representation). In the case of our NMR spectrum we find that almost all peaks lose their symmetry, whereby Tim10 peaks seem to be more affected than Tim9 peaks (see figure 3.20A/B), which would support exchange of one Tim9 subunit by one Tim12 subunit. This is still work in progress and we hope to answer this question in the future by using two different labelling schemes and NOE based experiments.

Crystallisation of TIM91012

As I mentioned before no structural data on Tim12 is available, neither in the free nor in the TIM91012 complex form. We assumed that solving the X-ray structure of apo TIM91012 could be feasible and would eventually answer the question about the stoichiometry of the complex. We considered TIM91012 to be the physiologically relevant form and thus only started crystallisation trials with the complex and not with the free subunit. Additionally, the large degree of unfolded regions would likely impair crystallisation of Tim12. From our NMR experiments we knew that Tim12 has a long unfolded C-terminus which could hamper crystallisation of the complex. Therefore, we tried to use a construct without this additional segment. Unfortunately, Tim12 is not able to form a complex with Tim9 and Tim10 without this C-terminal

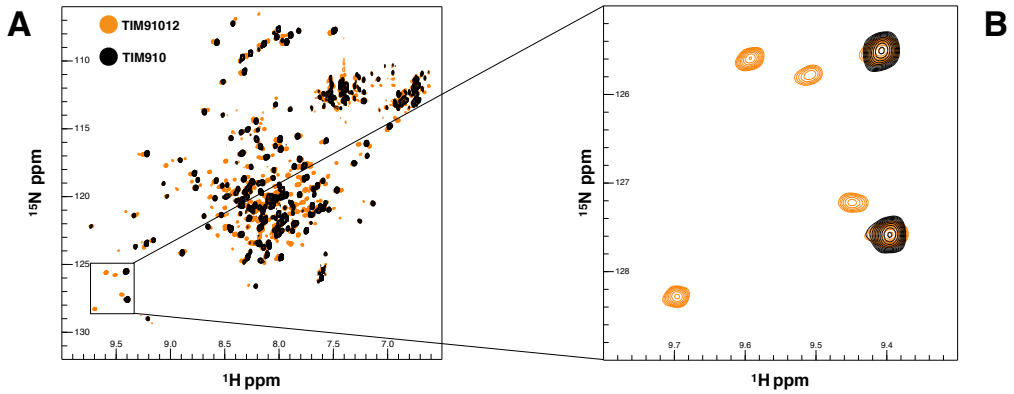


Figure 3.20: **Loss of symmetry in TIM91012 NMR samples A:** ^{1}H-^{15}N correlated NMR spectra of TIM910 (black) and TIM91012 (orange). In both cases only TIM910 are isotope labelled and Tim12 is unlabelled. **B:** Close up of **A**. Two isolated residues belonging to the Tim10 subunit are shown. While in the TIM910 spectrum only a single set of peaks can be observed, those peaks are split into three in the TIM91012 spectrum, indicating a break of symmetry of the system.

sequence. Thus, we set up crystallisation trials with TIM91012 including this flexible part. A first high throughput screening of crystallisation conditions revealed three conditions where TIM91012 was able to crystalize (see Materials and Methods for conditions, images of the crystals in figure 3.21). Condition 1 gave already diffraction of 2 Å. Subsequently we tried to reproduce and optimize these conditions to obtain bigger or better diffracting crystals. Unfortunately, we were not able to reproduce crystals of condition 1 and 3. Crystals from condition 2 were reproducible but only resulted in needle like crystals that did not diffract. Addition of chromium chloride to condition 2 changed the shape of these crystals from needle to a star shaped form. Nevertheless, these crystals were so fragile that they could not be harvested. A major problem in these crystallisation trials might be the high flexibility of Tim12, which on the one hand has a long unfolded C-terminus and likely has some exchange dynamics in the bound form which manifests itself in peak broadening of NMR spectra.

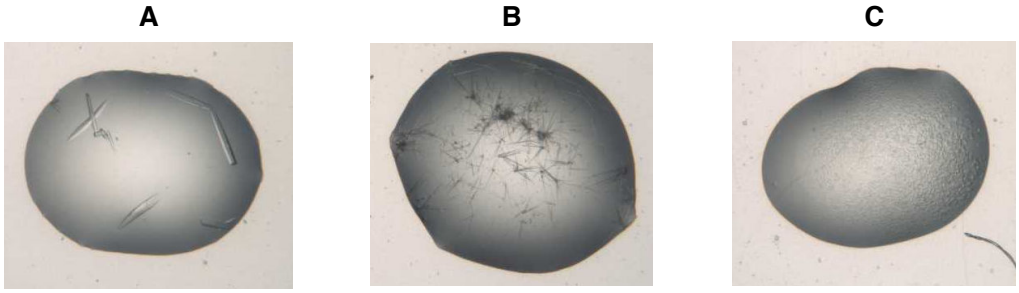


Figure 3.21: **Positive hits of a crystallisation screen for TIM91012. A:** Crystals obtained from condition 1 (components see chapter Materials and Methods). These crystals of TIM91012 were the only ones obtained that diffracted. **B:** Crystals from condition 2. Crystallisation was reproducible, but only needle or star shaped crystals could be obtained. **C:** Crystals from condition 3 were not reproducible.

3.4 TIM910 subunit exchange

3.4.1 Tim subunits can freely exchange between apo and substrate-bound TIM910

How substrate loaded TIM910 can be tethered to the inner mitochondrial membrane is still poorly understood. Two possible scenarios would be either a handover of the substrate from soluble TIM910 to membrane attached TIM91012 complex, or an exchange of Tim12 between substrate loaded and free TIM910.

We investigated here the ability of individual Tim subunits to exchange between substrate-bound and apo TIM910 complexes and found that the rate of exchange is nearly the same in both cases. We tested the subunit exchange by measuring the build-up or loss of complex or monomer specific peaks in a series of short NMR experiments after mixing of free subunit with TIM910 or substrate-bound TIM910. The resulting fitting curves and rates can be found in figure 3.22. Although the *in vitro* measured rates are too slow to explain an exchange process *in vivo* where import of inner membrane proteins occurs on a scale of minutes (correspondance with Nils Wiedemann), we cannot exclude that the discrepancy between our *in vitro* data and *in vivo* measurements are also due to the different experimental conditions and the presence of additional binding partners in the cell. Interestingly, Tim9 and Tim10 are able to integrate in TIM910-GGC complexes while Tim12 fails to do so. A possible explanation might be that Tim12 alone is only found at the MIA complex and a soluble TIM91012 complex only during transport between MIA and TIM22. Thus, an exchange with one of these

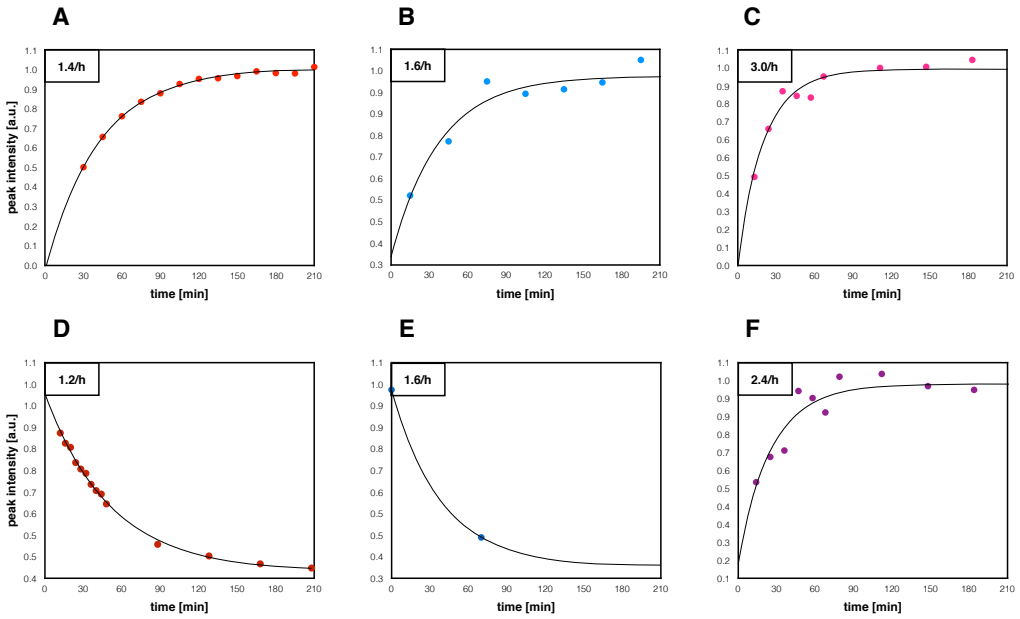


Figure 3.22: **Subunit exchange in apo and substrate-bound TIM910.** Subunit exchange in the different TIM complexes measured by real-time NMR. Immediately after mixing of unlabelled TIM910 or TIM910-GGC with isotope labelled Tim monomer or TIM910, a series of ^{1}H-^{15}N correlated NMR spectra is measured. The graphs show the build-up or decrease of signal intensity of the observed state. Rate constants are given in the upper left of each graph. **A:** Incorporation of isotope-labelled Tim9 into TIM910. **B:** Incorporation of isotope-labelled Tim10 into TIM910. **C:** Incorporation of isotope-labelled Tim12 into TIM910. **D:** Incorporation of isotope-labelled Tim9 into TIM910-GGC. **E:** Incorporation of isotope-labelled Tim10 into TIM910-GGC. **F:** Incorporation of isotope-labelled TIM910 into TIM91012.

intermediate states in the cell, would not lead to membrane insertion of the substrate and thus might be prohibited. Possibly exchange of the substrate between TIM910 and TIM91012 or exchange of Tim12 between TIM91012 and TIM910-GGC needs an additional factor like TIM22 associated Tim54 or a lipid environment. To this end we tried to express the N-terminal domain of Tim54, but were not able to obtain soluble protein. If subunit exchange is the underlying principle of substrate handover at the inner membrane remains to be answered. In any case exchange of free subunits with the hexameric complex seems to be an inherent characteristic of this system, which can be also appreciated from the constant equilibrium between monomeric and hexameric species in TIM910 as seen by NMR (see section 3.1.2).

Chapter 4

Discussion

4.1 Discussion

4.1.1 Structure of Tim monomers

The behaviour of the isolated subunits Tim9 and Tim10 is in full agreement with published data. Tim9 forms preferably dimers and Tim10 monomers, dimers and trimers (according to gel filtration data). Both subunits are completely unfolded when the two disulfide bonds of the CX_3C motif are broken and become partially folded as soon as the disulfide bridges are formed, making the CX_3C motif the sole determinant for their structure. A fully folded rigid state is only reached upon formation of the hexameric complex. Subunits with reduced CX_3C motif are not able to form a hexameric complex. Interestingly, residues around the CX_3C motif are not visible in NMR spectra of oxidized Tim9 and Tim10, whereby this effect is more pronounced in Tim9 than in Tim10. Residues close to this invisible region show millisecond dynamics, seen as non-flat CPMG RD dispersion curves. The detection of millisecond motions in these flanking regions suggests that the the non-observable residues become broadened by the same conformational exchange process. We could not determine the nature of this exchange process. Possible scenarios would be exchange between different oligomeric species, folding/unfolding or opening and closing of the disulfide bridges. However, the CX_3C motif is considered to be stable, making an opening of the disulfide bridges rather unlikely.

4.1.2 The hexameric TIM910 complex

Like the individual subunits, also the TIM910 hexameric complex behaves as expected from the X-ray structure. Helices are formed in the same regions as in the X-ray structure (TALOS+ analysis, not shown), the loop connecting the two CX_3C motifs is rather rigid and the N- and C-terminal parts invisible in the X-ray structure are unfolded and flexible. The most structured part of the complex is the middle region including the CX_3C motif, the connecting loop and the helical part adjacent to the CX_3C motif. The closer one gets to the termini the more flexible the helix becomes. A phenomenon not detected previously is the simultaneous existence of hexameric and monomeric species. Although the main species is always the hexamer, a small population of monomers is always present and cannot be removed by gel filtration. This is likely based on an equilibrium between hexamer and monomer, which manifests itself also in a slight asymmetry of the gel filtration peak towards smaller species. In NMR spectra this equilibrium can be seen as a doubling of the last 5-10 N-terminal and C-terminal peaks. This peak doubling is not

due to the presence of a second conformation of the complex but evoked by a small sub population of monomers. We could demonstrate this by performing 2D DOSY experiments, showing that the minor population is diffusing according to a monomeric protein and the major population according to a hexameric protein. Furthermore, peak positions of the doubled peaks perfectly superimpose with peaks in isolated Tim9 or Tim10 spectra. We will discuss this phenomenon still later on in the section "Subunit exchange".

4.1.3 Substrate binding of TIM910

The key element for our studies was the formation of a complex between TIM910 and a substrate protein. Cell-free experiments in which we expressed GGC in the presence of different amounts of TIM910 show that the chaperone is functional and able to keep its substrates in solution in a concentration dependent manner. In principle we could have used this technique to produce TIM910/substrate complexes by purifying the resulting complex after cell-free expression. The high cost of protein production in cell-free in relation with the high demand in protein for most structural biology techniques was the main driving force for finding a better way of producing these samples. To optimise our yields we decided to produce the substrate in inclusion bodies and refold it in the presence of the TIM910 chaperone. The main difficulty in making these samples is the high aggregation probability of the substrate. In all trials that included refolding by dialysis or flash dilution, the membrane protein substrate aggregated before TIM910 was able to bind, even at very low substrate and high chaperone concentrations. The cause of this problem lies in the nature of the TIM910 chaperone. Although TIM910 is very stable and can be easily refolded when the disulfide bridges are formed, the integrity of the hexamer is easily influenced by denaturants. Even small amounts of denaturants (above 50 mM of guanidine-HCl) are able to dissociate the hexamer into individual subunits. As our study shows, the hexamer, rather than the monomer, is the active species, no substrates can be bound unless the denaturant concentration is below the treshold where hexamer formation is permitted. The treshold for aggregation will be for most substrates way higher than the 50 mM guanidine-HCl necessary for TIM910 to reform a hexamer. Thus, the main objective for developing the method we used for complex formation was the prevention of aggregation of the substrate, while reducing the denaturant to a concentration where hexamer formation of TIM910 is possible. To this end we came up with a protocol that included binding of the substrate to a NiNTA column under denaturing conditions and then gradually removing the denaturant in the presence of TIM910. With this setup the aggregation of the substrate can be delayed while TIM910 is

constantly present to bind the substrate as soon as it reaches the treshold for hexamer formation. Additionally, we tested if we could form TIM910-GGC complexes from GGC in detergent micelles (DPC), since TIM910 shows high stability towards DPC. Surprisingly, complexes could be formed also only by refolding on a NiNTA column and not by dialysis or flash dilution and the resulting yield was similar to the one in denaturant. Therefore, we decided to continue using denaturant as a more cost effective method.

Although we did not succeed in perfoming size exclusion chromatography on TIM910/substrate complexes, we analysed the samples by AUC and found that they are stable and homogeneous. We do not have an explanation why an otherwise stable complex cannot withstand a gel filtration run. A possible reason might be that the substrate interacts with the gel filtration resin and thus promotes dissociation of the complex. While the complex always dissociated on Superdex columns, the complex staid intact on Sec5 Agilent columns. Unfortunately, the resolution power of these columns was not enough to separate apo TIM910 from TIM910/substrate complexes and thus useless in determining the homogeneity of samples. The dominating secondary structure in apo TIM910 is the helix, with some short unstructured parts at the N- and C-termini. Accordingly, CD spectra of apo TIM910 show α-helical properties. At 25°C CD spectra of apo TIM910 and TIM910 bound to AAC or GGC are almost identical. Only when spectra are recorded at 35°C, the unstructured part in apo TIM910 increases relative to 25°C, in agreement with NMR data, which show that the population of folded hexamer with respect to only partially folded monomer decreases with increasing temperature until the complex fully unfolds. CD spectra of TIM910 with bound GGC or AAC at 35°C on the other hand show the same characteristics as spectra at 25°C, implying that binding of the substrates stabilizes a more folded conformation of TIM910. Although it is not possible to distinguish the state of the substrate from that of the chaperone by CD, we speculate that the substrate is in at least a partially helical conformation, since the spectrum does not show characteristics for unfolded or β-sheet proteins. Naturally, all CD spectra of TIM910/substrate complexes are dominated by the contribution of TIM910 due to the differences in size of chaperone and substrate and thus assumptions about the secondary structure of the substrate are difficult to make. Attempts to determine the secondary structure of TIM910 and GGC in the same sample by infrared (IR) measurements with differential isotope labelling of the carbons (^{12}C on TIM910 versus ^{13}C on GGC) did not give any conclusive results due to a strong broadening of the GGC signal. Correct isotope labelling has been confirmed by NMR, thus not explaining the low signal observed in IR.

To determine the substrate binding site of TIM910 we produced isotope

labelled samples for NMR, that were either NMR active on the chaperone or the substrate side. ^{1}H-^{15}N correlated NMR spectra of TIM910 bound to unlabelled GGC contained only about half of the peaks that apo TIM910 spectra did, while spectra of GGC bound to unlabelled TIM910 contained less than one tenth of the expected peaks. Although we expected the size of the complex to be a challenge for NMR, that we tried to overcome by deuterating our proteins, the poor quality of our spectra could not only be explained by the size of the protein. Furthermore, not all parts of the complex seemed to be affected by peak broadening in a uniform way. Obtaining residue-resolved information on these samples was impossible and thus we decided to use selective labelling of methyl sidechains to determine the binding site. We chose to label Leucine and Valine residues of TIM910, because of their distribution on the TIM910 X-ray structure and Isoleucine and Alanine residues on GGC because of their non-overlapping frequencies in NMR spectra. Later on we also prepared TIM910 samples that were labelled on Alanine, Leucine and Valine to increase the number of probes for NMR experiments. Methyl spectra of TIM910 bound to GGC, AAC or AAC_{short} provided enough signal to noise to investigate the binding site by NMR. For TIM910 only one set of peaks was observed in both apo and substrate bound spectra, showing that the symmetry of the TIM910 is not broken upon substrate binding. GGC methyl spectra do not show individually resolved peaks, but rather a mixture of broadenend peaks indicating the existence of different conformations or states. Relaxation dispersion measurements revealed that several of the residues in TIM910 and GGC experience exchange on the millisecond timescale. The fact that both binding partners, TIM910 and GGC, undergo these dynamics supports the idea that it is the same process that can be experienced on both sides. This is further corroborated by fitting of the CPMG data showing that the rate of exchange in both proteins is the same. Thus, the reason for the low quality of NMR spectra might be the dynamic binding of the substrate, which can be experienced as millisecond exchange dynamics on the chaperone and the substrate.

The location of residues experiencing this millisecond dynamics on TIM910 was already an indication of where the substrate is bound. If the ΔR_{2eff} of the CPMG dispersion curves is taken as a reference of the effect of substrate binding, the binding site is situated in the cleft formed by the inner and outer helices in TIM910. Furthermore, we used PRE, sPRE and magnetization transfer (HETSOFAST) measurements to validate the location of the binding site. In short, all methods converge on the same substrate binding site, although the precision between the different methods is quite different. The most precise method is the HETSOFAST measurement, where only a small set of residues is perturbed and they coincide well with residues that

experience millisecond dynamics. Attachment of an MTSL tag also succeeds in finding the same site, but the effect is more spread, most likely due to the dynamic binding of the substrate on TIM910. This also becomes evident when the MTSL tag is attached at different positions. If the binding of the substrate was static, the effect between two differentially positioned tags should be different. In our case positioning the tag on the extremity of GGC or in the middle results in very similar binding profiles. According to solvent PRE data also the N- and C-termini are protected in substrate bound TIM910. One reason for this behaviour could be the stoichiometry of the complex, which will be discussed in the next paragraph. This stoichiometry could lead to a partial protection of the termini even though they are not part of the substrate binding site.

Comparison of conserved residues in TIM910 and the substrate binding site determined by NMR, reveals that indeed especially the hydrophobic residues in this part are highly conserved. Moreover, the cleft formed between the inner and outer helix in TIM910 is highly hydrophobic, whereas the inner pore of TIM910 is lined by hydrophylic residues. Thus, conservation and hydrophobicity of the binding site support our hypothesis that TIM910 binds its substrates on the outside in a hydrophobic cleft formed by the N- and C-terminal helix.

One of the main questions in this project was the determination of the structure of a TIM910/substrate complex. We ruled out X-ray crystallography, since we already knew from NMR that these complexes are very dynamic, furthermore the complex is too small to be studied by cryo-EM and too big to be studied by NMR. Hence we had to rely on biochemical data and low resolution structural techniques to get a picture of how this complex looks like. The first question we tried to answer is the stoichiometry of the complex. As has been already noted by Webb et al,[143] the cavity of TIM910 is in comparison to the cavity of its homologue in bacteria, Skp,[18] too small to accomodate its substrates. Therefore, either several TIM910 hexamers are necessary to bind a substrate or the complex has to rearrange upon binding. We showed that TIM910 does not undergo major structural changes upon substrate binding (only minor changes in peaks positions between apo and holo complex and no change in order parameters between bound and free TIM910) and that the substrate length determines the stoichiometry of the complex. Full-length GGC and AAC contain six transmembrane helices and have a molecular weight of around 30 kDa. AAC_{short} is a short version of AAC and contains only transmembrane helices 2 and 3 and has a molecular weight of about 15 kDa. AUC sedimentation profiles of TIM910 bound to full-length GGC are in agreement with a 2:1 complex between TIM910 and GGC, while TIM910 bound to AAC_{short} sediments like a 1:1 complex. We

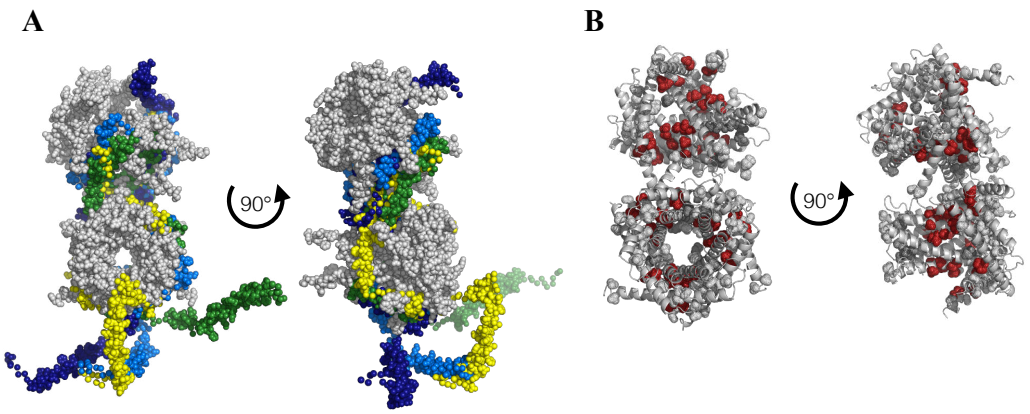


Figure 4.1: **The TIM910 substrate binding site determined by NMR and SAXS.** Comparison between the SAXS model of the TIM910-GGC complex (**A**) and the substrate-binding site of TIM910 determined by NMR (**B**).

obtained the same result using NMR DOSY experiments, TIM910 bound to full-length AAC or GGC diffused like a 2:1 complex, while TIM910 bound to AAC_{short} had a diffusion rate very similar to apo TIM910. Diffusion rates of TIM910 bound to full-length GGC are the same either if they are measured on TIM910 or GGC, underlining the quality of our NMR samples. To be able to build a model of our TIM910-GGC complexes we collected SAXS data using an online NiNTA protocol to avoid aggregation of the substrate and provide high quality SAXS data. Already the SAXS envelope shows a 2:1 stoichiometry, in agreement with AUC, NMR and SDS PAGE data. Furthermore, the two TIM910 molecules in the SAXS envelope still contain free cavities corroborrating our substrate binding model, whereby the central cavity of TIM910 is not occupied by the substrate. To build the SAXS model, two TIM910 molecules are fixed around the cavities of the SAXS envelope while GGC is modelled in the extra density of the envelope. To account for the dynamic nature of the bound GGC and to obtain an ensemble of structures, several GGC models have been generated. All models that fit the original SAXS data have the GGC molecule located outside of the TIM910 chaperone. Positioning the two TIM910 molecules with facing N-termini or facing CX_3 motif did not yield satisfactory fits of the experimental data. Rotation of the two TIM910 molecules with respect to each other on an axis perpendicular to the cavity does not change the quality of the fits, thus any of these positions might represent the measured data. The localization of the GGC chain around TIM910 in the SAXS models is in perfect agreement with the substrate binding site determined by NMR.

We therefore propose that TIM910 binds its substrates in a hydrophobic cleft on the outside of the chaperone in a modular fashion. The number of TIM910 molecules needed to transport a substrate depends on the length of the substrate. This modularity would offer TIM910 the flexibility to interact with a wide range of substrates of variable sizes. This is in agreement with a chaperone that has to interact with substrates that differ markedly in their membrane-inserted structure and size, where molecular weights are ranging from 17 kDa to about 50 kDa.

TIM910 is required, not only to transport membrane protein precursors to the inner mitochondrial membrane, but also to the outer mitochondrial membrane. While all inner-membrane substrates of TIM910 form helices in their membrane-inserted form, the outer-membrane substrates of TIM910 are all β-barrel proteins. We never managed to reconstruct a soluble and stable complex of TIM910 with an outer-membrane protein. There could be several reasons explaining the difficulty of making soluble complexes. First of all, outer-membrane protein transport depends not only on TIM910 but also on TIM813.[41,147] Therefore, correct formation of the complex might need both chaperone complexes. An argument against this hypothesis is the fact that TIM813 is not an essential protein and the functions of TIM813 can be taken over by TIM910 in the cell.[98] Another possible reason why these complexes are more difficult to form is the place of insertion for the different complexes. While complexes of TIM910 with inner-membrane proteins need to cross the intermembrane space, outer-membrane substrates can be directly inserted in the outer membrane and thus may never be in a soluble form in the intermembrane space. The fact that SAM and TOM can form a supercomplex,[95] thus minimizing the distance an outer-membrane substrate would have to be translocated, supports this hypothesis.

In order to increase the solubility of the binding partner and thus still obtain some information about the binding site, we decided to work on short peptides of β-barrel proteins in collaboration with Doron Rapaport from the University of Tübingen. We tested the binding of two peptides of human VDAC: a linear version of the β-signal of VDAC and a cyclic version of this peptide. Jores et al showed that the cyclic VDAC peptide is more structured than the linear version.[55] NMR titration experiments of labelled TIM910 with unlabelled VDAC peptide showed that the linear peptide is unable to bind TIM910 while the cyclic version is able to bind TIM910, albeit with a rather low affinity. Chaperones can bind their substrates by engaging a multitude of weak interactions that in sum result in a high binding affinity.[18] Thus, binding of short fragments can only result in a limited amount of contacts and therefore a low binding affinity. Spectra of TIM910 bound to cyclic VDAC suffer considerably less from peak broadening, making even the

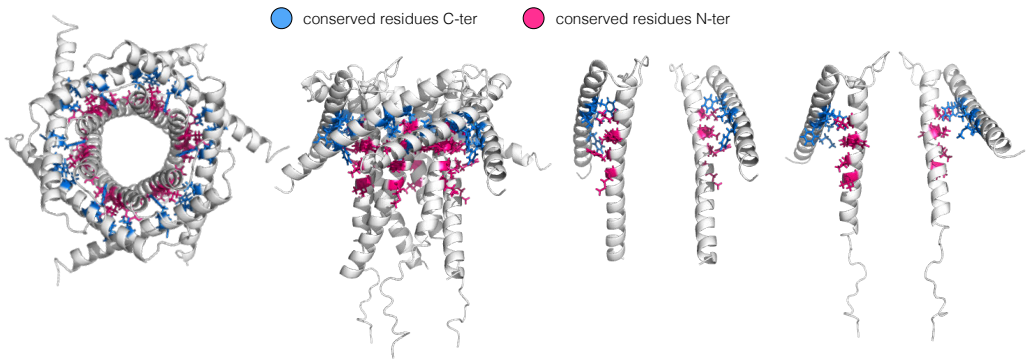


Figure 4.2: **Conserved hydrophobic residues in TIM910.** Conservation between small Tim proteins calculated using the Geneious software. Conserved residues in the N-terminal helix are shown in pink, in the C-terminus in blue. TIM910 most likely uses hydrophobic interactions to bind its client proteins. Residues in the TIM910 substrate binding site are especially conserved.

analysis of backbone amide spectra feasible. The binding site of cyclic VDAC can be easily determined by following peak shifts and intensity changes of methyl and amide peaks of TIM910. The availability of a larger amount of probes in amide spectra provides a better insight into the binding site of TIM910. While in methyl spectra only residues deeper in the binding site are effected, amide spectra show peak shifts even close to the N- and C-termini. Similar to mitochondrial carriers, binding of the cyclic VDAC peptide induces millisecond dynamics in TIM910 in residues of the binding site. Due to the low affinity of the VDAC peptide, this non-flat CPMG RD curves could be also caused by on/off binding dynamic of the peptide. In summary, outer-membrane substrates of TIM910 use the same binding site as inner membrane substrates despite their differences in structure in the membrane-inserted form. Furthermore, TIM910 preferentially binds substrates that contain already some secondary structure, as can be seen by comparing the behaviour of cyclic and linear VDAC peptide.

Our final goal was to confirm the binding site by performing *in vivo* import experiments in yeast. To this end we started a collaboration with Nils Wiedemann from the University of Freiburg. To validate the binding site we chose residues that, according to our NMR data, would be crucial for substrate binding and mutated these hydrophobic amino acids into hydrophylic amino acids. A list of these residues with the matching point mutation and their effect on yeast growth can be found in figure 3.16. Several point mutations already resulted in non-viable yeast cells ($Tim9_{M21S}$, $Tim9_{Y28Q}$, $Tim9_{F74K}$, $Tim10_{L26K}$, $Tim10_{V29K}$, $Tim10_{M32K}$, $Tim10_{M80K}$). Yeast cells that

suffered from a growth defect were typically deficient in mitochondrial carrier (AAC2) and outer-membrane protein (Tom40) import. One example ($Tim10_{F33Q}$) is shown in the appendix in figure A.5 (experiments performed by Caroline Lindau). Control mutations in the connecting loop of TIM910, between N- and C-terminal helix, did have no effect on yeast growth. To confirm that lethal mutations do not corrupt hexamer stability, we expressed these mutants *in vitro* and performed gel filtration runs. A selection of these runs compared with wild type TIM910 can be seen in figure 3.16A. None of the mutations caused the dissociation of the hexamer, thus the reason for yeast lethality is not a compromised structure of TIM910. Furthermore, we conducted substrate/chaperone complex reconstitution assays with these mutants and wild type TIM910. We could show that, while wild type TIM910 is able to hold the substrate in solution, none of the TIM910 mutants is able to bind the substrate. We conclude that lethality is caused by the inability of these point mutants to interact with the substrate and not because of a loss of TIM910 complex integrity. To summarize, single point mutations in the substrate binding site can cause either lethality or import defects in yeast. This is rather unexpected, since we speculated that the substrate interacts with TIM910 in a multitude of weak interactions. Thus, mutation of a single residue should not perturb the binding significantly.

4.1.4 Tim12

Tim12 is the only small Tim protein that is not found in the soluble fraction in the intermembrane space but is primarily attached to the inner membrane.[38] Nevertheless, Tim12 is able to form complexes with TIM910 that are key in translocating the substrate protein from the TIM910 complex to the TIM22 insertion machinery.

We found that we could only refold Tim12 in the presence of TIM910. One reason for this behaviour could be that the biogenesis of Tim12 in the cell proceeds in a similar manner. After import and folding of the CX_3C motif by the Mia40 complex, Tim12 is bound by TIM910 and transported to the TIM22 complex.[38] Thus, Tim12 is only soluble in complex with TIM910 during transport to the TIM22 complex.

Isolated Tim12 behaves similarly to Tim9 and Tim10. The reduced form of Tim12 is completely unfolded, while the oxidized form contains the two typical inner and outer helices. Although the length of the helices in isolated Tim12, Tim9 and Tim10 are similar, the dynamic behaviour of these subunits seems to differ. While Tim10 and especially Tim9 suffer from peak broadening, due to millisecond dynamics, nearly all Tim12 peaks are NMR visible. Suggesting that Tim12 does not undergo millisecond dynamics in its

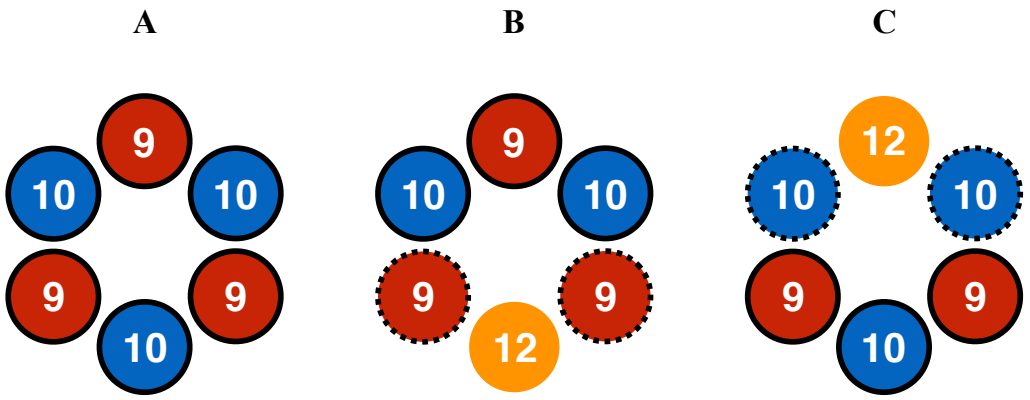


Figure 4.3: **TIM91012 subunit stoichiometry.** Considerations for peak symmetry in NMR spectra of TIM91012. In apo TIM910 (**A**) all three Tim9 and Tim10 subunits share the same neighbours and thus give rise to a single set of peaks. In the TIM91012 hexamer one of the two TIM910 subunits must be exchanged by a Tim12 subunit. Depending on if Tim10 is exchanged (**B**) or Tim9 (**C**) the subunit which is not exchanged will experience three different chemical environments which should give rise to a splitting of peaks.

isolated form. Surprisingly the behaviour of Tim12 in the hexameric complex is completely different from Tim9 and Tim10. In TIM91012 complexes all the Tim12 peaks are broadened and only the very N- and C-terminal residues are visible. Although the extra C-terminal sequence is predicted to be helical, both N- and C-termini are unfolded according to TALOS+ analysis based on our assignment. Possibly helix formation only occurs after binding to the TIM22 complex or to the membrane. Unfortunately, all our attempts to bind Tim12 to lipids, in nanodisks or liposomes, remained unsuccessful (data not shown, conducted by Audrey Hessel). Even though all lipid preparations were enriched with cardiolipin, the main driving force for lipid binding according to Tokatlidis et al.[74] Expression and purification of the soluble intermembrane space domain of Tim54 as a putative binding site for Tim12 was not successful.

A major question in the Tim12 project was also the stoichiometry of the complex. Gebert et al suggested that the stoichiometry of the soluble transport form of TIM91012 ratio between 9:10:12 is 3:1:2,[38] while Adam et al claimed it to be 3:2:1.[1] Although our AUC and gel filtration data show unquestionnably that TIM91012 form a hexamer, we were not able to solve the stoichiometry of this complex. Native mass spectrometry data on our TIM91012 complexes revealed that only one Tim12 subunit is found in each complex, but due to the similarity in size the number of Tim9 and Tim10 subunits could not be determined. The fact that yeast Tim12 is evolution-

arily closer to Tim10 than to Tim9, suggests that a substitution of Tim10 by Tim12 is more likely. Furthermore, we showed that Tim12 preferentially forms complexes with Tim9 but not with Tim10 (see Results section "Tim12"). Thus, in general our data is in agreement with the literature data. However, NMR spectra of TIM91012 suggest that the complex contains three Tim10 subunits. When we prepared NMR samples labelled on Tim9 and Tim10, but not on Tim12, we could observe splitting of peaks in two or three subpopulations. Mostly these splitted peaks belong to Tim10, although theoretically substitution of Tim10 by Tim12 would result in a change of environment for Tim9 (see figure 4.3). This is still ongoing work and to clarify the peak assignment of these splitted peaks we will apply a differential labelling scheme, which should allow the discrimination of Tim9 and Tim10 peaks.

As a side project we aimed at solving the crystal structure of the hexameric TIM91012 complex. Although a first crystallisation trial revealed three different possible crystallisation conditions, further optimisation tests did not give diffracting crystals. The first condition, that diffracted already to 2 Å could not be reproduced, the second condition could be reproduced, but did not diffract and the third condition could not be reproduced. A major hindrance in crystallizing TIM91012 could be the inherent dynamics of the complex and the presence of the unfolded C-terminal extension. An attempt to remove this C-terminal extension resulted in an inability of Tim12 to form complexes and thus was not applicable for further studies.

4.1.5 Subunit exchange

TIM910 and TIM91012 complexes can exchange subunits without influencing the integrity of these complexes. TIM910 is in constant exchange with subunits giving rise to two sets of peaks in NMR spectra. These exchange processes occur on a timescale of 1-2 h, posing the question of the physiological relevance of this exchange, where substrate import is achieved in much shorter times. If this slow timescale is only a result of a non-physiological environment in the NMR samples or if this exchange is only a natural rearrangement of subunits in these complexes cannot be determined. It is tempting though to hypothesize that subunit exchange might be the underlying principle of exchanging between free TIM910-GGC complexes and membrane tethered TIM91012-GGC complexes. Our finding that subunit exchange can also occur on substrate loaded TIM910 complexes could support this hypothesis. However, we failed to exchange Tim12 with GGC-loaded TIM910, the main interaction partners in this hypothetical scenario. The inability of Tim12 to integrate into TIM910-GGC complexes could also be

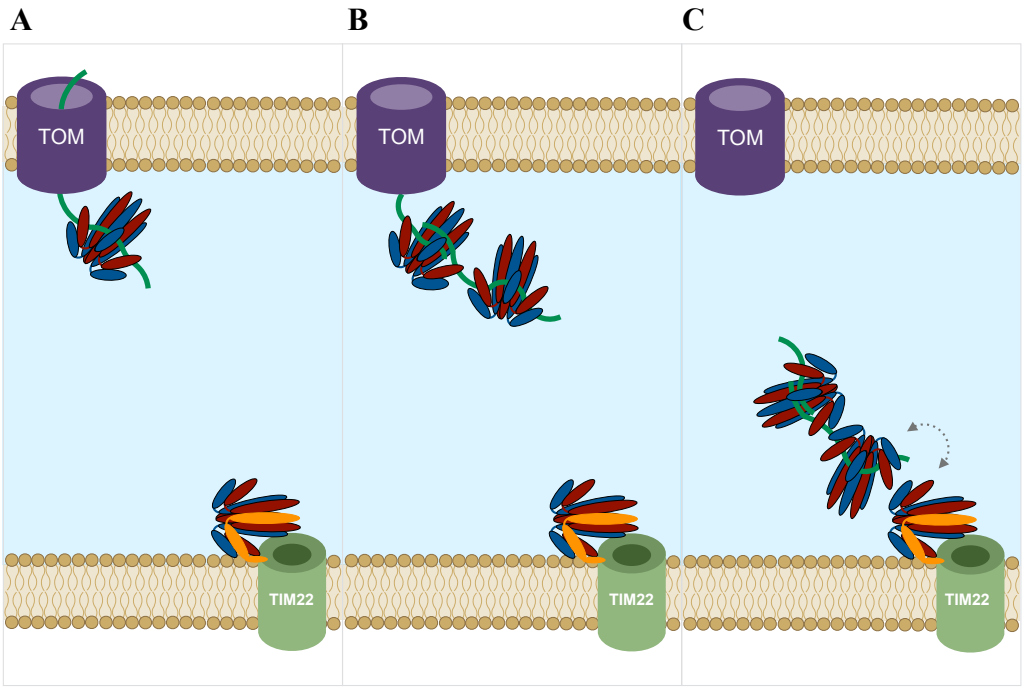


Figure 4.4: **Suggested mechanism for mitochondrial carrier import by TIM910** **A:** The membrane protein precursor is imported into the IMS via the TOM import complex. **B:** One or two TIM910 molecules bind to the substrate protein depedning on its length. **C:** The substrate is taken over by the TIM22 associated TIM91012 complex, either by subunit exchange or handover of the substrate and subsequently inserted into the inner membrane. Tim9 in red, Tim10 in blue and Tim12 in orange.

due to a lack of additional binding partners present in the TIM22 complex or the missing lipid environment. Nevertheless, the fact that TIM910 exists in a dynamic equilibrium between hexamers and monomers might be an important characteristic for a protein that has to form complexes with different subunits and interact dynamically with its binding partners.

4.2 Conclusion

The aim of this PhD thesis was to elucidate the mechanism of chaperoning of mitochondrial inner- and outer-membrane proteins by the intermembrane space chaperone TIM910 and to define the structure of these substrate/chaperone complexes. We could show here that the substrate binding site of TIM910 is located on the exterior of the chaperone in a conserved hydrophobic cleft formed by the inner and outer helices. Inner- as well as

outer-membrane substrates engage the same binding site. A certain level of secondary structure of the substrate protein is beneficial for substrate binding. TIM910 binds substrates in a modular manner, thus adapting to the size of the substrate. Mitochondrial carriers, the major clients of the Tim chaperones, are bound by two TIM910 hexamers. Our NMR and SAXS data on this complexes allowed us to build a detailed model of the structure of these complexes. Residue-resolved millisecond dynamics measurements revealed that both substrate and chaperone undergo millisecond exchange processes when the substrate is bound. We hypothesize that this exchange is due to a dynamic binding behaviour of the substrate, that exchanges between different binding sites or conformations on the millisecond timescale. Furthermore, we provided valuable insight into the structure and dynamics of the hexameric chaperone and its isolated subunits. Exchange between the hexameric TIM910 complex and the membrane attached Tim12 subunit might help explain the mechanism of substrate handover at the inner-mitochondrial membrane.

Chapter 5

Materials and Methods

5.1 Materials and Methods

5.1.1 Protein expression and purification

Cloning, expression and purification of scTIM910

Saccharomyces cerevisiae TIM9 (MDALNSKEQQEFQKVVEQKQMKDF MRLYSNLVERCFTDCVNDFTTSKLTTNKEQTCIMKCSEKFLKHSER VGQRFQEQNAALGQGLGR) was cloned via NdeI/XhoI into the multiple cloning site (MCS)2 of pETDuet-1, subsequently *Saccharomyces cerevisiae* TIM10 (MGSSHHHHHHSQDPENLYFQGSFLGFGGGQPQLSSQQKIQA AEAELDLVTDMFNKLVNNCYKKCINTSYSEGELNKNESSCLDRCVA KYFETNVQVGENMQKMGQSFNAAGKF) containing a N-terminal TEV (Tobacco etch virus) cleavage site was cloned via BamHI/HindIII into the MCS1.

Tim9 and His-tagged Tim10 were co-expressed in SHuffle T7 E. coli cells. After induction at OD_{600} 0.8 with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), cells were incubated at 20°C overnight and harvested by centrifugation.

The cell pellet was resuspended in 25 mL 50 mM Tris-HCl pH 7.4, 150 mM NaCl (buffer A) per litre of cell culture and sonicated. After a heat shock for 15 minutes at 80°C the soluble fraction was separated from the insoluble fraction by centrifugation at 39121xg for 30 min. The supernatant was loaded on a NiNTA affinity column pre-equilibrated with buffer A, then washed with ten column volume (CV)s buffer A supplemented with 20 mM Imidazole and 0.5 M NaCl (buffer B). Hexameric TIM910 complex was eluted with buffer A supplemented with 0.5 M Imidazole (buffer C). The His-tag was cleaved overnight in buffer A supplemented with 0.5 mM dithiothreitol (DTT) and 0.5 mM ethylenediaminetetraacetic acid (EDTA) by adding 1 mg TEV protease to 50 mg protein, cleaved TIM910 was separated from uncleaved TIM910 on a NiNTA column using the above described protocol. The load and wash fractions were then further purified on a HiLoad 16/600 Superdex S200 PG, where TIM910 eluted as a single peak. For NMR experiments gel filtration was performed with 20 mM MES (2-(N-morpholino)ethanesulfonic acid) pH 6.5, 50 mM NaCl (NMR buffer).

To purify the individual subunits of the TIM910 complex, an additional washing step was included into the NiNTA purification protocol: after washing with buffer B, the column was washed with ten CVs of buffer B supplemented with 3 M guanidine-HCl to dissociate the TIM910 complex. While His-tagged Tim10 was still bound to the NiNTA column, Tim9 was found only in the wash fraction. His-tagged Tim10 was eluted with buffer C sup-

plemented with 3 M guanidine-HCl and refolded by flash dilution into buffer A, the His-tag was cleaved as described above. The wash fraction including Tim9 was dialysed against buffer A and both (Tim9 and Tim10) subjected to gel filtration on a HiLoad 16/600 75 PG. Tim9 was found to elute as a dimer, Tim10 mainly as a monomer.

Expression and purification of of TIM910 substrates (scGGC1, scGGC1-C222S, scGGC1-Q8C-C222S, scAAC3-C62A-C233A-C260A-C277S and scAAC3$_{\mathbf{short}}$)

Saccharomyces cerevisiae GGC1 constructs were purchased from GeneCust in pET21b, AAC3 constructs in pET21a plasmids. GGC (MPHTDKKQQSGL ARLLGSASAGGIMEIAVFHPVDTISKRLMSNHTKITSGQELNRVIFRD HFSEPLGKRLFTLFPGLGYAASYKVLQRVYKYGGQPFANEFLNKHY KKDFDNLFGEKTGKAMRSAAAGSLIGIGEIVLLPLDVLKIKRQTNPE SFKGRGFIKILRDEGLFNLYRGWGWTAARNAPGSFALFGGNAFAKE YILGLKDYSQATWSQNFISSIVGACSSLIVSAPLDVIKTRIQNRNFDNP ESGLRIVKNTLKNEGVTAFFKGLTPKLLTTGPKLVFSFALAQSLIPRF DNLLSKLEHHHHHH) constructs were designed with a C-terminal 6xHis-tag and AAC3-C62A-C233A-C260A-C277S (MASMTGGQQMGRGSHHH HHHHHHHLVPMSSDAKQQETNFAINFLLGGVSAAIAIAKTAASPIERV KLIQNQDEMIKQGTLDKKYSGIVDAFKRTAKQEGLISFWRGNTANVI RYFPTQALNFAFKDKIKLMFGFKKEEGYGKWFAGNLASGGAAGAL SLLFVYSLDFARTRLAADAKSSKKGGARQFNGLTDVYKKTLKSDGIA GLYRGFMPSVVGIVVYRGLYFGMFDSLKPLVLTGSLDGSFLASFLLG WVVTTGASTASYPLDTVRRRMMMTSGQAVKYNGAIDALKKIVASE GVGSLFKGSGANILRSVAGAGVISMYDQLQMILFGKKFK) with a N-terminal 9xHis-Tag with an additional sequence described in Bruschweiler et al.[16] AAC3$_{\text{short}}$ contains residues 50–180 of Saccharomyces cerevisiae AAC3 with a C62A mutation, comprising transmembrane helices 2 and 3 (MGHH HHHHGTLDKKYSGIVDSFKRTAKQEGLISFWRGNTANVIRYFPTQAL NFAFKDKIKLMFGFKKEEGYGKWFAGNLASGGAAGALSLLFVYSLD FARTRLAADAKSSKKGGARQFNGLTDVYKKTLKSDGIAGLYR).

Saccharomyces cerevisiae Tom40 and POR1 were purchased from GeneCust in a pET21a plasmid. POR1 (MSPPVYSDISRNINDLLNKDFYHATPAA FDVQTTTANGIKFSLKAKQPVKDGPLSTNVEAKLNDKQTGLGLTQG WSNTNNLQTKLEFANLTPGLKNELITSLTPGVAKSAVLNTTFTQPFF TARGAFDLSLKSPTFVGDLTMAHEGIVGGAEFGYDISAGSISRYAMA LSYFAKDYSLGATLNNEQITTVDFFQNVNAFLQVGAKATMNSKLPN SNVNIEFATRYLPDASSQVKAKVSDSGIVTLAYKQLLRPGVTLGVGS SFDALKLSEPVHKLGWSLSFDALEHHHHHH) contained two Cys muta-

tions (C130S, C210S) and a C-terminal 6xHisTag. Tom40 (MSAPTPLAEAS QIPTIPALSPLTAKQSKGNFFSSNPISSFVVDTYKQLHSHRQSLELVNP GTVENLNKEVSRDVFLSQYFFTGLRADLNKAFSMNPAFQTSHTFSIG SQALPKYAFSALFANDNLFAQGNIDNDLSVSGRLNYGWDKKNISKVN LQISDGQPTMSQLEQDYQASDFSVNVKTLNPSFSEKGEFTGVAVASF LQSVTPQLALGLETLYSRTDGSAPGDAGVSYLTRYVSKKQDWIFSGQ LQANGALIASLWRKVAQNVEAGIETTLQAGMVPITDPLMGTPIGIQP TVEGSTTIGAKYEYRQSVYRGTLDSNGKVASFLERKVLPTLSVLFSG EIDHFKNDTKIGSGLQFETAGNQELLMLQQGLDADGNPLQALPQLLE HHHHHH) four Cys mutations and C165S, C326S, C341S, C355S and also a C-terminal 6xHisTag.

All TIM910 substrates were expressed in BL21(DE3) *E.coli* cells as inclusion bodies. Cells were grown at 37°C until an OD_{600} of 0.7 was reached, and induced with 1 mM IPTG for 3 hours. The cell pellet was resuspended in buffer A, sonicated and centrifuged at 39121xg for 30 min. The inclusion body pellet was resuspended in buffer A supplemented with 6 M of guanidine-HCl (for GGC wild type 2 mM DTT were added to each purification step) and incubated at 4°C for 16 hours. The solubilized protein was separated from the insoluble fraction by centrifugation and loaded on a NiNTA resin. Proteins were purified as described for TIM910 with addition of 3 M guanidine-HCl to all buffers.

Expression and purification of scTim12

pET28a plasmids containing N-terminal 6xHis-tagged scTim12-C40A-C61A and wild type Tim12 (MGHHHHHHENLYFQGSFFLNSLRGNQEVSQE KLDVAGVQFDAMASTFNNILSTCLEKCIPHEGFGEPDLTKGEQACID RCVAKMHYSNRLIGGFVQTRGFGPENQLRHYSRFVAKEIADDSKK) were purchased from GeneCust and expressed in SHuffle T7 *E. coli* cells as described for TIM910. Tim12 was found in the insoluble fraction and first purified as described for the TIM910 substrates guanidine-HCl. Denatured Tim12 was bound to a NiNTA column, washed with five CVs of buffer A and incubated with a threefold excess of TIM910 for 30 min and subsequently eluted. After TEV-cleavage (see TIM910 purification), a gel filtration profile on a HiLoad 16/600 Superdex S200 PG in buffer A shows two separate peaks: Tim12 and Tim12 in a complex with TIM910.

Isotope labelling

For NMR experiments TIM910 was expressed in D_2O M9 minimal medium and either labelled with 1 g/L [^{15}N] NH_4Cl and 2 g/L D-[^{2}H,^{13}C] glucose

or specifically labelled on alanine, leucine and valine side chains using the SLAM-Aβ and DLAM-LVproS kit from NMR-Bio according to the manufacturers instructions. Accordingly scGGC1-C222S was labelled on alanine and isoleucine side chains using the DLAM-Aβ/Iδ1 from NMR-Bio.

Site-specific spin-labelling of scGGC1

GGC was incubated for one hour at room temperature in 6 M guanidine-HCl with 10 mM DTT. Subsequently DTT was removed on an illustra NAP25 column and the protein incubated with a threefold molar excess of MTSL dissolved in dimethyl sulfoxide (DMSO) for 16 hours at 4°C in the dark. Free MTSL was removed on an illustra NAP25 column. Efficiency of the spin labelling reaction was confirmed by continuous-wave electron paramagnetic resonance (EPR) spectroscopy (data not shown).

Reconstitution of chaperone/substrate complexes

Purified substrate protein in 6M guanidine-HCl was bound to NiNTA resin and washed with five CVs of 1 M guanidine-HCl. Subsequently a twofold excess of TIM910 was added to the column. The flow through containing TIM910 was collected, diluted 1:1 with buffer A and reloaded on the column. This step was repeated until ten CVs are reached and the guanidine-HCl concentration was below 0.05 M. The column was washed with another five CVs of buffer A, the TIM910-substrate complex was eluted with five CVs of buffer C and dialysed into buffer A or NMR buffer.

5.1.2 NMR experiments

All NMR experiments were performed on Bruker Avance III spectrometers equipped with cryogenically cooled TCI probeheads, operating at magnetic field strengths corresponding to ^{1}H Larmor frequencies of 950, 850, 700 and 600 MHz, respectively. The sample temperature was set to 308 K, unless stated otherwise.

Sequence-specific resonance assignments of TIM910

Resonance assignment experiments were performed on TIM910 complexes with individually labelled subunits. To this end [^{2}H]-labelled His-tagged Tim10 was mixed with a 1.2-fold excess of [^{2}H,^{13}C,^{15}N]-labelled Tim9 and purified on a Ni-NTA column as described above. Unbound Tim9 was found in the wash fraction, while hexameric TIM[^{2}H,^{13}C,^{15}N]9[^{2}H]10 was eluted. The His-tag was cleaved as described above and the complex subjected to

gel filtration on a HiLoad 16/600 Superdex S200 PG in NMR buffer. To obtain TIM[^{2}H]9[^{2}H,^{13}C,^{15}N]10, the protocol was repeated with [^{2}H,^{13}C,^{15}N] His-tagged Tim10 and [^{2}H] Tim9.

For assignment the following experiments were performed: 2D ^{15}N-^{1}H-BEST-TROSY HSQC, 3D BEST-TROSY HNCO, 3D BEST-TROSY HNcaCO, 3D BEST-TROSY HNCA, 3D BEST-TROSY HNcoCA, 3D BEST-TROSYHNcocaCB and 3D BEST-TROSY HNcaCB.[33]

Methyl groups of Ala$^{\beta}$, Leu$^{\delta 1}$ and Val$^{\gamma 1}$ were assigned using 3D ^{1}H-^{1}H-NOESY-^{13}C-HSQC and 4D ^{15}N-HSQC-NOESY-^{13}C-HSQC with 300 ms mixing time. The assignment of Ala$^{\beta}$ was additionally confirmed through the ^{13}C$^{\beta}$ chemical shift, obtained from HNCACB experiments (see above).

Data processing and analysis was performed using the NMRPipe software package and CCPN software.[25]

Relaxation experiments

^{15}N R_1 and R_2 relaxation rate constants and hetNOE data were derived from 2D ^{15}N-^{1}H-TROSY HSQC experiments. 12 relaxation delays for T_1 experiments were ranging from 0.01 to 1.21 s and 14 delays from 0 to 85 ms for T_2.

Diffusion ordered spectroscopy

Diffusion constants were derived from 1D ^{13}C filtered DOSY experiments performed at 600 MHz ^{1}H Larmor frequency. All samples were measured at 70 µM at 308 K in NMR buffer with 10% (v/v) D_2O. The effective diffusion time for TIM910 was 0.1 s or 0.2 s and for [TIM910 GGC] complexes 0.3 s. Three samples were measured: $^{13}CH_3$-ALV TIM910, $^{13}CH_3$-ALV TIM910 in complex with unlabelled GGC and unlabelled TIM910 in complex with $^{13}CH_3$-AI GGC.

Methyl order parameter measurements from relaxation-violated coherence transfer NMR

A triple-quantum based relaxation-violated coherence transfer experiment was performed to measure the fast methyl dynamics in TIM910 and in TIM910-GGC1, i.e. the product of site-specific methyl-axis order parameter and the overall-tumbling correlation time constant.[121] TIM910, labelled at $^{13}CH_3$-ALV groups in an otherwise perdeuterated environment was either measured in the apo form or with bound GGC, in D_2O 20 mM MES-d_{13} 50 mM NaCl pD 6.1 (D_2O NMR buffer). The experiments were recorded at a ^{1}H Larmor frequency of 950 MHz, according to Sun et al.[121] 12 different values

of the 3Q delay (delay T in Figure 2 of Sun et al[121]) were acquired, from 1.1 to 35 ms. The ratio of peak integrals in the relaxation-violated coherence transfer (I_a) and the reference experiment (I_b) were fitted according to eq. of Sun et al:

$$\frac{I_a}{I_b} = \frac{3}{4}\frac{\eta \tanh(\sqrt{\eta^2+\delta^2}T)}{\sqrt{\eta^2+\delta^2} - \delta \tanh(\sqrt{\eta^2+\delta^2}T)}$$

where δ is an empirical fit constant, and η is proportional to $S^2_{\tau c}$. The dependency of $\frac{I_a}{I_b}$ on the relaxation delay T was fitted numerically using a grid-search procedure implemented in an in-house written python program to extract η and δ. Error margins of the fitted parameters were obtained using a standard Monte Carlo procedure. Hereby, 1000 synthetic data sets were generated, by adding random noise (within the experimentally determined noise levels) onto the data points that were back-predicted from the best-fit parameters. These 1000 data sets were fitted analogously to the experimental data, to extract the standard deviation over these fitted parameters

CPMG experiments

^{1}H-^{13}C multiple-quantum relaxation dispersion experiments[66] were recorded to probe millisecond motion of methyl groups in TIM910, TIM910-GGC, TIM910-AAC3 and GGC. The samples were either $^{13}CH_3$-ALV TIM910, or $^{13}CH_3$ TIM910 in complex with unlabelled GGC, or $^{13}CH_3$-$A^{\beta}/I^{\delta 1}$ GGC bound to unlabelled TIM910. The CPMG relaxation period was set to 40 ms (TIM910 and TIM910-GGC) or 60 ms (TIM910-AAC). The CPMG frequency applied for 40 ms was 0, 50, 100, 150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900 and 1000 Hz and for 60 ms 66, 133, 200, 267, 333, 400, 466, 533, 600, 733, 867, 1000 Hz. Additionally a reference experiment with no CPMG relaxation period was performed. Experiments were performed at different static magnetic field strengths (600, 700, 850, 950 MHz), as stated in the figure captions.

CPMG profiles, $R_{2,\text{eff}}$ as a function of ν_{CPMG}, were determined from peak volumes, and fitted with a numerical approach, using the program ChemEx (version 0.6.0, available at https://github.com/gbouvignies/chemex/releases), in order to obtain a global exchange rate constant, k_{ex}, minor-state population p_B and site-specific chemical-shift differences $|\Delta\omega_{13C}|$. Populations where set to 50%, since we expected the substrate to sample a multitude of different states.

^{1}H intermolecular cross-relaxation (HETSOFAST) experiments

To detect the contact surface of GGC with TIM910, we probed the effect of ^{1}H spin saturation transfer, i.e. a NOE-type transfer, from unlabelled GGC to a deuterated $^{13}CH_3$-ALV-labelled TIM910. The principle of this measurement is identical to previously proposed HETSOFAST experiments,[104] except that we apply this concept here to methyls, rather than backbone amides. The effect of saturation transfer, i.e. cross-relaxation from the protons of GGC to TIM910 is monitored by measuring the intensity of TIM910 methyl peaks in two experiments. The first data set (I_{ref}) is a standard ^{1}H-^{13}C SOFAST experiment,[105] using a recycle delay of 0.2 s. In a second experiment (I_{sat}), an inversion pulse is applied at the beginning of the recycle delay, which inverts protons of GGC. If a given TIM910 methyl site is in close spatial proximity, the effective ^{1}H longitudinal relaxation during the recycle delay will be slowed down in this latter experiment, because of the intermolecular cross-relaxation. The inversion pulse (ISNOB2[67]) was applied at 4 ppm with a band width of 2 ppm, i.e. in a spectral region that is devoid of TIM910 resonances, due to the >98% deuteration. The differential longitudinal relaxation of methyl ^{1}H spins on TIM910 that arises from the intermolecular cross-relaxation can be monitored by the difference of the intensities in the two experiments.[104] To ensure that artefacts due to residual protonation of TIM910 do not introduce bias, we recorded these experiments not only on $^{13}CH_3$-ALV-TIM910 bound to unlabeled GGC, but also on $^{13}CH_3$-ALV-TIM910 in the absence of GGC. In this latter experiment no inter-molecular cross-relaxation can occur.

In all cases, the samples were in D_2O NMR buffer, and the experiments were recorded at a ^{1}H Larmor frequency of 950 MHz.

PRE experiments

A complex between $^{13}CH_3$-ALV-TIM910 and MTSL-labelled GGC (wild type, containing a Cys at position 222 or scGGC1-Q8C-C222S) was prepared as described above. Methyl ^{1}H R_1 relaxation was measured as described above for this TIM910-GGC_{MTSL} complex, and a sample in which the MTSL spin label was reduced by addition of 5 mM sodium ascorbate. The effect of the paramagnetically-tagged GGC was monitored via the difference between R_1 relaxation rate constants between the paramagnetic and the ascorbate-reduced sample.

solvent PRE experiments

In order to probe the solvent exposed areas of TIM910, and those regions that become inaccessible to solvent when membrane proteins are bound, we performed additional solvent paramagnetic relaxation experiments, i.e. the enhancement of proton relaxation that is induced by a soluble gadolinium complex, gadolinium diethylenetriaminepentaacetic acid–bismethylamide (Gd(DTPA-BMA)). The approach has been described elsewhere,[50] and was implemented as follows.

We have measured the ^{1}H longitudinal relaxation of methyl groups in deuterated $^{13}CH_3$-ALV-TIM910 via saturation recovery. To this end we recorded a series of ^{1}H-^{13}C SOFAST experiments in which a ^{1}H saturation element and a recovery delay were applied prior to the initial excitation pulse. The recovery delay was incremented from zero to 3 s (10 points), and the site-specific buildup rate constants were fitted by a mono-exponential function.

The recycle delay, i.e. the delay between detection of the free indutcion decay (FID) and the saturation element of the subsequent scan is inconsequential, because proton polarization is removed by the saturation element, and was set to 50 ms.

This experiment was performed with $^{13}CH_3$-ALV-TIM910, and $^{13}CH_3$-ALV-TIM910-GGC, and $^{13}CH_3$-ALV-TIM910-AAC, using each time several concentrations of Gd(DTPA-BMA), ranging from 0 to 3 mM.

Residue-wise R_1 buildup rate constants were plotted as a function of the concentration of the paramagnetic agent Gd(DTPA-BMA), and the concentration-dependence of R_1 was fitted with a linear function. The resulting slope, d R_1/d [Gd(DTPA-BMA)] is indicative of the solvent accessibility of a given site.

Subunit exchange experiments

Unlabelled TIM910 or TIM910-GGC complex was mixed in a 1:3 ratio with [^{2}H,^{15}N,^{13}C] Tim9 or Tim10. Immediately after mixing (dead time of 3-4 minutes), a series of ^{15}N-^{1}1H-BEST-TROSY HSQC experiments was started for 4 hours whereby one single experiment lasted 3 min 40 s.

Data was analysed by following the build up of TIM910 complex peaks or the loss of signal intensity in Tim9/Tim10 free subunit peaks. The sum of the peak intensity at each time point was plotted against the time and a mono exponential fit performed.

Oxidative refolding of Tim9

Purified Tim9 was unfolded with 6 M guanidine-HCl and 10 mM TCEP. The buffer was exchanged to NMR buffer on an illustra Nap-5 column and immediately measured by NMR. A series of Best-HSQC spectra was recorded for about two days. The dead time for the first experiment was 15 min and the approximate duration of a single experiment 8 min. To measure folding kinetics always ten spectra were summed up to increase the signal/noise ratio. Peak intensity decrease was fitted to an exponential decay using an in-house written python script. The protein concentration during the experiment was 0.2 mM.

VDAC peptide titration

VDAC peptides (obtained from Doron Rapaport, for sequence see figure 3.13) were dissolved in DMSO . The DMSO concentration was reduced by stepwise addition of NMR buffer (1:1 in each step). At a DMSO concentration of 10% (v/v) TIM910 in NMR buffer was added instead of pure NMR buffer. The final concentration of DMSO was 6%. The following samples were measured: 0.15 mM TIM910 + 0.15 mM VDAC (linear or cyclic), 0.15 mM TIM910 + 0.3 mM VDAC, 0.15 mM TIM910 + 0.45 mM VDAC, 0.15 mM TIM910 + 0.6 mM VDAC, 0.15 mM TIM910 + 6% DMSO.

5.1.3 Small-angle X-ray scattering

SAXS data were collected at ESRF BM29[92] with a Pilatus 1M detector (Dectris) at the distance of 2.872 m from the 1.8 mm diameter flow-through capillary. The scattering of pure water was used to calibrate the intensity to absolute units. Data collection was performed continuously throughout the chromatography run at a frame rate of 1 Hz. The X-ray energy was 12.5 keV and the accessible q-range 0.032 nm^{-1} to 4.9 nm^{-1}. The incoming flux at the sample position was in the order of 1012 photons/s in 700x700 μ^2. All images were automatically azimuthally averaged with pyFAI (Ashiotis et al., 2015).

SAXS data of pure TIM910 was collected at 1, 2.5 and 5 mg/mL using the BioSAXS sample changer.[100] 10 frames of 1 s were collected for each concentration. Exposures with radiation damage were disposed, the remaining frames averaged and the buffer background subtracted by the online processing pipeline.[15] Data from the three concentrations were merged following standard procedures to create an idealized scattering curve of TIM910, using Primus from the ATSAS package.[93] The pair distribution function p(r) was calculated using GNOM.[122] 40 ab-initio models each were calculated in C1,

C3 and C6 symmetry, using DAMMIF[35] and averaged, aligned and compared using DAMAVER.[138] The models in C1 and C3 symmetry agreed well with the data, whereas the quality of the fit was considerably worse in C6 symmetry. Additionally, 10 ab-initio models in C1 symmetry were calculated with GASBOR,[123] using 864 beads ($\sim$ 95 kDa). The GASBOR models agreed well with the DAMMIF models in C1 symmetry. For comparison of the SAXS data to the TIM910 crystal structure (pdb access code 3DXR[6] the missing C- and N-termini were modelled with CORAL.[93]

Online purification of the TIM-GGC complex using a NiNTA column was performed with a high pressure liquid chromatography (HPLC) system (Shimadzu, France), as described in.[15] 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. The stoichiometry of the individual peaks in the chromatogram was determined molecular mass estimates based on the correlated and Porod volumes[93] and by calculating the ratios of maximum forward scattering and maximum absorbance at 280 nm for each peak and comparing to the theoretical values of different stoichiometries.

For creating background corrected scattering curves for the TIM910:GGC 2:1 complex peak, measurements of buffer with the lowest (200 mM) and the highest (300 mM) possible imidazole concentration, as well as interpolated curves between these two extremes, were subtracted individually from each frame. The protein concentration independence of the resulting curves was confirmed as in Hynson et al.[51] Subtraction of the highest imidazole concentration provided a stable signal throughout the center of the peak. 50 frames from the middle of the peak were averaged to create a single SAXS curve for the TIM910:GGC 2:1 complex. For all subsequent analys, the robustness of the main findings (general shape and dimensions) was confirmed by analogous treatment of curve resulting from the subtraction of the lowest possible imidazole concentration. The pair distribution function p(r) was calculated using GNOM.[122] 40 ab-initio models each were calculated in C1, C3 and C32 symmetry, using DAMMIF[35] and averaged, aligned and compared using DAMAVER.[138] Only the models in C1 symmetry agreed well with the data. Additional GASBOR models[123] in C1, using 1418 beads ($\sim$ 156 kDa), corresponded well to the DAMMIF models.

Based on the known crystal structure (pdb access code 3DXR[6]) and the CORAL results for pure TIM910 complex a model for TIM910 was built using iTASSER. GGC was split up in 8 α-helical regions separated by flexible linkers. While this arrangement most likely does not correspond to reality, it reduces the available conformation space sufficiently to allow us to model the distribution of GGC in the TIM910-GGC complex with reasonable computational effort. The position of two TIM910 complexes was fixed in agreement

the DAMMIF models and rotated in respect to each other to generate 7 sets of starting conditions. The distance of the centers of mass was 71.3 Å. Freeing the positions of the TIM910s did not significantly improve fit quality (χ^2 = 1.13 as opposed to 1.16).

Data-collection parameters	TIM9-10	TIM-GGC
Instrument:	ESRF BM29	
Wavelength (Å)	0.99	
q-range ($Å^{-1}$)	0.0032 – 0.49	
Sample-to-detector distance (m)	2.864	
Exposure time (sec)	10 x 1	1 per frame
Concentration range (mg/mL)	1 - 5	0 - 2
Temperature (K)	293	
Detector	Pilatus 1M (Dectris)	
Flux (photons/s)	5×10^{11}	2×10^{12}
Beam size (μm^2)	700 × 700	
Structural parameters		
I_0 (cm^{-1}) [from Guinier]	0.0400	n.a.
R_g (Å) [from Guinier]	26.7	45.1
$q_{min}R_g$ – $q_{max}R_g$ used for Guinier	0.24- 1.29	0.76 – 1.11
D_{max} (Å)	100	160
I_0 (cm^{-1}) [from p(r)]	0.0402	n.a.
R_g (Å) [from p(r)]]	27.3	46.0
Porod volume V_p (10^3 $Å^3$) [from Scåtter]	103	272
Molecular mass M_r (kDa) [from V_p]	61	160
Correlated volume V_c ($Å^2$) [from Scåtter]	417	897
Molecular mass M_r (kDa) [from V_c]	53	145
Calculated monomeric M_r from sequence (kDa)	61	156
Software employed		
Primary data reduction	EDNA, pyFAI	
Data processing	Primus, GNOM, Scåtter	
Ab initio analysis	Dammif, Gasbor	
Validation and averaging	*DAMAVER*	
Rigid-body modelling	Coral	
Three-dimensional graphics representations	PyMOL	

Figure 5.1: **Parameters of SAXS data acquisition and analysis.**

5.1.4 Additional biochemical and biophysical assays

Analytical ultracentrifugation

Measurements are performed on the Plateforme Biophysique AUC-PAOL , IBS, Grenoble. Samples are measured in NMR buffer at 10°C at three different concentrations (0.2-0.002mM), recording interference and A_{275nm}. Data analysis is done using GUSSI software.

Circular dichroism spectroscopy

Far ultraviolet (FUV) CD spectra were recorded in a 1 mm cell at 0.1 mg/mL protein concentration at 20°C and 35°C in buffer A on a Jasco J-810 spectropolarimeter continuously purged with nitrogen.

Native Mass spectrometry

The samples were subjected to gel filtration on an Agilent Bio SEC-5 column in 250 mM ammonium acetate. Afterwards, the samples were analysed by native mass spectrometry (MS).[14] Protein ions were generated using a nanoflow electrospray ionization (ESI) source. Nanoflow platinum-coated borosilicate electrospray capillaries were bought from Thermo Electron SAS (Courtaboeuf, France). MS analyses were carried out on a quadrupole time-of-flight mass spectrometer (Q-TOF Ultima, Waters Corporation, Manchester, U.K.). The instrument was modified for the detection of high masses.[115,131] The following instrumental parameters were used: capillary voltage = 1.2-1.3 kV, cone potential = 40 V, RF lens-1 potential = 40 V, RF lens-2 potential = 1 V, aperture-1 potential = 0 V, collision energy = 30-140 V, and microchannel plate (MCP) = 1900 V. All mass spectra were calibrated externally using a solution of cesium iodide (6 mg/mL in 50% isopropanol) and were processed with the Masslynx 4.0 software (Waters Corporation, Manchester, U.K.) and with Massign software package.[85]

SDS PAGE analysis

SDS PAGE analysis of 27 independently prepared TIM910-GGC complex samples was performed using Image Lab Software. The bands were integrated and the percentage between the integrals of TIM910 and GGC calculated. According to the number of amino acids the expected percentage of a 1 : 1 TIM910 : GGC complex would be 36% : 64%, a 2 : 1 complex would be 22% : 78%.

Cell free expression

scGGC1 with a C-terminal HisTag in pIVEX 2.3d was expressed under cell free conditions for 2 h at 30°C.

condition	1	2	3	4	5	6
used vector	GGCHis pIVEX 2.3d	GGCHis pIVEX 2.3d	GGCHis pIVEX 2.3d	GGCHis pIVEX 2.3d	GGCHis pIVEX 2.3d	GGCHis pIVEX 2.3d
DNA initial (µg/ml)	1050	1050	1050	1050	1050	1050
S30 extract	Brij35 0,5%	x	x	x	x	x
MgOAc mM	16	16	16	16	16	16
DNA final (µg/ml)	10.00	10.00	10.00	10.00	10.00	10.00
Condition S30 pour 50 µl	20	20	20	20	20	20
10X Buffer	5	5	5	5	5	5
creatine phosphate (1M)	4	4	4	4	4	4
potassium glutamate (4M)	3	3	3	3	3	3
magnesium acetate (1,07M)	0.49	0.49	0.49	0.49	0.49	0.49
aa mix 15mM	4	4	4	4	4	4
creatine kinase (10 mg/ml)	1	1	1	1	1	1
T7 pol (2.5 mg/ml)	1	1	1	1	1	1
S30 LI12F	20	20	20	20	20	20
PEG 50%	0	0	0	0	0	0
vector	0.48	0.48	0.48	0.48	0.48	0.48
tRNA (17.5 mg/ml)	1	1	1	1	1	1
Brij 35 20%	1	0	0	0	0	0
TIM910 (22 mg/ml)	0	0	1	3	5	10
water	9.94	11.19	10.69	8.69	6.19	1.19
total volume (µl)	50	50	50	50	50	50

Figure 5.2: **Cell free conditions.** Experimental conditions for Cell-free expression of GGC in the presence of TIM910.

5.1.5 Crystallization

For a first screen of crystallization conditions we used the PSB HTXlab. The crystallisation temperature was 20°C and the protein concentration was 15 mg/mL. Two ratios (1:1 and 1:2) were tested. The following crytallization screens were used: PEGs-I (Quiagen), Salt-Grid (Hampton Research), Wizard I+II (Rigaku), Classics Suite (Quiagen), JGSG (MD) and PACT (MD). Crystallization conditions were optimized in 24-well hanging-drop plates at 20°C at two different ratios and different starting protein concentrations.

Successful conditions from the HTXlab crystallization screen were:

- Condition 1: 0.1 M sodium cacodylate, 0.2 M calcium acetate, 40% (v/v) PEG 300, pH 6.5
- Condition 2: 0.1 M phosphate/citrate, 40% v/v Ethanol, 5% (w/v) PEG 1K, pH 4.2
- Condition 3: 0.2 M ammonium sulfate, 0.1 M sodium acetate, 30% (w/v) PEG 2000 MME, pH 4.6

Chapter 6

Bibliography

Bibliography

[1] Adam, a., Endres, M., Sirrenberg, C., Lottspeich, F., Neupert, W., and Brunner, M. Tim9, a new component of the TIM22.54 translocase in mitochondria. *The EMBO journal 18* (1999), 313–9.

[2] Alcock, F. H., Grossmann, J. G., Gentle, I. E., Likić, V. a., Lithgow, T., and Tokatlidis, K. Conserved substrate binding by chaperones in the bacterial periplasm and the mitochondrial intermembrane space. *The Biochemical Journal 409* (2008), 377–387.

[3] Alderson, T. R., Kim, J. H., and Markley, J. L. Dynamical Structures of Hsp70 and Hsp70-Hsp40 Complexes. *Structure 24* (2016), 1014–1030.

[4] Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science 181* (1973), 223–230.

[5] Baker, M. J., Tatsuta, T., and Langer, T. Quality control of mitochondrial proteostasis. *Cold Spring Harb Perspect Biol 3* (2011), 1–19.

[6] Baker, M. J., Webb, C. T., Stroud, D. A., Palmer, C. S., Frazier, A. E., Guiard, B., Chacinska, A., Gulbis, J. M., and Ryan, M. T. Structural and Functional Requirements for Activity of the Tim9 – Tim10 Complex in Mitochondrial Protein Import. *Molecular Biology of the Cell 20* (2009), 769–779.

[7] Bar-Lavan, Y., Shemesh, N., and Ben-Zvi, A. Chaperone families and interactions in metazoa. *Essays In Biochemistry 60* (2016), 237–253.

[8] Bausewein, T., Mills, D. J., Langer, J. D., Nitschke, B., Nussberger, S., and Kühlbrandt, W. Cryo-EM Structure of the TOM Core Complex from Neurospora crassa. *Cell 170* (2017), 693–700.e7.

[9] Bayrhuber, M., Meins, T., Habeck, M., Becker, S., Giller, K., Villinger, S., Vonrhein, C., Griesinger, C., Zweckstetter, M., and Zeth, K. Structure of the human voltage-dependent anion channel. *Proc. Natl. Acad. Sci. (U.S.A.) 105* (2008), 15370–15375.

[10] Becker, T., Pfannschmidt, S., Guiard, B., Stojanovski, D., Milenkovic, D., Kutik, S., Pfanner, N., Meisinger, C., and Wiedemann, N. Biogenesis of the mitochondrial tom complex: Mim1 promotes insertion and assembly of signal-anchored receptors. *Journal of Biological Chemistry 283* (2008), 120–127.

[11] Bertelsen, E. B., Chang, L., Gestwicki, J. E., and Zuiderweg, E. R. P. Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proceedings of the National Academy of Sciences 106* (2009), 8471–8476.

[12] Beverly, K. N., Sawaya, M. R., Schmid, E., and Koehler, C. M. The Tim8-Tim13 Complex Has Multiple Substrate Binding Sites and Binds Cooperatively to Tim23. *Journal of Molecular Biology 382* (2008), 1144–1156.

[13] Bitto, E., and McKay, D. B. Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure 10* (2002), 1489–1498.

[14] Boeri Erba, E., and Petosa, C. The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Science 24* (2015), 1176–1192.

[15] Brennich, M. E., Round, A. R., and Hutin, S. Online size-exclusion and ion-exchange chromatography on a saxs beamline. *Journal of Visualized Experiments : JoVE* (2017), 54861.

[16] Brüschweiler, S., Yang, Q., Run, C., and Chou, J. J. Substrate-modulated ADP / ATP-transporter dynamics revealed by NMR relaxation dispersion. *Nature structural & molecular biology 22* (2015), 636–41.

[17] Bukau, B., Weissman, J., and Horwich, A. Molecular chaperones and protein quality control. *Cell 125* (2006), 443–51.

[18] Burmann, B. M., Wang, C., and Hiller, S. Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. *Nature structural & molecular biology 20* (2013), 1265–72.

[19] Clare, D. K., and Saibil, H. R. ATP-driven molecular chaperone machines. *Biopolymers 99* (2013), 846–859.

[20] Claypool, S. M., McCaffery, J. M., and Koehler, C. M. Mitochondrial mislocalization and altered assembly of a cluster of Barth syndrome mutant tafazzins. *Journal of Cell Biology 174* (2006), 379–390.

[21] Cohen, S. I. A., Arosio, P., Presto, J., Kurudenkandy, F. R., Biverstål, H., Dolfe, L., Dunning, C., Yang, X., Frohm, B., Vendruscolo, M., Johansson, J., Dobson, C. M., Fisahn, A., Knowles, T. P. J., and Linse, S. A molecular chaperone breaks the catalytic cycle that generates toxic aβoligomers. *Nat Struct Mol Biol 22* (2015), 207–213.

[22] Curran, S., and Leuenberger, D. The Tim9p–Tim10p complex binds to the transmembrane domains of the ADP/ATP carrier. *The EMBO Journal 21* (2002).

[23] Davis, A. J., Alder, N. N., Jensen, R. E., and Johnson, A. E. The Tim9p/10p and Tim8p/13p Complexes Bind to Specific Sites on Tim23p during Mitochondrial Protein Import. *Molecular Biology of the Cell 18* (2006), 475–486.

[24] de Marcos-Lousa, C., Tokatlidis, K., and Baud, C. Molecular interactions of the mitochondrial Tim12 translocase subunit. *Protein and Peptide Letters 14* (2007), 597–600.

[25] Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. NMRPIPE - a multidimensional spectral processing system based on Unix pipes. *J. Biomol. NMR 6* (1995), 277–293.

[26] Demishtein-Zohary, K., Günsel, U., Marom, M., Banerjee, R., Neupert, W., Azem, A., and Mokranjac, D. Role of tim17 in coupling the import motor to the translocation channel of the mitochondrial presequence translocase. *eLife 6* (2017), e22696.

[27] Dimmer, K. S., Papić, D., Schumann, B., Sperl, D., Krumpe, K., Walther, D. M., and Rapaport, D. A crucial role for mim2 in the biogenesis of mitochondrial outer membrane proteins. *Journal of Cell Science 125* (2012), 3464–3473.

[28] Dolezal, P., Likic, V., Tachezy, J., and Lithgow, T. Evolution of the molecular machines for protein import into mitochondria. *Science 313* (2006), 314–318.

[29] Duran, E. C., Weaver, C. L., and Lucius, A. L. Comparative Analysis of the Structure and Function of AAA + Motors ClpA , ClpB , and Hsp104: Common Threads and Disparate Functions. *Frontiers in Molecular Biosciences 4* (2017), 1–19.

[30] Durigon, R., Wang, Q., Ceh Pavia, E., Grant, C. M., and Lu, H. Cytosolic thioredoxin system facilitates the import of mitochondrial small Tim proteins. *EMBO reports 13* (2012), 916–922.

[31] Echeverría, P. C., Bernthaler, A., Dupuis, P., Mayer, B., and Picard, D. An interaction network predicted from public data as a discovery tool: Application to the hsp90 molecular chaperone machine. *PLOS ONE 6* (2011), 1–11.

[32] Ellis, R., and Minton, A. Protein aggregation in crowded environments. *Biological Chemistry 387* (2006), 485–497.

[33] Favier, A., and Brutscher, B. Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J. Biomol. NMR 49* (2011), 9–15.

[34] Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature 431* (2004), 9396–9401.

[35] Franke, D., and Svergun, D. I. Dammif, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography 42* (2009), 342–346.

[36] Gao, W., Kim, J.-y., Anderson, J. R., Akopian, T., Hong, S., Jin, Y.-y., Kandror, O., Kim, J.-w., Lee, I.-a., Lee, S.-y., Mcalpine, J. B., Mulugeta, S., Sunoqrot, S., Wang, Y., Yang, S.-h., Yoon, T.-m., Goldberg, A. L., Pauli, G. F., Suh, J.-w., and Franzblau, S. G. The Cyclic Peptide Ecumicin Targeting ClpC1 Is Active against Mycobacterium tuberculosis In Vivo. *Antimicrobial Agents and Chemotherapy 59* (2015), 880–889.

[37] Gavrish, E., Sit, C. S., Cao, S., Kandror, O., Spoering, A., Peoples, A., Ling, L., Fetterman, A., Hughes, D., Bissell, A., Torrey, H., Akopian, T., Mueller, A., Epstein, S., Goldberg, A., Clardy, J., and Lewis, K. Article Lassomycin , a Ribosomally Synthesized Cyclic Peptide , Kills Mycobacterium tuberculosis by Targeting the ATP-Dependent Protease ClpC1P1P2. *Chemistry & Biology 21* (2014), 509–518.

[38] Gebert, N., Chacinska, A., Wagner, K., Guiard, B., Koehler, C. M., Rehling, P., Pfanner, N., and Wiedemann, N. Assembly of the three small Tim proteins precedes docking to the mitochondrial carrier translocase. *EMBO reports 9* (2008), 548–554.

[39] Gentle, I. E., Perry, A. J., Alcock, F. H., Likić, V. a., Dolezal, P., Ng, E. T., Purcell, A. W., McConnville, M., Naderer, T., Chanez, A.-L., Charrière, F., Aschinger, C., Schneider, A., Tokatlidis, K., and Lithgow, T. Conserved motifs reveal details of ancestry and structure in the small TIM chaperones of the mitochondrial intermembrane space. *Molecular biology and evolution 24* (2007), 1149–60.

[40] Gornicka, A., Bragoszewski, P., Chroscicki, P., Wenz, L.-S., Schulz, C., Rehling, P., and Chacinska, A. A discrete pathway for the transfer of intermembrane space proteins across the outer membrane of mitochondria. *Molecular Biology of the Cell 25* (2014).

[41] Habib, S. J., Waizenegger, T., Lech, M., Neupert, W., and Rapaport, D. Assembly of the TOB complex of mitochondria. *Journal of Biological Chemistry 280* (2005), 6434–6440.

[42] Hagn, F., Lagleder, S., Retzlaff, M., Rohrberg, J., Demmer, O., Richter, K., Buchner, J., and Kessler, H. Structural analysis of the interaction between hsp90 and the tumor suppressor protein p53. *Nat Struct Mol Biol 18* (2011), 1086–1093.

[43] Hartl, F. U. Molecular chaperones in cellular protein folding. *Nature 381* (1996), 571–580.

[44] Hasson, S. A., Damoiseaux, R., Glavin, J. D., Dabir, D. V., Walker, S. S., and Koehler, C. M. Substrate specificity of the TIM22 mitochondrial import pathway revealed with small molecule inhibitor of protein translocation. *Proceedings of the National Academy of Sciences 107* (2010), 9578–9583.

[45] He, L., Sharpe, T., Mazur, A., and Hiller, S. A molecular mechanism of chaperone-client recognition. *Science Advances 2* (2016), 1601625–1601625.

[46] Herndon, J. D., Claypool, S. M., and Koehler, C. M. The Taz1p transacylase is imported and sorted into the outer mitochondrial membrane via a membrane anchor domain. *Eukaryotic Cell 12* (2013), 1600–1608.

[47] Higurashi, T., Hines, J. K., Sahi, C., Aron, R., and Craig, E. A. Specificity of the j-protein sis1 in the propagation of 3 yeast prions. *Proceedings of the National Academy of Sciences 105* (2008), 16596–16601.

[48] Hiller, S., Garces, R. G., Malia, T. J., Orekhov, V. Y., and Wagner, G. Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science 321* (2009), 1206–1210.

[49] Hipp, M. S., Park, S.-H., and Hartl, F. U. Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends in Cell Biology 24* (2014), 506–514.

[50] Hocking, H. G., Zangger, K., and Madl, T. Studying the Structure and Dynamics of Biomolecules by Using Soluble Paramagnetic Probes. *ChemPhysChem 14* (2013), 3082–3094.

[51] Hynson, R. M. G., Duff, A. P., Kirby, N., Mudie, S., and Lee, L. K. Differential ultracentrifugation coupled to small-angle X-ray scattering on macromolecular complexes. *Journal of Applied Crystallography 48* (2015), 769–775.

[52] Ieva, R., Schrempp, S. G., Opaliński, Ł., Wollweber, F., Höss, P., Heisswolf, A. K., Gebert, M., Zhang, Y., Guiard, B., Rospert, S., Becker, T., Chacinska, A., Pfanner, N., and van der Laan, M. Mgr2 functions as lateral gatekeeper for preprotein sorting in the mitochondrial inner membrane. *Molecular Cell 56* (2014), 641–652.

[53] Jehle, S., Vollmar, B. S., Bardiaux, B., Dove, K. K., Rajagopal, P., Gonen, T., Oschkinat, H., and Klevit, R. E. N-terminal domain of B-crystallin provides a conformational switch for multimerization and structural heterogeneity. *Proceedings of the National Academy of Sciences 108* (2011), 6409–6414.

[54] Joachimiak, L. A., Walzthoeni, T., Liu, C. W., Aebersold, R., and Frydman, J. The structural basis of substrate recognition by the eukaryotic chaperonin TRiC/CCT. *Cell 159* (2014), 1042–1055.

[55] Jores, T., Klinger, A., Gross, L. E., Kawano, S., Flinner, N., Duchardt-Ferner, E., Wöhnert, J., Kalbacher, H., Endo, T., Schleiff, E., and Rapaport, D. Characterization of the targeting signal in mitochondrial β-barrel proteins. *Nature Communications 7* (2016), 12036.

[56] Kalisman, N., Schröder, G. F., and Levitt, M. The crystal structures of the eukaryotic chaperonin CCT reveal its functional partitioning. *Structure 21* (2013), 540–549.

[57] Kampinga, H. H., and Craig, E. A. The Hsp70 chaperone machinery: J-proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology 8* (2010), 579–592.

[58] Kang, P.-J., Ostermann, J., Shilling, J., Neupert, W., Craig, E. A., and Pfanner, N. Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. *Nature 348* (1990), 137–143.

[59] Karagöz, G. E., Duarte, A. M. S., Akoury, E., Ippel, H., Biernat, J., Morán Luengo, T., Radli, M., Didenko, T., Nordhues, B. A., Veprintsev, D. B., Dickey, C. A., Mandelkow, E., Zweckstetter, M., Boelens, R., Madl, T., and Rüdiger, S. G. D. Hsp90-tau complex reveals molecular basis for specificity in chaperone action. *Cell 156* (2014), 963–974.

[60] Kemper, C., Habib, S. J., Engl, G., Heckmeyer, P., Dimmer, K. S., and Rapaport, D. Integration of tail-anchored proteins into the mitochondrial outer membrane does not require any known import components. *Journal of Cell Science* (2008), 1990–1998.

[61] Kirschke, E., Goswami, D., Southworth, D., Griffin, P. R., and Agard, D. A. Glucocorticoid receptor function regulated by coordinated action of the Hsp90 and Hsp70 chaperone cycles. *Cell 157* (2014), 1685–1697.

[62] Kirstein, J., Schlothauer, T., Dougan, D. A., Lilie, H., Tischendorf, G., Mogk, A., Bukau, B., and Turgay, K. Adaptor protein controlled oligomerization activates the aaa+ protein clpc. *The EMBO Journal 25* (2006), 1481–1491.

[63] Kityk, R., Kopp, J., Sinning, I., and Mayer, M. P. Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones. *Molecular Cell 48* (2012), 863–874.

[64] Koehler, C. M., Jarosch, E., Tokatlidis, K., Schmid, K., Schweyen, R. J., and Schatz, G. Import of Mitochondrial Carriers Mediated by Essential Proteins of the Intermembrane Space. *Science 279* (1998), 369–373.

[65] Koldewey, P., Stull, F., Horowitz, S., Martin, R., and Bardwell, J. C. Forces Driving Chaperone Action. *Cell 166* (2016), 369–379.

[66] Korzhnev, D. M., Kloiber, K., Kanelis, V., Tugarinov, V., and Kay, L. E. Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *J. Am. Chem. Soc. 126* (2004), 3964–3973.

[67] Kupce, E., Boyd, J., and Campbell, I. D. Short selective pulses for biochemical applications. *Journal of Magnetic Resonance, Series B 106* (1995), 300–303.

[68] Kutik, S., Stojanovski, D., Becker, L., Becker, T., Meinecke, M., Krüger, V., Prinz, C., Meisinger, C., Guiard, B., Wagner, R., Pfanner, N., and Wiedemann, N. Dissecting Membrane Insertion of Mitochondrial B-Barrel Proteins. *Cell 132* (2008), 1011–1024.

[69] Lasorsa, F. M., Pinton, P., Palmieri, L., Fiermonte, G., Rizzuto, R., and Palmieri, F. Recombinant expression of the ca2+-sensitive aspartate/glutamate carrier increases mitochondrial atp production in agonist-stimulated chinese hamster ovary cells. *Journal of Biological Chemistry 278* (2003), 38686–38692.

[70] Leitner, A., Joachimiak, L. A., Bracher, A., Mönkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B., Ludtke, S., Chiu, W., Hartl, F. U., Aebersold, R., and Frydman, J. The molecular architecture of the eukaryotic chaperonin TRiC/CCT. *Structure 20* (2012), 814–825.

[71] Li, J., Soroka, J., and Buchner, J. The Hsp90 chaperone machinery: Conformational dynamics and regulation by co-chaperones. *Biochimica et Biophysica Acta - Molecular Cell Research 1823* (2012), 624–635.

[72] Liang, F.-C., Kroon, G., McAvoy, C. Z., Chi, C., Wright, P. E., and Shan, S.-o. Conformational dynamics of a membrane protein chaperone enables spatially regulated substrate capture and release. *Proceedings of the National Academy of Sciences 113* (2016), 1615–1624.

[73] Libich, D. S., Tugarinov, V., and Clore, G. M. Intrinsic unfoldase/foldase activity of the chaperonin GroEL directly demonstrated using multinuclear relaxation-based NMR. *Proceedings of the National Academy of Sciences 112* (2015), 8817–8823.

[74] Lionaki, E., Lousa, C. D. M., Baud, C., Vougioukalaki, M., Panayotou, G., and Tokatlidis, K. The essential function of Tim12 in vivo is ensured by the assembly interactions of its C-terminal domain. *Journal of Biological Chemistry 283* (2008), 15747–15753.

[75] Liu, J., Mei, Z., Li, N., Qi, Y., Xu, Y., Shi, Y., Wang, F., Lei, J., and Gao, N. Structural dynamics of the MecA-ClpC complex: A type II AAA+ protein unfolding machine. *Journal of Biological Chemistry 288* (2013), 17597–17608.

[76] Lopez, T., Dalton, K., and Frydman, J. The mechanism and function of group II chaperonins. *Journal of Molecular Biology 427* (2016), 2919–2930.

[77] Lu, H., Golovanov, A. P., Alcock, F., Grossmann, J. G., Allen, S., Lian, L.-Y., and Tokatlidis, K. The structural basis of the TIM10 chaperone assembly. *The Journal of biological chemistry 279* (2004), 18959–18966.

[78] Mainz, A., Peschek, J., Stavropoulou, M., Back, K. C., Bardiaux, B., Asami, S., Prade, E., Peters, C., Weinkauf, S., Buchner, J., and Reif, B. The chaperone αB-crystallin uses different interfaces to capture an amorphous and an amyloid client. *Nature Structural & Molecular Biology 22* (2015), 898–905.

[79] Mayer, M. P., and Le Breton, L. Hsp90: Breaking the symmetry. *Molecular Cell 58* (2015), 8–20.

[80] Melero, R., Moro, F., Perez-Calvo, M. A., Perales-Calvo, J., Quintana-Gallardo, L., Llorca, O., Muga, A., and Valpuesta, J. M. Modulation of the chaperone dnak allosterism by the nucleotide exchange factor grpe. *Journal of Biological Chemistry 290* (2015), 10083–10092.

[81] Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., Hell, K., and Herrmann, J. M. A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell 121* (2005), 1059–1069.

[82] Milenkovic, D., Gabriel, K., Guiard, B., Schulze-Specking, A., Pfanner, N., and Chacinska, A. Biogenesis of the essential Tim9-Tim10 chaperone complex of mitochondria: Site-specific recognition of cysteine residues by the intermembrane space receptor Mia40. *Journal of Biological Chemistry 282* (2007), 22472–22480.

[83] Milenkovic, D., Ramming, T., Müller, J. M., Wenz, L.-s., Gebert, N., Schulze-Specking, A., Stojanovski, D., Rospert, S., and Chacinska, A. Identification of the signal directing Tim9 and Tim10 into the intermembrane space of mitochondria. *Molecular biology of the cell 20* (2009), 2530–9.

[84] Mokranjac, D., and Neupert, W. Cell biology: Architecture of a protein entry gate. *Nature 528* (2015), 201–202.

[85] Morgner, N., and Robinson, C. V. Massign: An assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies. *Analytical Chemistry 84* (2012), 2939–2948.

[86] Nakatsukasa, K., Huyer, G., Michaelis, S., and Brodsky, J. L. Dissecting the er-associated degradation of a misfolded polytopic membrane protein. *Cell 132* (2007), 101–112.

[87] Palmieri, F. Mitochondrial transporters of the SLC25 family and associated diseases : a review. *Journal of Inherited Metabolic Diseases* (2014).

[88] Papić, D., Krumpe, K., Dukanovic, J., Dimmer, K. S., and Rapaport, D. Multispan mitochondrial outer membrane protein ugo1 follows a unique mim1-dependent import pathway. *The Journal of Cell Biology 194* (2011), 397–405.

[89] Park, S. J., Kostic, M., and Dyson, H. J. Dynamic interaction of Hsp90 with its client protein p53. *Journal of Molecular Biology 411* (2011), 158–173.

[90] Paschen, S. A., Rothbauer, U., Kaldi, K., Bauer, M. F., Neupert, W., and Brunner, M. The role of the TIM8-13 complex in the import of Tim23 into mitochondria. *The EMBO Journal 19* (2000), 6392–6400.

[91] Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trézéguet, V., Lauquin, G. J.-M., and Brandolin, G. Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature 426* (2003), 39–44.

[92] Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., Gobbo, A., Gordon, E., Huet, J., Kieffer, J., Lentini, M., Mattenet, M., Morawe, C., Mueller-Dieckmann, C., Ohlsson, S., Schmid, W., Surr, J., Theveneau, P., Zerrad, L., and McSweeney, S. Upgraded esrf bm29 beamline for saxs on macromolecules in solution. *Journal of Synchrotron Radiation 20* (2013), 660–664.

[93] Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V., and Svergun, D. I. New developments in the atsas program package for small-angle scattering data analysis. *Journal of Applied Crystallography 45* (2012), 342–350.

[94] Popov-Čeleketić, J., Waizenegger, T., and Rapaport, D. Mim1 functions in an oligomeric form to facilitate the integration of tom20 into the mitochondrial outer membrane. *Journal of Molecular Biology 376* (2008), 671–680.

[95] Qiu, J., Wenz, L. S., Zerbes, R. M., Oeljeklaus, S., Bohnert, M., Stroud, D. A., Wirth, C., Ellenrieder, L., Thornton, N., Kutik, S., Wiese, S., Schulze-Specking, A., Zufall, N., Chacinska, A., Guiard, B., Hunte, C., Warscheid, B., Van Der Laan, M., Pfanner, N., Wiedemann, N., and Becker, T. Coupling of mitochondrial import and export translocases by receptor-mediated supercomplex formation. *Cell 154* (2013), 596–608.

[96] Rahman, B., Kawano, S., Yunoki-Esaki, K., Anzai, T., and Endo, T. Nmr analyses on the interactions of the yeast tim50 c-terminal region with the presequence and tim50 core domain. *FEBS Letters 588* (2014), 678–684.

[97] Rapaport, D., and Neupert, W. Biogenesis of Tom40, Core Component of the TOM Complex of Mitochondria. *The Journal of Cell Biology 146* (1999).

[98] Roesch, K., Curran, S. P., Tranebjaerg, L., and Koehler, C. M. Human deafness dystonia syndrome is caused by a defect in assembly of the DDP1/TIMM8a-TIMM13 complex. *Human molecular genetics 11* (2002), 477–486.

[99] Roise, D., Horvath, S. J., Tomich, J. M., Richards, J. H., and Schatz, G. A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphiphilic helix and perturb natural and artificial phospholipid bilayers. *The EMBO Journal 5* (1986), 1327–1334.

[100] Round, A., Felisaz, F., Fodinger, L., Gobbo, A., Huet, J., Villard, C., Blanchet, C. E., Pernot, P., McSweeney, S., Roessle, M., Svergun, D. I., and Cipriani, F. Biosaxs sample changer: a robotic sample changer for rapid and reliable high-throughput x-ray solution scattering experiments. *Acta Crystallographica Section D: Biological Crystallography 71* (2015), 67–75.

[101] Rousseau, F., Serrano, L., and Schymkowitz, J. W. H. How evolutionary pressure against protein aggregation shaped chaperone specificity. *Journal of Molecular Biology 355* (2006), 1037–1047.

[102] Ruprecht, J. J., Hellawell, A. M., Harding, M., Crichton, P. G., McCoy, A. J., and Kunji, E. R. S. Structures of yeast mitochondrial adp/atp carriers support a domain-based alternating-access transport mechanism. *Proceedings of the National Academy of Sciences 111* (2014), E426–E434.

[103] Saio, T., Guan, X., Rossi, P., Economou, A., and Kalodimos, C. G. Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone. *Science 344* (2014), 1250494.

[104] Schanda, P., Forge, V., and Brutscher, B. HET-SOFAST NMR for fast detection of structural compactness and heterogeneity along polypeptide chains. *Magn. Reson. Chem. 44 Spec No* (2006), S177–84.

[105] Schanda, P., Kupce, E., Brutscher, B., Kupce, E., and Brutscher, B. SOFAST-HMQC experiments for recording two-dimensional deteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR 33* (2005), 199–211.

[106] Schiffrin, B., Calabrese, A. N., Devine, P. W. A., Harris, S. A., Ashcroft, A. E., Brockwell, D. J., and Radford, S. E. Skp is a multivalent chaperone of outer-membrane proteins. *Nature Structural & Molecular Biology 23* (2016), 786–793.

[107] Schlecht, R., Erbse, A. H., Bukau, B., and Mayer, M. P. Mechanics of Hsp70 chaperones enables differential interaction with client proteins. *Nature Structural & Molecular Biology 18* (2011), 345–351.

[108] Schmitt, E. K., Riwanto, M., Sambandamurthy, V., Roggo, S., Miault, C., Zwingelstein, C., Krastel, P., Noble, C., Beer, D., Rao, S. P. S., Au, M., Niyomrattanakit, P., Lim, V., Zheng, J., Jeffery, D., Pethe, K., and Camacho, L. R. The natural product cyclomarin kills mycobacterium tuberculosis by targeting the ClpC1 subunit of the caseinolytic protease. *Angewandte Chemie - International Edition 50* (2011), 5889–5891.

[109] Schröder, H., Langer, T., Hartl, F. U., and Bukau, B. Dnak, dnaj and grpe form a cellular chaperone machinery capable of repairing heat-induced protein damage. *The EMBO Journal 12* (1993), 4137–4144.

[110] Sekhar, A., Rosenzweig, R., Bouvignies, G., and Kay, L. E. Hsp70 biases the folding pathways of client proteins. *Proceedings of the National Academy of Sciences 113* (2016), 2794–2801.

[111] Shiau, A. K., Harris, S. F., Southworth, D. R., and Agard, D. A. Structural Analysis of E. coli hsp90 Reveals Dramatic Nucleotide-Dependent Conformational Rearrangements. *Cell 127* (2006), 329–340.

[112] Shiota, T., Imai, K., Qiu, J., Hewitt, V. L., Tan, K., Shen, H.-H., Sakiyama, N., Fukasawa, Y., Hayat, S., Kamiya, M., Elofsson, A., Tomii, K., Horton, P., Wiedemann, N., Pfanner, N., Lithgow, T., and Endo, T. Molecular architecture of the active mitochondrial protein gate. *Science 349* (2015), 1544–1548.

[113] Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F., and Moarefi, I. Structure of the Molecular Chaperone Prefoldin. *Cell 103* (2000), 621–632.

[114] Skjærven, L., Cuellar, J., Martinez, A., and Valpuesta, J. M. Dynamics, flexibility, and allostery in molecular chaperonins. *FEBS Letters 589* (2015), 2522–2532.

[115] Sobott, F., Hernández, H., McCammon, M. G., Tito, M. A., and Robinson, C. V. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Analytical Chemistry 74* (2002), 1402–1407.

[116] Sriram, M., Osipiuk, J., Freeman, B., Morimoto, R., and Joachimiak, A. Human hsp70 molecular chaperone binds two calcium ions within the atpase domain. *Structure 5* (1997), 403–414.

[117] Stengel, F., Baldwin, A. J., Painter, A. J., Jaya, N., Basha, E., Kay, L. E., Vierling, E., Robinson, C. V., and Benesch, J. L. P. Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proceedings of the National Academy of Sciences 107* (2010), 2007–2012.

[118] Street, T. O., Lavery, L. A., and Agard, D. A. Substrate Binding Drives Large-Scale Conformational Changes in the Hsp90 Molecular Chaperone. *Molecular Cell 42* (2011), 96–105.

[119] Stroud, D. A., Becker, T., Qiu, J., Stojanovski, D., Pfannschmidt, S., Wirth, C., Huntea, C., Guiard, B., Meisinger, C., Pfanner, N., and Wiedemann, N. Biogenesis of mitochondrial b-barrel proteins: the POTRA domain is involved in precursor release from the SAM complex. *Molecular Cell 22* (2011), 2823–2833.

[120] Suh, W.-C., Burkholder, W. F., Lu, C. Z., Zhao, X., Gottesman, M. E., and Gross, C. A. Interaction of the hsp70 molecular chaperone, dnak, with its cochaperone dnaj. *Proceedings of the National Academy of Sciences 95* (1998), 15223–15228.

[121] Sun, H., Kay, L. E., and Tugarinov, V. An Optimized Relaxation-Based Coherence Transfer NMR Experiment for the Measurement of Side-Chain Order in Methyl-Protonated, Highly Deuterated Proteins. *J. Phys. Chem. B 115* (2011), 14878–14884.

[122] Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography 25* (1992), 495–503.

[123] Svergun, D. I., Petoukhov, M. V., and Koch, M. H. Determination of domain structure of proteins from x-ray solution scattering. *Biophysical Journal 80* (2001), 2946–2953.

[124] Szyperski, T., Pellecchia, M., Wall, D., Georgopoulos, C., and Wüthrich, K. Nmr structure determination of the escherichia coli dnaj molecular chaperone: secondary structure and backbone fold of the n-terminal region (residues 2-108) containing the highly conserved j domain. *Proceedings of the National Academy of Sciences 91* (1994), 11343–11347.

[125] Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K. D., Karras, G. I., and Lindquist, S. Quantitative analysis of Hsp90-client interactions reveals principles of substrate recognition. *Cell 150* (2012), 987–1001.

[126] Taipale, M., Tucker, G., Peng, J., Krykbaeva, I., Lin, Z.-Y., Larsen, B., Choi, H., Berger, B., Gingras, A.-C., and Lindquist, S. A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell 158* (2014), 434–448.

[127] Taylor, R. C., and Dillin, A. Aging as an event of proteostasis collapse. *Cold Spring Harbor Perspectives in Biology 3* (2011).

[128] Thoma, J., Burmann, B. M., Hiller, S., and Müller, D. J. Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. *Nature Structural & Molecular Biology 22* (2015), 795–802.

[129] Trentini, D. B., Suskiewicz, M. J., Deszcz, L., and Mechtler, K. Arginine phosphorylation marks proteins for degradation by the ClpCP protease. *Nature 539* (2015), 1–41.

[130] Truscott, K. N., Kovermann, P., Geissler, A., Merlin, A., Meijer, M., Driessen, A. J. M., Rassow, J., Pfanner, N., and Wagner, R. A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by tim23. *Nat Struct Mol Biol 8* (2001), 1074–1082.

[131] van den Heuvel, R. H. H., van Duijn, E., Mazon, H., Synowsky, S. A., Lorenzen, K., Versluis, C., Brouns, S. J. J., Langridge, D., van der Oost, J., Hoyes, J., and Heck, A. J. R. Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Analytical Chemistry 78* (2006), 7473–7483.

[132] Vasiljev, A., Ahting, U., Nargang, F. E., Go, N. E., Habib, S. J., Kozany, C., Panneels, V., Sinning, I., Prokisch, H., Neupert, W., Nussberger, S., and Rapaport, D. Reconstituted TOM core complex and Tim9/Tim10 complex of mitochondria are sufficient for translocation of the ADP/ATP carrier across membranes. *Molecular Biology of the Cell 15* (2004), 1445–1458.

[133] Vasudevan, D., Rao, S. P. S., and Noble, C. G. Structural basis of mycobacterial inhibition by Cyclomarin A. *Journal of Biological Chemistry 288* (2013), 30883–30891.

[134] Verba, K. A., Wang, R. Y.-R., Arakawa, A., Liu, Y., Shirouzu, M., Yokoyama, S., and Agard, D. A. Atomic structure of hsp90-cdc37-cdk4 reveals that hsp90 traps and stabilizes an unfolded kinase. *Science 352* (2016), 1542–1547.

[135] Vergnolle, M. a. S., Baud, C., Golovanov, A. P., Alcock, F., Luciano, P., Lian, L. Y., and Tokatlidis, K. Distinct domains of small Tims involved in subunit interaction and substrate recognition. *Journal of Molecular Biology 351* (2005), 839–849.

[136] Vial, S., Lu, H., Allen, S., Savory, P., Thornton, D., Sheehan, J., and Tokatlidis, K. Assembly of Tim9 and Tim10 into a functional chaperone. *The Journal of Biological Chemistry 277* (2002), 36100–36108.

[137] Vögtle, F. N., Wortelkamp, S., Zahedi, R., Becker, D., Leidhold, C., Gevaert, K., Kellermann, J., Voos, W., Sickmann, A., Pfanner, N., and Meisinger, C. Global analysis of the mitochondrial n-proteome identifies a processing peptidase critical for protein stability. *Cell 139* (2009), 428–439.

[138] Volkov, V. V., and Svergun, D. I. Uniqueness of *ab initio* shape determination in small-angle scattering. *Journal of Applied Crystallography 36* (2003), 860–864.

[139] Wagner, K., Gebert, N., Guiard, B., Brandner, K., Truscott, K. N., Wiedemann, N., Pfanner, N., and Rehling, P. The assembly pathway of the mitochondrial carrier translocase involves four preprotein translocases. *Molecular and cellular biology 28* (2008), 4251–4260.

[140] Walton, T. a., Sandoval, C. M., Fowler, C. A., Pardi, A., and Sousa, M. C. The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proceedings of the National Academy of Sciences of the United States of America 106* (2009), 1772–7.

[141] Walton, T. a., and Sousa, M. C. Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Molecular cell 15* (2004), 367–74.

[142] Wang, F., Mei, Z., Qi, Y., Yan, C., Hu, Q., Wang, J., and Shi, Y. Structure and mechanism of the hexameric MecA–ClpC molecular machine. *Nature 471* (2011), 331–335.

[143] Webb, C. T., Gorman, M. a., Lazarou, M., Ryan, M. T., and Gulbis, J. M. Crystal structure of the mitochondrial chaperone TIM9.10 reveals a six-bladed alpha-propeller. *Molecular cell 21* (2006), 123–33.

[144] Weissman, J. S., Hohl, C. M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H. R., Fenton, W. A., and Horwich, A. L. Mechanism of groel action: Productive release of polypeptide from a sequestered position under groes. *Cell 83* (1995), 577 – 587.

[145] Wiedemann, N., and Pfanner, N. Mitochondrial Machineries for Protein Import and Assembly. *Annual Review of Biochemistry 86* (2017), 060815–014352.

[146] Wiedemann, N., Pfanner, N., and Ryan, M. T. The three modules of ADP/ATP carrier cooperate in receptor recruitment and translocation into mitochondria. *EMBO Journal 20* (2001), 951–960.

[147] Wiedemann, N., Truscott, K. N., Pfannschmidt, S., Guiard, B., Meisinger, C., and 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. *Journal of Biological Chemistry 279* (2004), 18188–18194.

[148] Yam, A. Y., Xia, Y., Lin, H.-T. J., Burlingame, A., Gerstein, M., and Frydman, J. Defining the tric/cct interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nat Struct Mol Biol 15* (2008), 1255–1262.

[149] Yamano, K., Yatsukawa, Y.-i., Esaki, M., Hobbs, A. E. A., Jensen, R. E., and Endo, T. Tom20 and tom22 share the common signal recognition pathway in mitochondrial protein import. *Journal of Biological Chemistry* (2008), 3799–3807.

[150] Young, J. C., Agashe, V. R., Siegers, K., and Hartl, F. U. Pathways of chaperone-mediated protein folding in the cytosol. *Nature Reviews Molecular Cell Biology 5* (2004), 781–791.

[151] Zara, V., Ferramosca, A., Robitaille-Foucher, P., Palmieri, F., and Young, J. C. Mitochondrial carrier protein biogenesis: role of the chaperones Hsc70 and Hsp90. *The Biochemical journal 419* (2009), 369–375.

[152] Zara, V., Palmieri, F., Mahlke, K., and Pfanner, N. The cleavable presequence is not essential for import and assembly of the phosphate carrier of mammalian mitochondria but enhances the specificity and efficiency of import. *The Journal of biological chemistry 267* (1992), 12077–81.

[153] Zhang, S., Chen, J., Shi, W., Cui, P., Zhang, J., Cho, S., Zhang, W., and Zhang, Y. Mutation in clpc1 encoding an atp-dependent atpase involved in protein degradation is associated with pyrazinamide resistance in mycobacterium tuberculosis. *Emerging Microbes & Infections 6* (2017), e8–.

[154] Zhuravleva, A., Clerico, E. M., and Gierasch, L. M. An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. *Cell 151* (2012), 1296–1307.

[155] Zhuravleva, A., and Gierasch, L. M. Substrate-binding domain conformational dynamics mediate Hsp70 allostery. *Proceedings of the National Academy of Sciences 112* (2015), 2865–2873.

[156] Zuiderweg, E. R. P., Hightower, L. E., and Gestwicki, J. E. The remarkable multivalency of the Hsp70 chaperones. *Cell Stress and Chaperones 22* (2017), 173–189.

[157] Żwirowski, S., Kłosowska, A., Obuchowski, I., Nillegoda, N. B., Piróg, A., Ziętkiewicz, S., Bukau, B., Mogk, A., and Liberek, K. Hsp70 displaces small heat shock proteins from aggregates to initiate protein refolding. *The EMBO Journal 36* (2017), 783–796.

Chapter 7

Glossary

Glossary

AAC ADP/ATP carrier 36, 38, 60–69, 72–74, 93, 94, 96, 131

ACD α-crystallin domain 24, 131

ADP adenosine diphosphate 14, 131

AI alanine, isoleucine 61, 131

ALV alanine, leucine, valine 55, 57, 58, 61, 64–66, 70, 73, 131

ATP adenosine triphosphate 7, 9, 12–27, 29, 41, 131

AUC analytical ultracentrifugation 1, 60, 80, 81, 93, 96, 100, 101, 131

CD circular dichroism 34, 72, 73, 93, 131

CDN [^{13}C,^{2}H,^{15}N] 74, 131

CPMG Carr-Purcell-Meiboom-Gill sequence 53, 55, 56, 66, 67, 73, 74, 131

CTD C-terminal domain 16, 24, 25, 131

CV column volume 107, 109, 110, 131

DMSO dimethyl sulfoxide 110, 115, 131

DNA deoxyribonucleic acid 131

DOSY diffusion-ordered spectroscopy 55, 92, 96, 111, 131

DTT dithiothreitol 107, 110, 131

EDTA ethylenediaminetetraacetic acid 107, 131

EM electron microscopy 16, 18, 19, 24, 36, 41, 131

EPR electron paramagnetic resonance 110, 131

ESI electrospray ionization 117, 131

FID free indutcion decay 114, 131

FUV far ultraviolet 117, 131

GGC GDP/GTP carrier 58–70, 72–74, 76–79, 85–87, 92–97, 100, 102, 131

hetNOE heteronuclear Overhauser effect 52, 58, 111, 131

HIV human immunodeficiency virus 19, 20, 131

HPLC high pressure liquid chromatography 116, 131

Hsp heat shock protein 131

IDP intrinsically disordered protein 18, 131

IMS intermembrane space 27–32, 38, 79, 103, 131

IPTG isopropyl β-D-1-thiogalactopyranoside 107, 109, 131

LHCP light harvesting chlorophyll a/b binding proteins 26, 131

LID lid domain 14, 15, 131

LV leucine, valine 61, 65–67, 69, 71, 131

MALS multi angle light scattering 131

MC mitochondrial carriers 27, 29–32, 35, 36, 131

MCS multiple cloning site 107, 131

MD middle domain 16, 131

MOM mitochondrial outer membrane 131

MS mass spectrometry 117, 131

MTSL S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate 68, 69, 95, 110, 113, 131

NBD nucleotide binding domain 14, 15, 131

NEF nucleotide exchange factor 14, 15, 131

NiNTA nickel nitrilotriacetic acid resin 50, 53, 59, 60, 63, 70, 77, 81, 92, 93, 96, 107, 131

NMR nuclear magnetic resonance 1, 9, 15, 24, 34, 36, 38, 41, 50–58, 60–65, 67, 68, 70–74, 77, 80–87, 91, 93–98, 100–102, 131

NOE nuclear Overhauser effect 83, 131

NOESY nuclear Overhauser effect spectroscopy 131

NTD N-terminal domain 16, 17, 21–24, 131

OD_{600} optical density at a wavelength of 600 nm 107, 109, 131

OMP outer membrane protein 25, 26, 131

PAM protein import motor 29, 131

POTRA polypeptide-transport-associated 32, 131

PPI Peptidyl-prolyl isomerase domain 25, 131

PRE paramagnetic relaxation enhancement 68, 95, 131

R_1 longitudinal relaxation rate 113, 131

R_2 transverse relaxation rate 131

$R_{2,eff}$ effective spin relaxation rate 53, 66, 67, 112, 131

RBD ribosome binding domain 25, 131

SAM sorting and assembly machinery 27, 32, 36, 131

SAXS small-angle X-ray scattering 1, 24, 34, 58, 96, 97, 104, 131

SBD substrate binding domain 14, 15, 131

SDS PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis 81, 131

SEC size-exclusion chromatography 50, 52, 55, 60, 77, 131

sHSP small heat shock protein 23–25, 41, 131

sPRE solvent paramagnetic relaxation enhancement 67, 68, 131

T_1 longitudinal relaxation time 56, 58, 111, 131

T_2 transverse relaxation time 56, 58, 111, 131

TCEP tris(2-carboxyethyl)phosphine 52, 115, 131

TEV tobacco etch virus 50, 107, 109, 131

TIM10 Tim10 subunit in the hexameric complex 131

Tim10 isolated Tim10 131

Tim12 isolated Tim12 131

TIM813 hexameric complex made of three Tim8 and three Tim13 subunits 32, 33, 38, 39, 97, 131

TIM9 Tim9 subunit in the hexameric complex 131

Tim9 isolated Tim9 131

TIM910 hexameric complex made of three Tim9 and three Tim10 subunits 7, 9, 10, 13, 23, 24, 27–35, 37–40, 49, 53–87, 91–104, 131

TIM91012 hexameric complex made of Tim9, Tim10 and Tim12 subunits 10, 33, 40, 79–87, 100–103, 131

TOM translocase of the outer membrane 27, 29–32, 36, 131

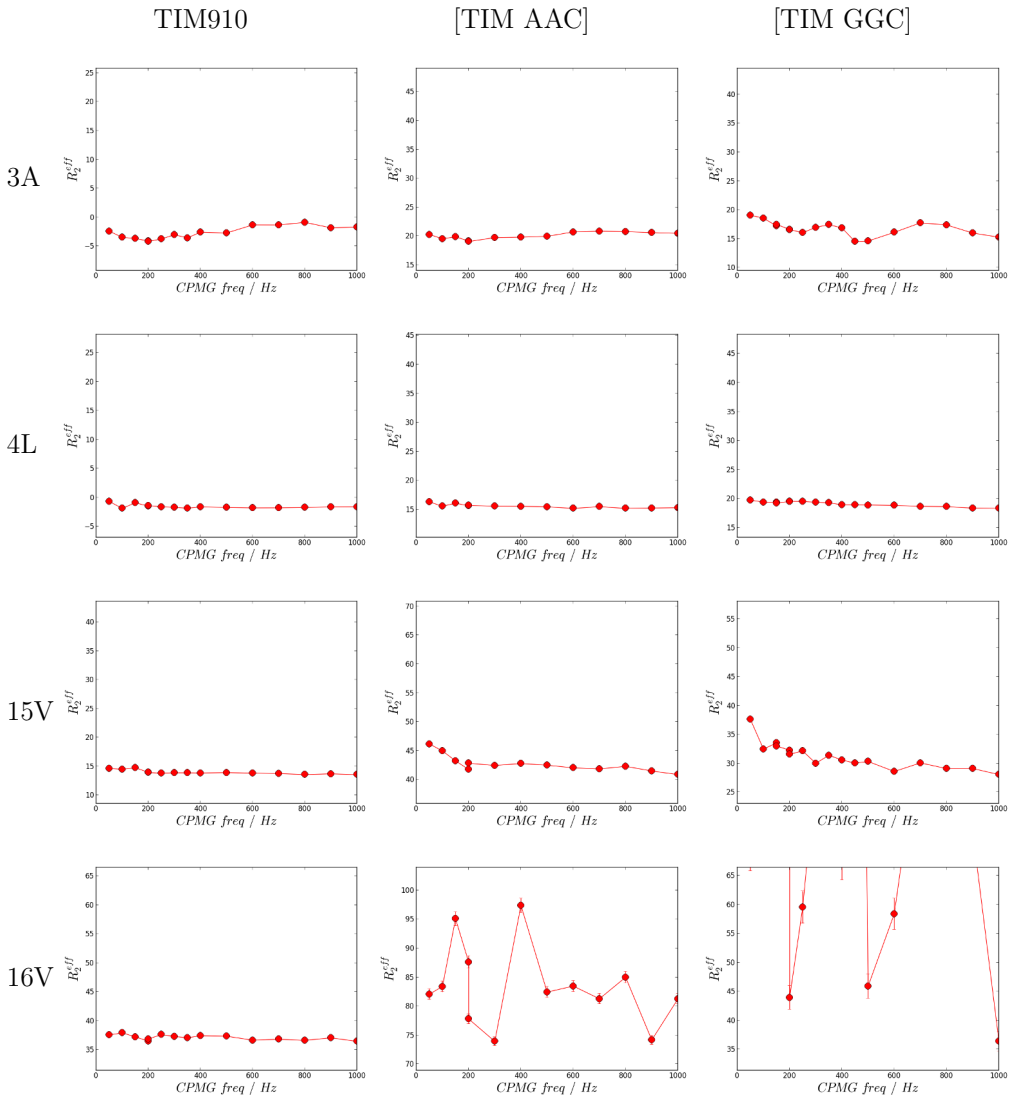
VDAC voltage-dependent anion channel 36, 70, 72, 98, 131

Appendices

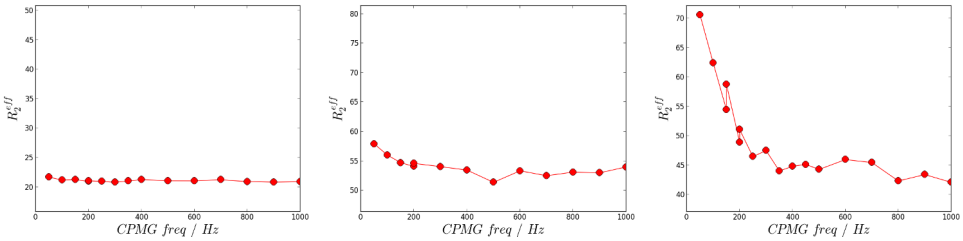
Appendix A

TIM910 supplementary figures

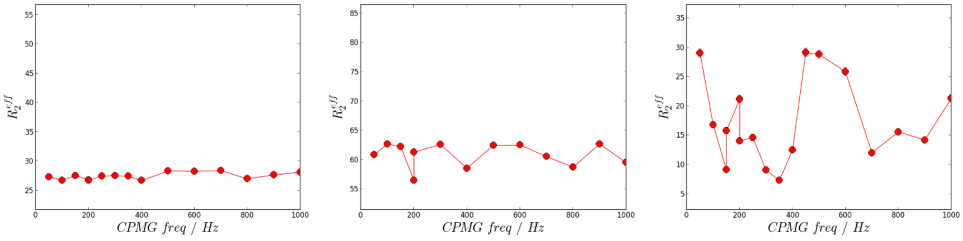
A.1 CPMG curves



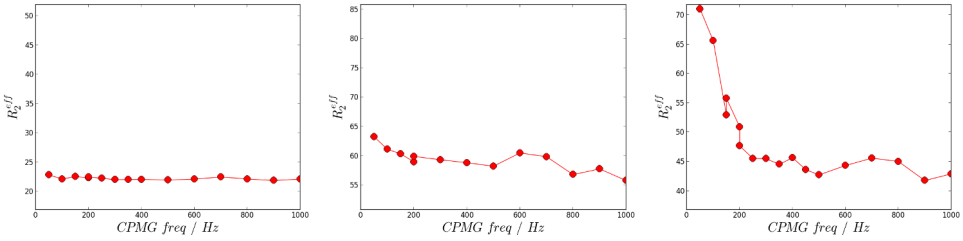
21A



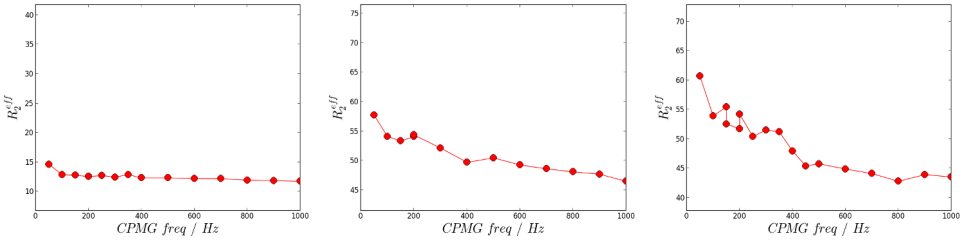
22A



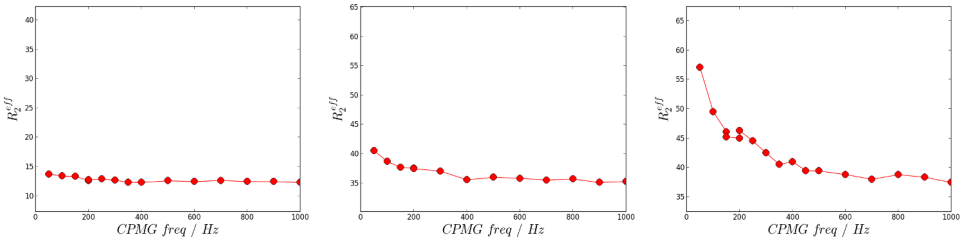
24A



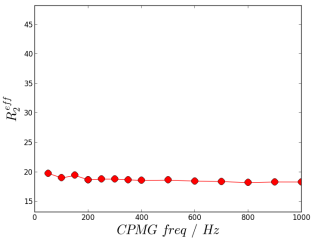
26L

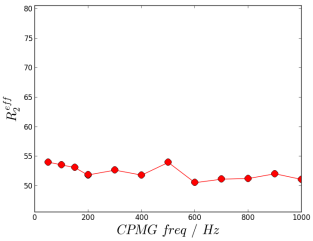


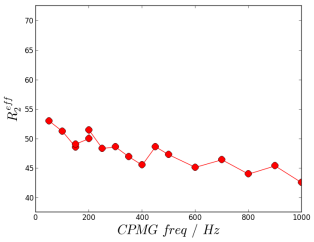
28L



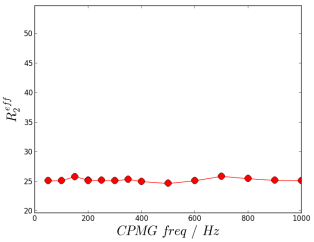
29V

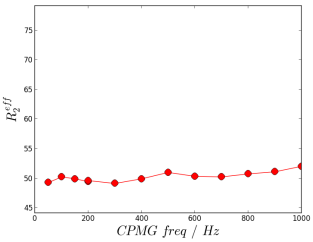


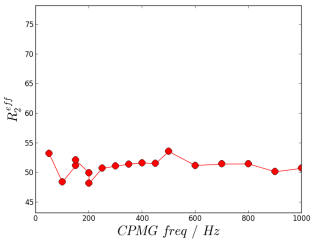




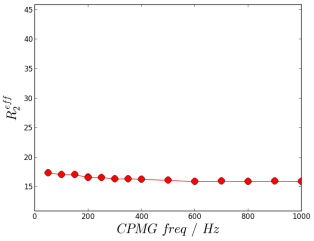
31L

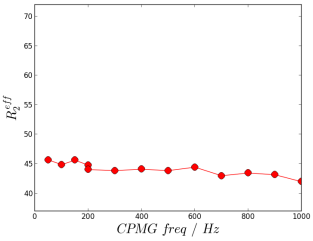


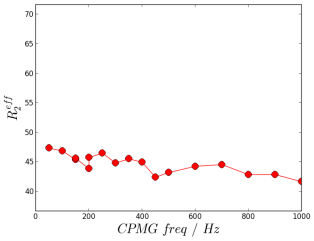




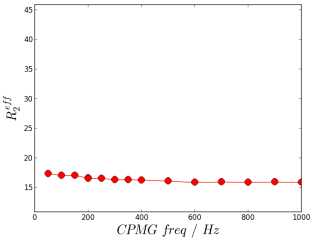
32V

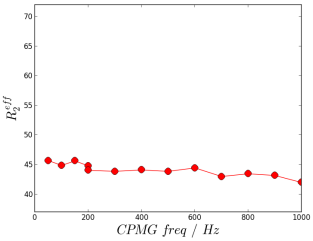


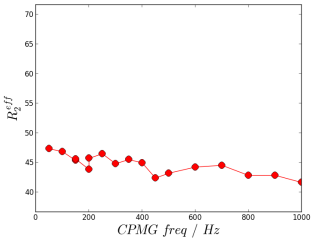




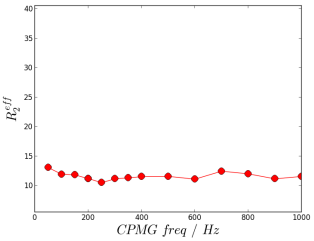
32V

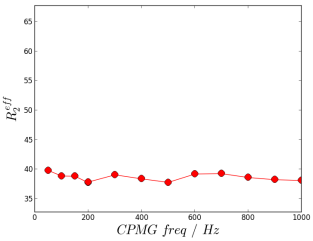


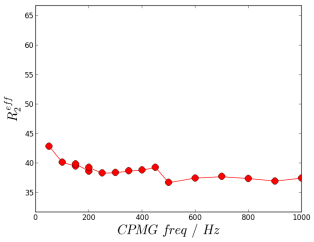




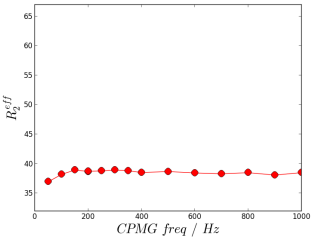
36L

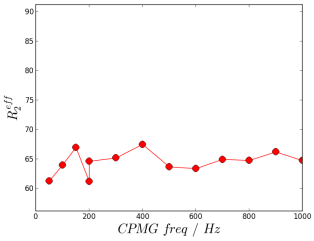


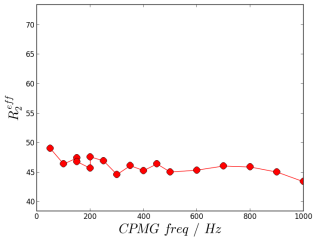




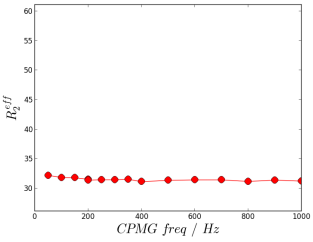
37V

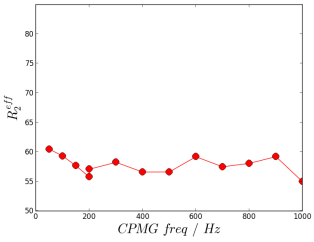


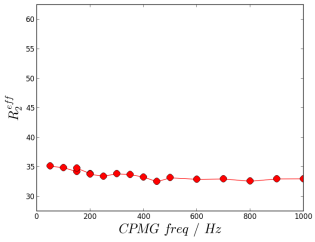




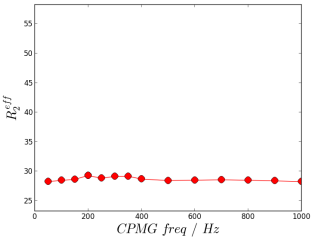
40V

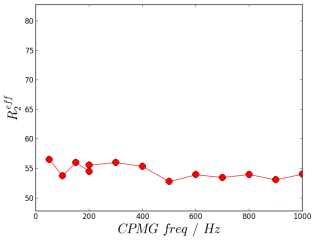


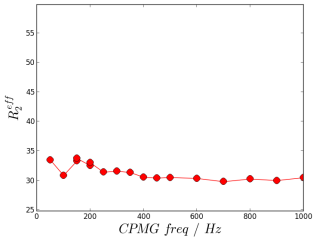




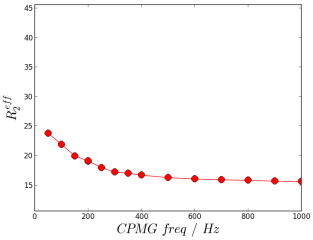
48L

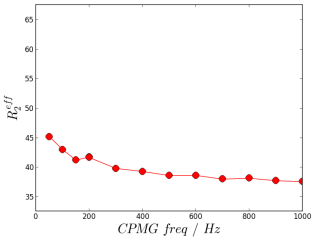


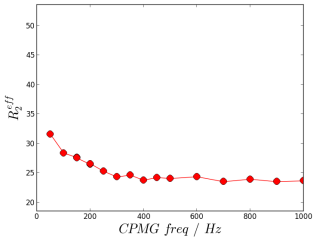




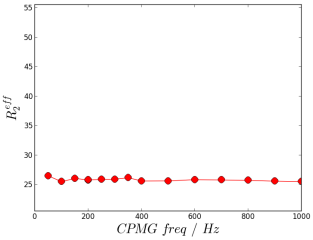
54L

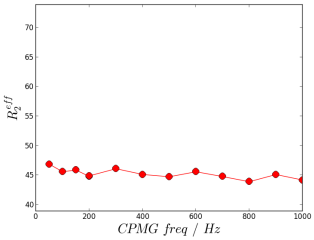


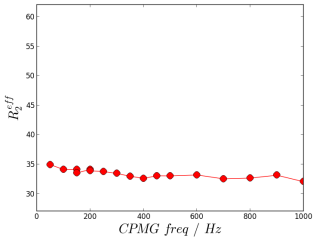




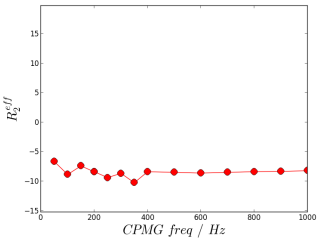
62L

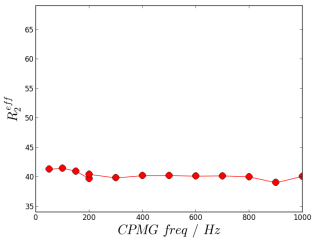


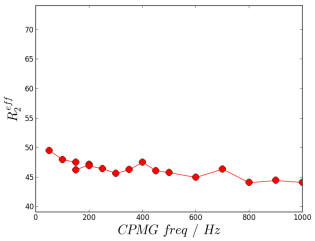




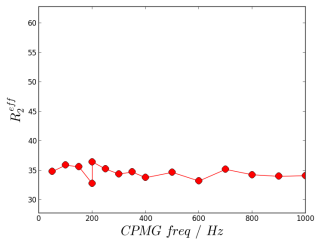
64L

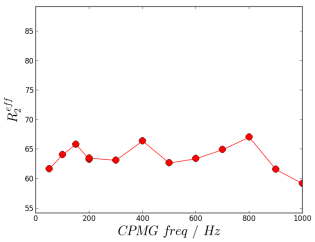


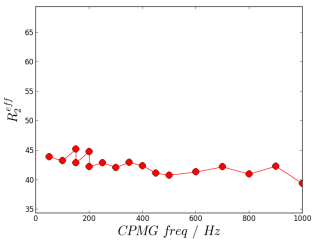




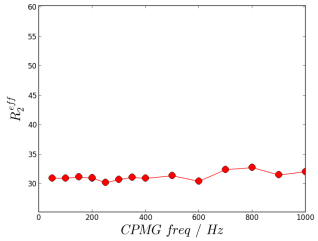
66V

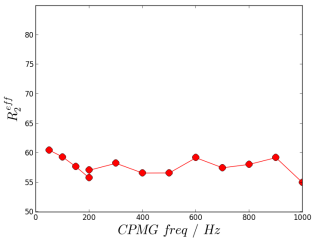


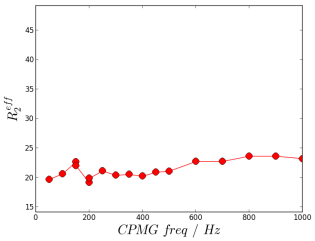




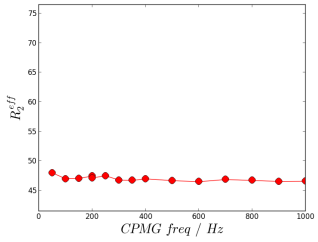
67A

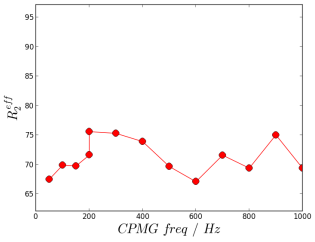


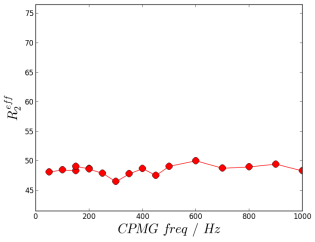




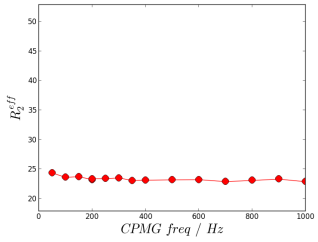
70V

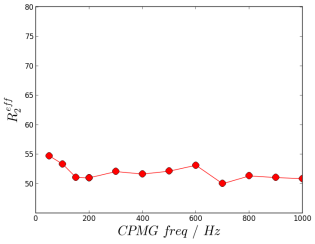


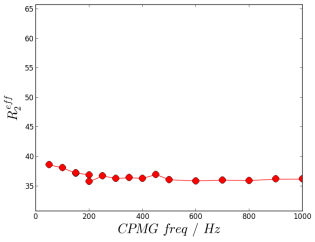




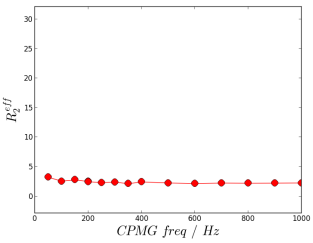
74V

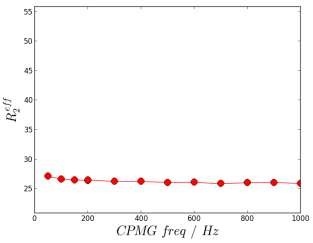


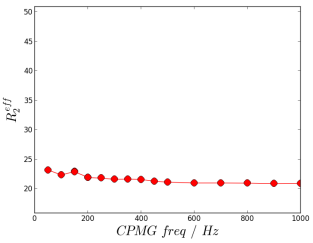




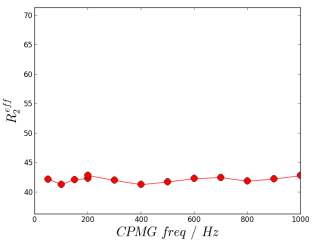
76V

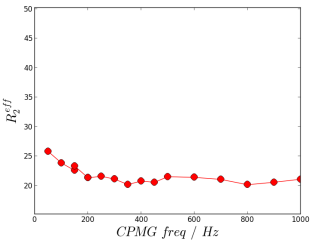




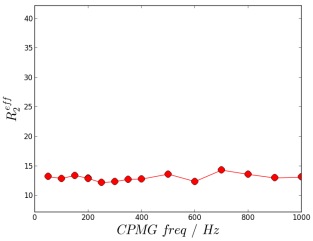


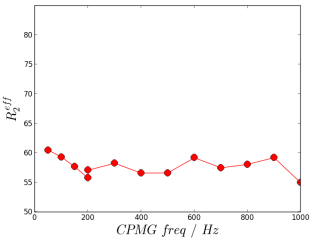
79A

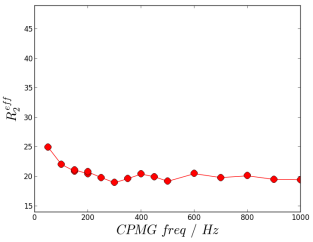




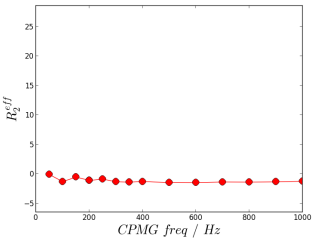
80A

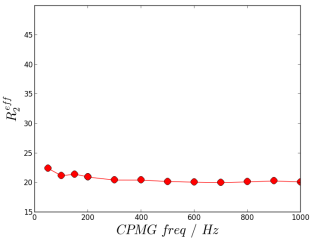


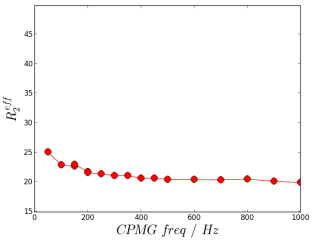




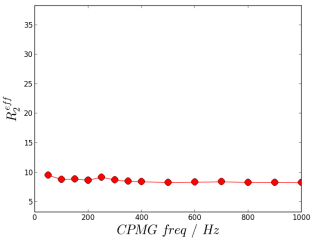
81L

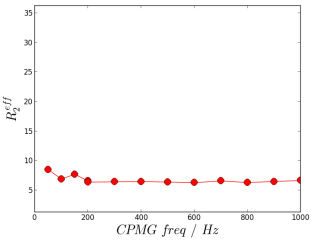


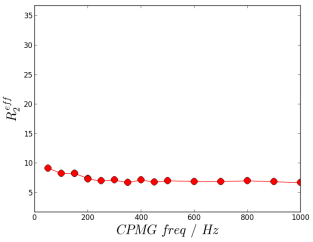




85L







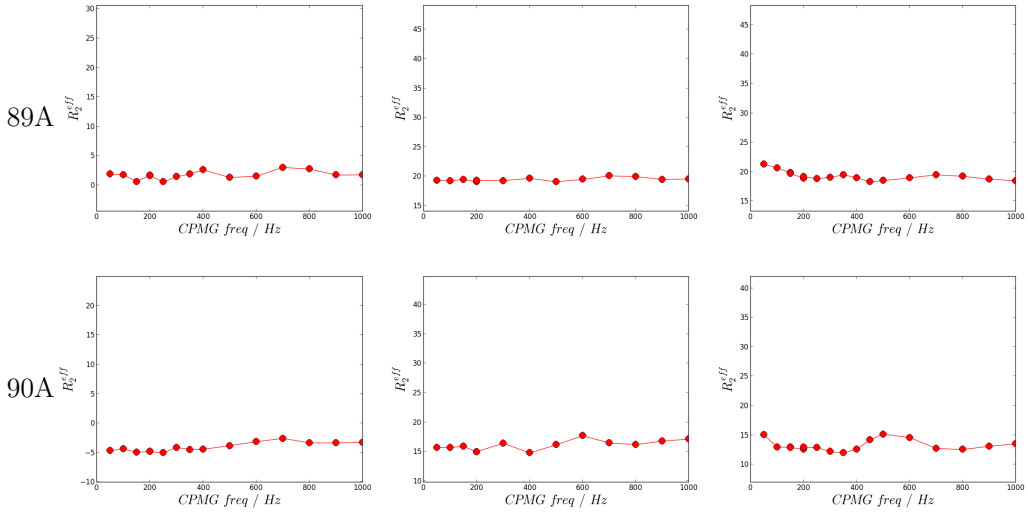


Figure A.1: **CPMG RD curves of all methyl residues in apo and GGC/AAC bound TIM910.** ALV methyl CPMG RD curves of apo TIM910, TIM910-AAC and TIM910-GGC complexes.

A.2 SDS PAGE TIM910-GGC complexes

A

Sample	% TIM	% GGC
1	79	21
2	79	21
3	83	17
4	76	24
5	83	17
6	63	37
7	80	20
8	79	21
9	80	20
10	82	18
11	82	18
12	82	18
13	82	18
14	83	17
15	72	28
16	71	29
17	78	22
18	74	26
19	75	25
20	73	27
21	78	22
22	83	17
23	78	22
24	82	18
25	81	19
26	76	24
27	78	22
Average	**78.2**	**21.8**
SDEV	4.7	4.7

B

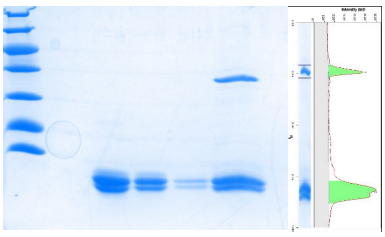


Figure A.2: **SDS-PAGE statistics for TIM910-GGC stoichiometry. A:** Average ratios of the integrated bands of 27 TIM910-GGC samples. **B:** Example peak integration of a TIM910-GGC SDS PAGE sample.

A.3 NMR intensity of TIM910/substrate complexes plotted on the structure

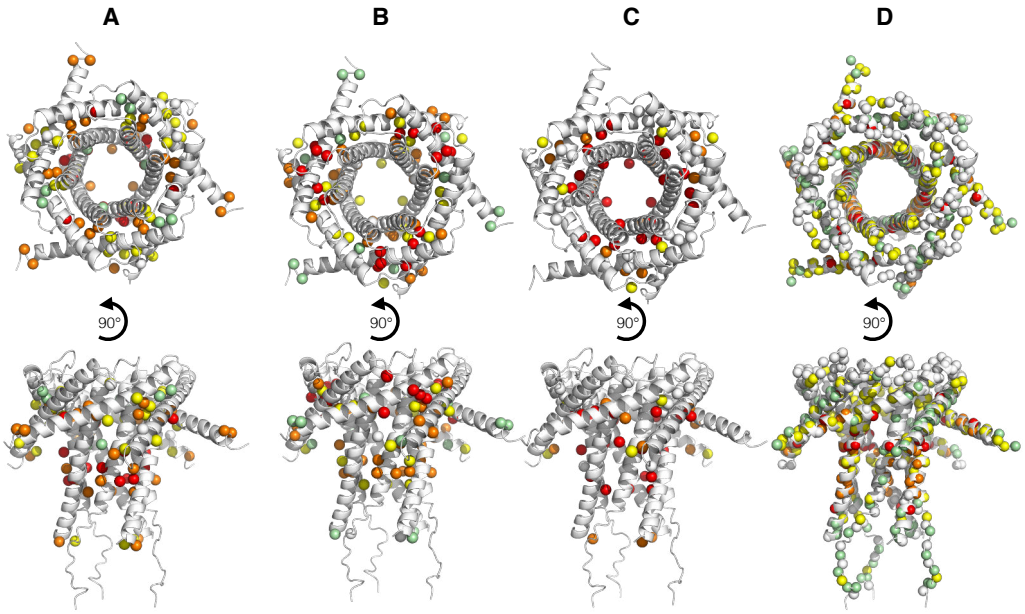


Figure A.3: **NMR peak intensities of TIM910/substrate complexes.** Peak height ratios between apo and substrate bound TIM910 plotted on an iTasser model of TIM910. **A:** ALV TIM910-GGC, **B:** ALV TIM910-AAC, **C:** LV TIM910-AAC_{short} and **D:** NH TIM910-AAC_{short}.

A.4 Cross-linking of TIM910 and VDAC peptide

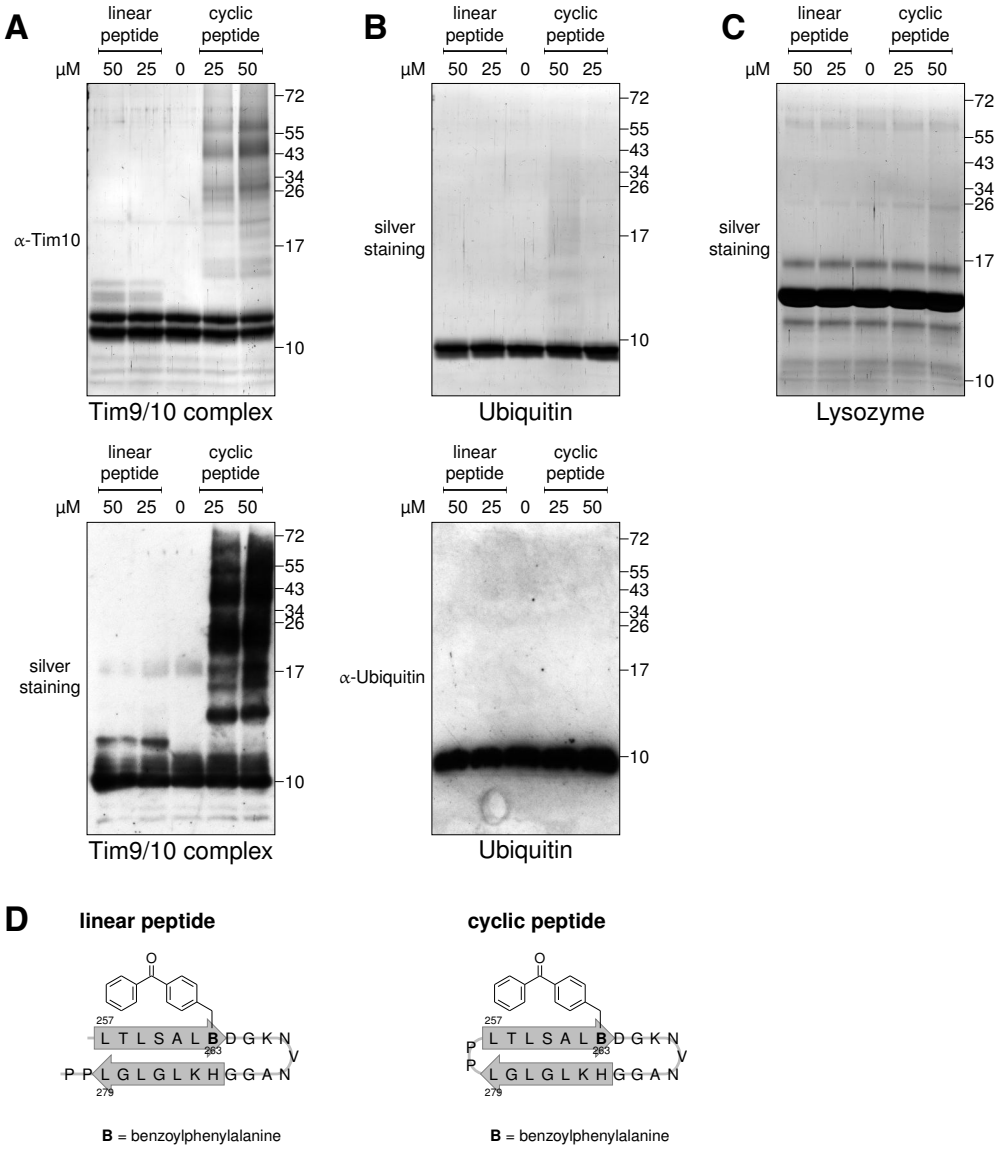


Figure A.4: **Cross-linking of TIM910 to VDAC peptides. A:** The linear and cyclic Bpa-peptides can be crosslinked to the TIM910 complex. The TIM910 complex was mixed with either the linear or the cyclic Bpa-peptide at the indicated concentrations, incubated on ice for 10 min and then crosslinked for 30 min under UV light. The samples were subjected to SDS PAGE followed by silver staining or immunodecoration with an antibody against Tim10. Cross-linking with TIM910 (**A**), Ubiquitin (**B**) and Lysozyme (**C**). **D:** The linear and cyclic VDAC peptide with a benzoylphenylalanine in position 263 for cross-linking purposes.

A.5 MC import in $Tim10_{F33Q}$ mitochondria

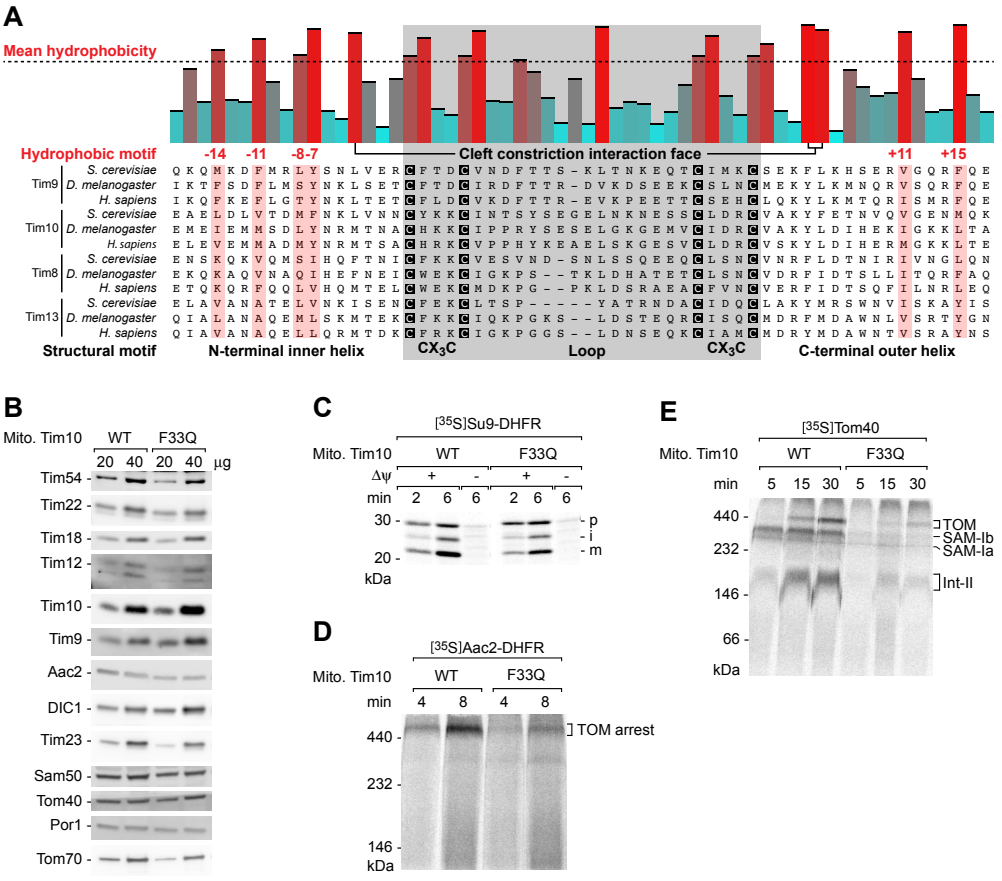


Figure A.5: **MC import in $Tim10_{F33Q}$ mitochondria. A:** Alignment of the soluble intermembrane space TIM chaperones. The bar graph indicates the mean hydrophobicity (Geneious) of the aligned sequences. A conserved hydrophobic motif in the N-terminal inner and C-terminal outer helices is highlighted in red. **B:** $Tim10_{F33Q}$ mutant mitochondria have comparable steady state protein levels to wildtype (WT) Tim10 in a tim8Δtim13Δdouble deletion background. Isolated mitochondria were analyzed by SDS PAGE and immunodecoration. **C:** Import of a presequence precursor protein into the mitochondrial matrix in $Tim10_{F33Q}$ mutant is comparable to WT. Import of a radiolabeled matrix targeted model precursor protein, *Neurospora crassa* ATP synthase subunit 9 presequence fused to the mouse dihydrofolate reductase (Su9-DHFR), into WT and $Tim10_{F33Q}$ mitochondria in the presence and absence of membrane potential ($\Delta\psi$). Samples were subjected to SDS PAGE and analyzed by autoradiography. p, precursor; i, intermediate; m, mature. **D:** The conserved hydrophobic residue at position -7 of the N-terminal inner helix of Tim10 is required for the binding of metabolite carrier precursor proteins to mitochondria. Impaired accumulation of the metabolite carrier fusion protein Aac2-DHFR (ADP/ATP carrier) at the translocase of the outer membrane (TOM) in $Tim10_{F33Q}$ mitochondria. The import of radiolabeled methotrexate-bound [^{35}S]Aac2-DHFR precursor in WT and $Tim10_{F33Q}$ mitochondria was analyzed by blue native-PAGE and autoradiography. **E:** The conserved hydrophobic residue at position -7 of the N-terminal inner helix of Tim10 is required for the import and assembly of barrel proteins into the outer membrane. The import and assembly of [^{35}S]Tom40 was analyzed by blue native-PAGE and autoradiography.

A.6 Tim12 aligment

Figure A.6: **Multiple sequence alignment of Tim12.** Multiple sequence alignment of Tim12 performed with Clustal Omega. All Tim12 sequences found in the Uniprot database have been used. The two additional cysteines of *Saccharomyces cerevisiae* Tim12, that are not part of the twin CX_3C motif, are not conserved in any other species. Position highlighted in blue.

Appendix B

ClpC1 manuscript

The natural product antibiotic Cyclomarin blocks Arginine-Phosphate induced millisecond dynamics in the N-terminal domain of ClpC1 from *Mycobacterium tuberculosis*

Katharina Weinhäupl[1], Martha Brennich[2], Uli Kazmaier[3], Joel Lelievre[4], Lluis Ballell[4], Alfred Goldberg[5] Paul Schanda[1*] and Hugo Fraga [1,5,6*]

1-Institut de Biologie Structurale, Univ. Grenoble Alpes, CEA, CNRS, IBS, 71 Avenue des martyrs, F-38044 Grenoble, France
2-European Molecular Biology Laboratory, Grenoble, France
3- Institute of Organic Chemistry, Saarland University, Campus C4.2, 66123 Saarbrücken, Germany
4- Diseases of the Developing World, Glaxo Smith and Kline, Severo Ochoa 2, 28760 Tres Cantos, Madrid, Spain
5- Department Cell Biology, Harvard Medical School, USA
6- Departamento de Bioquimica, Faculdade de Medicina da Universidade do Porto, Portugal

Running Title: Cyclomarin blocks Arginine-Phosphate induced dynamics

Keywords: nuclear magnetic resonance, Mycobacterium tuberculosis, natural product, antibiotic action, antibiotic resistance, small angle X-ray scattering, protease, chaperone

* to whom correspondence should be addressed: hfraga@med.up.pt (H.F); paul.schanda@ibs.fr (P.S.)

Abstract:

The ability of *Mycobacterium tuberculosis* to remain dormant in the host is the main obstacle to improving the treatment for tuberculosis. Recently, the natural product antibiotics cyclomarin, ecumicin and lassomycin were shown to efficiently kill *Mycobacterium tuberculosis* persisters and thus represent a new hope for future therapeutics. The target of these antibiotics is the N-terminal domain of the hexameric AAA+ ATPase ClpC1 that plays a critical role in protein degradation together with the associated ClpP1P2 protease complex. ClpC1 is responsible for the recognition, unfolding and translocation of protein substrates, such as proteins containing phosphorylated arginines into the ClpP1P2 protease for degradation. Surprisingly, these antibiotics do not inhibit the ClpC1 ATPase, and how they cause cell death is still unclear. Here we demonstrate using NMR, that arginine-phosphate binding to the ClpC1 N-terminal domain induces millisecond dynamics. Cyclomarin binding specifically blocks these N-terminal dynamics, which are caused by conformational changes and not from unfolding or oligomerization of this domain. Based on this data we propose a novel mechanism of action involving cyclomarin-induced restriction of dynamics, which prevents degradation of certain protein substrates.

Introduction:

Tuberculosis (TB) is a major public health problem with ten million people infected and two million dying each year (1). The main challenge in TB treatment is the long duration of therapy required for a cure, as the resistance of TB results from its ability to stay dormant for long periods in the host. Most antibiotics require bacterial replication for their action and this dormant state renders *Mycobacterium tuberculosis* (*Mtb*) resistant to bactericidal antibiotics. Aggravating this problem, *Mtb* has become increasingly resistant to existing antibiotics, and multidrug-resistant TB (MDR-TB) is now widespread (1).

The proteolytic complex formed by the proteins *Mtb*ClpP1 and *Mtb*ClpP2 and their hexameric regulatory ATPases, *Mtb*ClpX and *Mtb*ClpC1 are essential in mycobacteria and have emerged as attractive targets for anti-TB drug development. The Clp ATP-dependent protease complex is formed by two heptameric rings of protease subunits (*Mtb*ClpP1 and *Mtb*ClpP2) enclosing a central degradation chamber, and a hexameric ATPase complex, *Mtb*ClpC1 or *Mtb*ClpX (2). The ClpC1/ClpX ATPases recognize, unfold and translocate specific protein substrates into the *Mtb*ClpP1P2 proteolytic chamber, where degradation occurs. *Mtb*ClpC1 is a member of the class II AAA+ family of proteins, which contains a N-terminal domain (NTD) and two distinct ATP-binding modules, D1 and D2 (Fig. 1a). The active form of ClpC is a homohexamer, and in *Mtb*ClpC1 and *Synechococcus elongates* ClpC ATP alone is essential and sufficient for efficient protein degradation in association with ClpP (3). However, *B. subtilis* ClpC (*Bs*ClpC) requires the binding of both ATP and the adaptor protein MecA for formation of the active hexamer. No homologous adaptor protein has been described in *Mtb* (4) but it remains to be tested if *Mtb*ClpC1 can associate with a MecA-like protein.

Recently, Clausen and coworkers demonstrated that *Bs*ClpC specifically recognizes proteins phosphorylated on arginine residues by the arginine kinase McsB (5). These phosphorylation sites are often found in secondary structure elements and thus are accessible only when the protein is unfolded or misfolded. This innovative work revealed a new pathway for selective degradation of misfolded proteins in bacteria, but the structural consequences of arginine-phosphate (ArgP) binding to ClpC are unclear. Indeed, although the crystal structure of *Bs*ClpC NTD shows two ArgP molecules bound to the protein, no significant structural changes were observed, which is quite surprising since arginine-phosphorylated substrates (e.g. casein) can stimulate *Bs*ClpC ATPase activity and can promote complex oligomerization (5).

The potential importance of *Mtb*ClpC1 as a novel drug target against TB has been emphasized by the recent findings by two independent groups that pyrazinamide (PZA) resistant strains contain mutations (Fig. 1b) in ClpC1 (6, 7). PZA is a critical first-line TB drug used with isoniazid, ethambutol and rifampicin for the treatment of TB and is also frequently used to treat MDR-TB (1). In addition to PZA, three natural product antibiotics that specifically target *Mtb*ClpC1 have been recently discovered: cyclomarin (8, 9), ecumicin (10) and lassomycin (11). Cyclomarin, identified (and since synthetized by independent groups) is bactericidal against TB and is able to kill non-replicating bacteria (9). Despite the absence of resistance mutations, ClpC1 NTD was identified as the drug target using affinity chromatography with cyclomarin conjugated to sepharose (9). While the crystal structure of the NTD was identical with or without cyclomarin bound, observations in *Mycobacterium smegmatis* suggested that cyclomarin can increase proteolysis by the ClpC1P1P2 machine (8, 9). How cyclomarin binding to the NTD may lead to increased proteolysis is still not known. Ecumicin is another potent natural antibiotic that efficiently kills *Mtb* persisters, and resistance mutations in ClpC fall within its NTD (fig 1b). When tested *in vitro*, ecumicin increases ClpC1 ATPase activity several fold, while simultaneously compromising degradation of ClpC1P1P2 substrates (10, 12). Lassomycin, an

actinomycetes ribosomally encoded cyclic peptide, is yet another natural antibiotic able to kill *Mtb* persisters with efficiency (11). Despite differing structurally from ecumicin, lassomycin also activates ATP hydrolysis by ClpC1 ATPase and resistant mutants map to a basic domain in the protein's NTD (11).

Due to their high molecular weights and structural complexity, these natural products are challenging for structure-activity relationship studies, but compounds with similar modes of actions may be very attractive as drug candidates. Comprehension of the mechanism of action of these compounds will provide valuable insights for development of more effective TB drugs. Unfortunately, the intrinsic flexibility and exchange dynamics between different oligomeric states of ClpC1 impeded so far structural studies. For example, crystallization of *Bs*ClpC was only possible upon the removal of flexible loop regions, rendering the protein nonfunctional, and at the same time underlining the importance of dynamics for the function of these complexes (13). For this purpose, nuclear magnetic resonance (NMR) and small angle X-ray scattering (SAXS) offer important advantages to investigate protein conformation in solution and test the effects of ligands.

Here we study the interaction of these potent new antibiotics with ClpC1 by using state of the art NMR and SAXS in order to elucidate their mode of action. A proper comprehension of the ways these drugs influence ClpC1 mechanism may also clarify how the family of AAA+ ATPases functions upon substrate binding.

Results:

Drug Binding to ClpC1 NTD

It is now widely accepted that conformational heterogeneity of proteins can be an important factor in ligand binding and drug mechanism of action (14). Indeed, in the case of cyclomarin, no significant changes of the ClpC1 NTD X-ray structure were observed upon ligand binding. Therefore, it was proposed that hidden unexplored conformations could be the basis for the compound's specific actions (8). Although linking conformational heterogeneity or dynamics to protein function is a difficult task, NMR is a powerful method to elucidate such phenomena.

Full length *Mtb*ClpC1 contains 849 residues that form a functional hexamer of 561 kDa in the presence of ATP. The large size of this complex together with its low solubility and expression prohibit NMR studies of the full-length protein, even when perdeuterated and specifically methyl labelled samples are used. To circumvent this problem, we cloned and expressed separately the *Mtb*ClpC1 domains: NTD (1-145 aa), D1 (165-493 aa), D2 (494-849 aa), NTD-D1 (1-493 aa) and D1-D2 (165-849 aa). With the exception of the NTD, NTD-D1 and D2 constructs, all the other yielded insoluble proteins when expressed. Moreover, the D2 domain when purified did not show any detectable ATPase activity and did not form oligomers in the presence of ATP, while the NTD-D1 construct was soluble when expressed in ArticExpress cells at 4 °C, but precipitated when ATP was added.

Mutations in the NTD of *Mtb*ClpC1 have been shown to confer resistance to cyclomarin, ecumicin, lassomycin and PZA, indicating a pivotal role of this domain (Fig1b). This 16 kDa domain is easily accessible for solution state NMR even without perdeuteration. Therefore, we focused our work on *Mtb*ClpC1 NTD and tested the effects of the different antibiotics on its structure in solution. *Mtb*ClpC1 NTD behaved as a homogeneous monomeric protein upon size exclusion chromatography (FigS1a) and dynamic light scattering (DLS, FigS1a), and resulted in high quality NMR spectra with excellent peak dispersion, indication of a well-folded, globular protein (Fig2a). Two sets of assignment experiments were done. While at pH 7.5, we were unable to assign loop regions due to amide exchange, at pH 6.0 we were able to assign 95% of all residues allowing the mapping of almost the entire protein structure. ClpC1 NTD consists of

eight helices that fold as two repeats of a four-helix motif sharing 58% identity (Fig2b). A 14 amino acid loop between helix 4 and 5 connects the two motifs (8). Analysis of the backbone ^{1}H, ^{13}C and ^{15}N chemical shifts with TALOS+ (15) shows that the predicted secondary structure elements in the NTD in solution are highly similar to the elements in the crystal structure (Fig2b).

We proceeded testing the effect of cyclomarin binding on *Mtb*ClpC1 NTD spectra. When cyclomarin was added to the *Mtb*ClpC1 NTD, large changes in the spectrum were observed (Fig2a). Similar changes were observed with the analogue desoxycyclomarin (Fig S1b)(16). In fact, given the amplitude of chemical shift perturbations (CSP), a new set of assignment experiments was required to identify the shifted residues. The large magnitude of the CSP (Fig2c) observed is likely a consequence of the rich content of aromatic residues in the cyclomarin molecule and the corresponding ring-current effect. As the chemical shift is very sensitive to a change in the local chemical environment, we were able to map the compound binding site to the region between helix 1 and 5 (Fig2d) in agreement with the X-ray structure of this domain in the presence of bound cyclomarin (8). Analysis of the backbone amide, CO, Cα and Cβ chemical shifts with TALOS+ (15) shows that the predicted secondary structure elements in the cyclomarin bound NTD in solution are highly similar to the elements in the crystal structure (Fig S1d) excluding major changes in the domain secondary structure.

Based on the location of the resistance mutations (Fig1b), ecumicin was proposed to target the *Mtb*ClpC1 NTD as well (10). For this reason, we tested if we could observe similar effects for ecumicin, as we did for cyclomarin by NMR. Compared to cyclomarin, ecumicin only caused modest spectral changes (Fig S2). These small perturbations did not indicate a strong binding of ecumicin to the *Mtb*ClpC1 NTD and were inconsistent with previous biochemical data where ecumicin was found to be a potent inhibitor of *Mtb*ClpC1 degradation of casein, but a stimulator of *Mtb*ClpC1 ATPase activity (result confirmed here, Fig S2d) (10). Consequently, we tested ecumicin binding to the NTD using isothermal titration calorimetry (ITC). While we were able to confirm the nanomolar K_d for cyclomarin (8) (Fig S3), we were unable to obtain a saturation curve with ecumicin under the conditions used (up to 100 μM ecumicin, Fig S3). In contrast to cyclomarin and ecumicin, the addition of PZA did not result in any changes of NTD NMR spectra, furthermore no binding of PZA was observed by ITC (Fig S3).

NTD dynamics

Analysis of ^{15}N backbone relaxation rates can provide detailed information about protein dynamics on different time scales. In particular motions in the μs-ms timescale, associated with conformational changes between different states, often occur in sites important for protein function (17, 18). Studying the dynamics of *Mtb*ClpC1 NTD is particularly interesting, since it acts as a ligand recognition site that has to convey information about the bound state to the D1 and D2 rings for ATP driven translocation and unfolding. Given the very modest structural differences of the NTD in its apo and cyclomarin bond states, we reasoned that dynamics may be important in the allosteric process.

In order to investigate whether such μs-ms motions are indeed relevant for the NTD, we performed Carr-Purcell-Meiboom-Gill (CPMG) relaxation-dispersion (RD) NMR experiments. Briefly, RD profiles monitor the effective spin relaxation rate constant ($R_{2,eff}$) as a function of a variable repetition rate, ν_{CPMG}, of refocusing pulses applied during a relaxation delay. The presence of conformational dynamics is manifest as a dependence of $R_{2,eff}$ on ν_{CPMG}, *i.e.* "non-flat" RD profiles. Such dispersions arise when the local environment around the atom under consideration is fluctuating on a μs-ms time scale, either because of motion of the considered atom(s) or neighboring atoms, e.g. through binding events. When we applied this technique to apo NTD, we observed flat RD curves for all

backbone amides. Thus, this domain alone does not exhibit significant µs-ms motions. Furthermore, cyclomarin binding does not induce ms dynamics in the NTD.

Arginine-Phosphate and arginine phosphorylated proteins bind to *Mtb*ClpC1 NTD

In *B.subtilis* phosphorylation of arginines targets certain proteins for ClpCP-mediated degradation (5). Phosphorylated proteins, are first detected and bound by the NTD of ClpC and then transferred to the D1 domain for subsequent unfolding and degradation (5). Analyzing the *Mtb*ClpC1 NTD sequence, we found that the binding site of ArgP in *B. subtilis* is strictly conserved in *Mtb* (Fig3a). Furthermore, comparing the overall structures of the two NTDs, it is clear that they are structurally identical with an RMSD of 1.7 Å (Fig S3b). This striking similarity motivated us to test if *Mtb*ClpC1 NTD could also bind ArgP. Indeed, using ITC we confirmed that NTD could bind ArgP with a K_d of 5.2 µM (Fig3b), a similar value as reported for *Bs*ClpC NTD (5). In addition, a pull down showed that McsB phosphorylated lysozyme and casein bind to ClpC1 NTD while no binding was observed with the control substrate (Fig 3c). Furthermore, as previously reported for *Bs*ClpC, a large excess of free ArgP was required to inhibit this association likely reflecting a stronger binding of NTD to the protein substrate (Fig 3c). Despite some residual binding of non-phosphorylated lysozyme, a similar result was observed when full length ClpC1 was used instead of isolated NTD (Fig S3c).

Addition of ArgP to *Mtb*ClpC1 NTD resulted in strong perturbations of its backbone amide NMR spectrum (Fig 3d). This effect was specific to ArgP since no significant changes were observed with unmodified arginine or phosphoserine (Fig S4). In contrast to the rather small changes in peak intensity observed upon cyclomarin or ecumicin addition, ArgP caused large changes in peak intensity, with 45% of the peaks falling below the level of detection into the background noise (Fig 3e). As shown below, this decrease in intensity can indicate exchange events at the affected residues. Interestingly, these residues do not map exclusively to the putative binding site of ArgP, but are localised over a large part of the core of the structure, whereby helices 2, 3, 6 and 7 are most affected (Fig3e, 3f). Using CPMG RD experiments, we could prove that the observed decrease in intensity is indeed caused by millisecond dynamics (Fig 4a, 4b). Surprisingly, these dynamics can be seen throughout the domain. In Fig. 4a the affected residues and the ΔR_{2eff} are plotted on the structure of *Mtb*ClpC1 NTD. Furthermore, we assume that the part of the protein showing the highest degree of motion are the three helices with disappearing resonances (refer to Fig3f).

Although CPMG RD experiments provide residue resolved direct evidence of the motions, they do not actually reveal what the underlying motion corresponds to. In the case of low affinity binding, ms dynamics can be also caused by on/off binding dynamics of the compound. Considering the ArgP K_d, we can calculate the population of free NTD with the concentrations of ArgP used (2 mM), which results in full domain saturation (calculated free NTD 0.3 %). Fitting of the CPMG RD data with the software Chemex provides an exchange rate between the ground and excited state, the population of the excited state and the chemical shift difference between the two states. The fitted population of the excited state (5 %, Fig S5d) is not in agreement with the calculated free state of the NTD (0.3 %). Moreover, theoretical chemical shift differences of the excited state derived from fitting of the CPMG RD data with Chemex do not correlate with experimental chemical shift differences between apo and ArgP bound NTD (Fig S5a). To conclude, we confirmed that the excited state does not correspond to the free state of the NTD, but most likely to a different conformation.

Arginine has been previously reported to reduce the melting point of some protein but paradoxically has been also suggested as a stabilizer for protein preparations (19-22). In order to discard the possibility that the dynamics observed result from ArgP promoted unfolding we recorded far-UV circular dichroism (CD) spectra of *Mtb*ClpC1 NTD with and without ArgP

present. In both cases, the spectra were typical for α-helical proteins with characteristic minima at ≈210 nm and ≈222 nm (Fig. 5a). Additionally, we used chemical shift differences from the Chemex fitting of the CPMG RD data and compared these values with the chemical shift differences of the folded NTD and calculated random coil shifts derived from ncIDP (23) (FigS5b). Clearly, the fitted chemical shift differences do not correlate with the calculated random coil values. Thus, ArgP does not appear to unfold the NTD and also does not perturb the secondary structure of the domain. Alternatively, ArgP might induce domain oligomerization, which would explain peaks disappearing by transient interactions between two ClpC1 subunits. Analysis of the protein by DLS (FigS1) and Diffusion Ordered Spectroscopy (DOSY, FigS5e), however, excluded this hypothesis. Both ArgP and cyclomarin did not induce any change in *Mtb*ClpC1 NTD oligomerization.

Intrinsic aromatic fluorescence can be used to probe conformational changes upon ligand binding. While *Mtb*ClpC1 NTD does not contain any tryptophan residues, it contains 3 tyrosine residues (Y27, Y102 and Y145), which can be used as probes for domain conformational changes (Fig S5g). When we tested the effect of ArgP on NTD fluorescence, we observed an increase in tyrosine fluorescence (Fig S5f) associated with a stabilization of the protein. NTD displayed cooperative unfolding with an apparent Tm of 69 °C, this value was increased to 79 °C in the presence of ArgP 1 mM (Fig 5b).

An alternative approach to probe conformational changes is the use of fluorescent probes. 1-Anilinonaphthalene-8-Sulfonic Acid (ANS) binds to hydrophobic regions in the protein, and ANS fluorescence increases substantially when proteins undergo changes that expose hydrophobic surfaces as normally occurs during protein unfolding. ANS can, however, be used to detect subtle conformational changes, and we tested if ArgP binding leads to changes in ANS fluorescence. While ANS binding to *Mtb*ClpC1 NTD resulted in an increase in fluorescence, indicating the presence of exposed hydrophobic surface, ArgP binding did not significantly alter fluorescence (Fig5c). Because cyclomarin has intense intrinsic fluorescence, it cannot be used in the previous fluorescence studies.

Cyclomarin restricts ArgP-induced dynamics

The fact that cyclomarin binding does not affect *Mtb*ClpC1 NTD dynamics is inconsistent with the prior proposal that cyclomarin acts by causing conformational changes in this domain (8). By contrast, ArgP leads to a significant increase in domain dynamics. We therefore tested if cyclomarin binding could either block ArgP-recognition or ArgP-induced dynamics. Adding ArgP to *Mtb*ClpC1 NTD prebound to cyclomarin, we observed significant changes in the ^{1}H-^{15}N correlated HSQC spectrum in the binding site of ArgP (Fig S6). This is consistent with our pull-down results where no effect of cyclomarin on arginine phosphorylated proteins binding to the NTD was observed (Fig3c). However, the behavior of ArgP bound and ArgP/cyclomarin bound NTD was clearly different. Several peaks that disappeared upon ArgP binding, reappeared when cyclomarin was added. Instead of 45%, only 25% of all peaks disappeared when cyclomarin was added together with ArgP. Related to the spectral changes, the most striking difference between ArgP and ArgP/cyclomarin binding are the dynamic properties of the NTD. When ArgP is bound to the NTD most observed residues exhibit μs-ms dynamics. These dynamics can however be completely abolished by addition of cyclomarin. In ArgP/cyclomarin bound NTD, not a single residue showed μs-ms dynamics exactly as occurs in apo *Mtb*ClpC1 NTD (Fig 4). Cyclomarin binds to a hydrophobic "bed" formed by two phenylalanines in the symmetric axis of the domain (8), which by stabilizing the core of the domain, presumably explains how it can block completely NTD domain dynamics.

Cyclomarin prevents ArgP inhibition of FITC-casein degradation

The primary function of ClpC1 is to recognize certain cellular proteins and to unfold and translocate them into ClpP1P2 for degradation, although it is also able, at least *in vitro*, to catalyze the refolding of some proteins (2, 24). Until the discovery of ArgP in *Bs*ClpC no recognition signal for ClpC was known. In the presence of the cofactor MecA, *Bs*ClpC can catalyze the degradation of unfolded proteins, such as casein. However, ArgP (1mM) is able to block MecA-dependent proteolysis, apparently because the binding site of ArgP overlaps with the contact site of MecA and *Bs*ClpC's NTD (5). This observation clearly suggests that ClpC has two alternative mechanisms for substrate selection. One for unfolded proteins containing no specific tag and dependent on the MecA adaptor for efficient ClpC1 oligomerization and activation, and a second MecA-independent pathway that depends specifically on the presence of protein sequences containing phosphorylated arginines (5, 25). In contrast to *Bs*ClpC, *Mtb*ClpC1 does not require any cofactor for activity. In the presence of ATP it can, in association with *Mtb*ClpP1P2, efficiently degrade unfolded proteins like casein (2).

Despite these differences, the mechanism of these homologous enzymes seems to be conserved between species, and studying the effects of cyclomarin and ArgP on protein degradation by *Mtb*ClpC1P1P2 could give important mechanistic insights.

One interesting question that arises is whether ArgP binding is able to block casein degradation as it does in *Bs*ClpC. We therefore compared degradation of FITC-casein in the presence and absence of ArgP. ArgP caused a significant inhibition (up to 55%, Fig5d) of FITC-casein degradation, but it did not completely block this process even at very high ArgP concentrations. By contrast, in the presence of cyclomarin (20 μM), no inhibition of proteolysis was observed (Fig5d).

ClpC1 forms high oligomeric species in solution

We used SAXS as a complementary method to obtain information on the effect of drug binding on ClpC1 structure. Compared to X-ray diffraction, SAXS has a modest resolution but can provide information on several global parameters: the radius of gyration; the largest intraparticle distance; the particle shape; and the degree of folding, denaturation, or disorder (26). All these parameters can be good reporters for significant structural changes promoted by drug binding. In addition, SAXS does not require the preparation of highly concentrated deuterated samples, allowing the study of ClpC1 structure in native like solution conditions.

Our first SAXS measurements in batch format in the presence of ATP revealed the presence of very high molecular weight species incompatible with a ClpC1 hexamer. As *Mtb*ClpC1 is inherently prone to aggregation, known to seriously affect SAXS data interpretation, we concluded that part of the *Mtb*ClpC1 sample could be aggregated. To overcome this problem, we turned to size-exclusion chromatography coupled to SAXS (SEC-SAXS). In this system, the sample is separated according to size and shape before SAXS measurement, this way removing protein aggregates from the protein samples. Indeed, consistent with the aggregation hypothesis, when ClpC1 was loaded on a Superose 6 10/300 GL column, the chromatograms of ClpC1 showed two distinct peaks: a small peak directly after the void volume followed by a second broad peak (Fig 6a). The scattering signal at the second peak was relatively stable with a radius of gyration in the range of 8 nm (Table S1), but reduces significantly to 7.6 nm at the end of the peak, indicating either structural flexibility or overlapping oligomeric states. Surprisingly, this radius of gyration was again clearly inconsistent with a ClpC1 hexamer (radius of gyration 5.53 nm) representing instead bigger complexes. Apparently, in the conditions used (Hepes 50 mM pH 7.5, KCl 100 mM, glycerol 10%, MgCl2 4 mM and ATP 1 mM, ClpC 1 mg/ml), the hexamer is not the dominant species, and ClpC1 appears to exist in a rather distinct molecular organization.

Recently, Carroni et al, using cryo-electron microscopy (cryo-EM) and mutagenesis, showed that *S. aureus* ClpC (*Sa*ClpC) can exist in a decameric resting state formed through ClpC middle domains establishing intermolecular

head-to-head contacts (27). This head to head contacts allow the docking of two layers of ClpC molecules arranged in a helical conformation. Despite their oligomerization, these structures are however highly dynamic - as the peripheral subunits are likely in exchange – suggesting that higher order species can exist in equilibrium.

As the middle domain of *Mtb*ClpC1 is conserved compared to *Sa*ClpC, and to better understand how our SAXS data could relate to the resting state cryo-EM structure, we averaged frames from the center of the peak and the trailing end. SAXS estimates the molecular weight at the center of the peak between 1100 and 1400 kDa, compatible with 12- or 14-mers (Table S1). For the tail, the estimated weight between 860 and 1200 kDa is in better agreement with 10- or 12-mers. Bead modeling based on both curves resulted in curling stone shaped objects, whose main body matches the reported EM data in size and shape (Fig6b). An artifact resulting from the presence of several oligomer populations in the sample is observed in the form of an appendix. Direct comparison of the SAXS curve from the tail with the atomistic model gives a surprisingly good fit (χ^2 =3.4, Fig 6c), given that the SAXS curve represents an ensemble of states. Considering the similarities between our data and the previously reported cryo-EM structure, it is likely that *Mtb*ClpC1 can form a resting state with a large part of the population representing even higher oligomers than decamers. This difference could derive from a concentration dependent oligomeric equilibrium which would explain why *Mtb*ClpC1 seems larger in our study as compared to the SEC-MALS data presented by Carroni et al (27). In fact, crosslinking data from the same study already suggested the presence of complexes bigger than the EM decameric structure (27). While we could not detect the hexamer in solution, the fact that in the same conditions ClpC1 is active, catalyzing the degradation of GFPssra and casein in association with ClpP1P2, suggests that a part of the population exists as a hexamer.

We proceeded testing if the natural product antibiotics or ArgP targeting ClpC1 could affect the distribution of ClpC1 between a resting state and an active hexameric form. The addition of ecumicin (20 μM), cyclomarin (20 μM) and ArgP (200 μM) to the SEC buffer lead only to small changes in the averaged SAXS curve (Fig S8 & Table S1). As the curve is not stable, the differences are too small to allow any statement about local structural rearrangement. It appears however that the natural antibiotics tested do not affect the ClpC1 oligomer equilibrium significantly.

Discussion:

The structural characterization of AAA+ proteins involved in molecular recognition and unfolding is usually a complex task. Whereas their intrinsic heterogeneity normally results in difficult crystallization, their oligomeric organization and large size makes their study by NMR challenging. So far, no X-ray structure is available for *Mtb*ClpC1 and, as reported here, the full-length protein and its domains have intrinsic low solubility. Adding further complexity to the study of the system, we show here that *Mtb*ClpC1 can exist in an equilibrium between different oligomeric states. The existence of a resting decameric state formed by the association of head-to-head contacts between coiled-coil middle domains of ClpC was recently described by Carroni et al (27). The middle domains were proposed to repress the activity of ClpC by forming a highly dynamic resting state that can block substrate binding or ClpP interaction. Quite striking is the observation that a single point mutation in the middle domain can disrupt the resting state and result in the formation of an active hexamer even in the absence of the MecA adaptor (27). Consistent with the conservation of the key residues between *Mtb*ClpC1 and *Sa*ClpC (27), our SAXS data suggests that a similar structure is predominant versus the active hexameric form in *Mtb*ClpC1. However, contrary to the case of *S. aureus* or *B. subtilis*, where MecA was proposed to modulate this equilibrium, in the case of *Mtb* it is not clear how the distribution between resting state and hexameric

form occurs. N-terminal domains are packed between middle domains in this resting state, and were suggested to play a role in the complex stability by fluctuating between a hidden position to an exposed one, available to adaptors or substrates. One hypothesis is that the equilibrium could be modulated by substrates or natural product antibiotics. However, we could show that the addition of cyclomarin, ecumicin or ArgP ligand appears not to shift ClpC1 distribution. With the current data, we cannot exclude that binding of a bulkier substrate can shift the equilibrium towards a state where the NTDs are not constrained thus activating the unfoldase.

Whereas natural product antibiotics appear not to influence ClpC1 oligomerization equilibrium, we were unable to obtain convincing evidence for ecumicin or PZA binding to the isolated *Mtb*ClpC1 NTD. This result is intriguing, as mutations in the NTD domain have been associated with resistance to ecumicin and PZA (10). With regard to ecumicin, we and others (10) were able to demonstrate biochemical activity, namely activating ATPase activity. Most likely, ecumicin requires full length ClpC1 or an oligomeric structure for binding and ATPase activation. Possibly the binding site is located at the interface of the NTD and D1 domain, which can explain the ATPase activation by ecumicin. In this case, disrupting part of the binding site, the N-terminal interface in the resistant mutants, might be sufficient for reduced binding affinity *in vivo.* Small CSPs were evident on the NTD NMR spectrum for ecumicin in the regions where resistant mutations are located, suggesting that this region could be part of the putative binding site (FigS2). In the case of PZA, an efficient "dirty drug" with multiple reported cellular targets, we cannot exclude that resistance derives from the modulation of protein homeostasis for any of the other targets, for example preventing or increasing substrate degradation (28).

Using NMR, we could show that, although cyclomarin binds with high affinity to the ClpC1 NTD, the domain dynamics are not modified. *Mtb*ClpC1 NTD is a rather rigid domain, and showed no millisecond dynamics in the apo state. Also, cyclomarin binding does not result in peak broadening or loss of intensity that might indicate the presence of an alternative state. These observations rule out the existence of hidden conformations not captured by previous X-ray studies (8).

Finally, our finding that ArgP binding induces millisecond dynamics in the *Mtb*ClpC1 NTD domain is a new and important clue about ClpC1 mechanisms. Particularly when no alternative conformations were reported in the X-ray structure of *Bs*ClpC with ArgP-bound, and because no structural changes were observed with arginine or phosphoserine. While we excluded unfolding and transient-binding as potential explanations for the observed dynamics, we were unable to pursue structural determination since approximately half of the residues in the ArgP-bound NTD are NMR invisible. ArgP binding results in a significant increase in tyrosine fluorescence and a dramatic change in the stability of this domain. While the increased fluorescence could result from subtle changes in tyrosine side chains in a region densely packed with aromatic residues (three Phe and two Tyr, FigS5) the increased stability could derive from the ArgP binding preferentially to the folded state. The exact relationship between ArgP induced dynamics and the functional cycle of *Mtb*ClpC1 is currently not clear. Do dynamics promote target binding through ArgP recognition by allowing multiple transient interaction sites with the incoming substrate? Conformational heterogeneity and dynamics in substrate binding sites have been proposed to increase substrate recognition efficiency, and at the same time to facilitate substrate handover to downstream elements, by making a multitude of transient weak interactions with the substrate which can be easily broken (29). In fact, a single phosphorylated arginine in a protein has been shown to be sufficient for efficient ClpC-mediated degradation. Thus ClpC's molecular recognition mechanism must be highly efficient, for example 1 mM of free ArgP does not completely block protein binding, which may appear inconsistent

with the micromolar K_d that we and others report for ArgP binding to the isolated NTD (5). Despite this, the fact that cyclomarin cannot block arginine phosphorylated protein binding appears to contradict the hypothesis that dynamics are fundamental for substrate association to the NTD, not excluding however, that they are relevant for subsequent steps – for example substrate release to the D1 pore (Fig 7). Another possibility is that the observed conformational dynamics could modulate the positioning of the NTD related to the D1 pore. Indeed, studies with the type II eukaryotic homologue p97/Cdc48 ATPase complex which shares the NTD-D1-D2 architecture with ClpC1, have stressed the mechanistic relevance of the interface between NTD/D1 and the denominated up/down equilibrium of the highly mobile NTDs (30). In p97/Cdc48, NTDs have been shown to adopt, depending on the nucleotide bound, either a coplanar (down) or elevated (up) position with respect to the D1 domain (30). Increasing the life time of the NTDs "up state", thereby holding the substrate next to the D1 pore would promote substrate recognition, while supporting the down state would prevent substrate recognition, nevertheless exposing the D1 domain pore. In *E.coli* ClpA, a close bacterial homologue of ClpC, removal of NTD is known to seriously impair recognition of substrates bearing the SsrA-targeting sequence, but to have only a modest effect on degradation of unfolded proteins (31). In other words, NTDs may work as recognition domains for certain substrates, while at the same time blocking the D1 domain and preventing free access for unfolded proteins. This differential effect on certain substrates was the basis for the suggested role for the NTDs as an "entropic brush" that prevents nonspecific degradation of proteins by blocking acess to the D1 ring (32). In our view the competition between recognition of ArgP-labelled and unfolded protein, can simultaneously explain the inhibition of casein degradation by ArgP, and the effect of cyclomarin, which is able to completely abolish this inhibition of casein hydrolysis.

Put together, our work shows that ArgP binding to *Mtb*ClpC1NTD leads to widespread domain millisecond dynamics and that cyclomarin is able to block this process. While, it is surely not the absence of ClpC1 NTD dynamics that kills TB but likely its derived functional consequences, for example blocking the ArgP pathway, this work sheds light on ClpC mechanism. Rather than a static interaction, ArgP labelled proteins binding to ClpC must be understood as a highly dynamic process. Cyclomarin is therefore a unique example of a drug whose mode of action relies on the restriction of protein dynamics induced by substrate binding.

Experimental Procedures:

ArgP was obtained from Sigma. Cyclomarin and desoxycyclomarin were synthetized as previously described (33). *Mtb*ClpC1, *Mtb*ClpP1, *Mtb*ClpP2 were expressed and purified as previously described (2). *Bacillus stearothermophilus* mcsB was cloned into a pet28a+ vector and expressed and purified as previously described (5). *Mtb*ClpC1 domains: NTD corresponding to residues 1-145, D1 corresponding to residues 165-493, D2 corresponding to residues 494-849, NTD-D1 corresponding to 1-493 and D1-D2 corresponding to 165-849 were cloned into a pet28a+ vector by Genscript. Unless otherwise noted, the purification protocol consisted of an initial NiNTA affinity chromatography step taking advantage of the histidine tag, followed by a size exclusion step using a Hiload 16/600 Superdex 200 pg column. FITC-casein, GFPssra degradation and ATP hydrolysis were measured, as previously described (2). For DLS measurements, 200 μl of a 1.1 mg/ml ClpC1 NTD solutions with and without ArgP (1 mM) were used.

Far UV CD spectra were acquired on a Jasco J-810 spectropolarimeter continuously purged with nitrogen and thermostated at 20 °C. Briefly, a solution of ClpC1 NTD (5 μM) in Tris pH 7.5 NaCl 150 mM with/without ArgP (1 mM) was used to obtain CD spectra between 205 and 250 nM.

Intrinsic tyrosine fluorescence was measured in a Varian Cary Eclipse spectrofluorimeter using a 60

μM solution of ClpC1 NTD. Samples were excited at 280 nm, and fluorescence spectra were measured from 290 to 350 nm. Samples in the presence of ANS (50 μM) were excited at 370 nm, and fluorescence spectra were measured from 400 to 600 nm.

Degree of saturation of ClpC1 NTD

The degree of saturation of ClpC1 NTD with ArgP was calculated using the equation:

$$\frac{[PL]}{[P]} = \frac{1}{2}\left[1 + \frac{[L]_0}{[P]_0} + \frac{K_d}{[P]_0} - \sqrt{\left(1 + \frac{[L]_0}{[P]_0} + \frac{K_d}{[P]_0}\right)^2 - 4\frac{[L]_0}{[P]_0}}\right]$$

where $[L]_0$ is initial ligand, $[P]_0$ is initial protein and *PL* is protein ligand complex.

SAXS data collection and analysis

SEC-SAXS data were collected at ESRF BM29 (34, 35). The HPLC system (Shimadzu, Farance) was directly coupled to the 1.8 mm flow-through capillary of SAXS exposure unit. The flow rate for all online experiments was 0.3 mL/min. SAXS data collection was performed continuously throughout the chromatography run at a frame rate of 1 Hz with a Pilatus 1M detector (Dectris) at the distance of 2.876 m from the capillary. The scattering of pure water was used to calibrate the intensity to absolute units (36). The X-ray energy was 12.5 keV and the accessible q-range 0.07 nm^{-1} to 4.9 nm^{-1}. The incoming flux at the sample position was in the order of 10^{12} photons/s in 700 × 700 μm^2. A summary of the acquisition parameters is given in table 1. All images were automatically azimuthally averaged with pyFAI (37) and corrected for background scattering by the online processing pipeline (38). For each frame, the forward scattering intensity and radius of gyration were determined according to the Guinier approximation (https://github.com/kif/freesas). For each run, regions of 20 to 80 frames were averaged for further characterization. Data at small angles before the Guinier-region was removed before further data analysis to avoid experimental artefacts.

Pair distribution functions were calculated using GNOM (39). 20 ab-initio models each were calculated in C1 symmetry, using DAMMIF (40) and averaged, aligned and compared using DAMAVER (41). The scattering curve of the ClpC decamer (27) was predicted and fitted to experimental data using Crysol 3 (42).

NMR experiments

All NMR experiments were performed on Bruker Avance III spectrometers, equipped with cryogenically cooled TCI probeheads, operating at magnetic field strengths corresponding to ^{1}H Larmor frequencies of 850, 700 and 600 respectively. The sample temperature was set to 37 °C, unless stated otherwise.

Sequence-specific resonance assignments of ClpC1 NTD

Apo ClpC1 NTD was assigned in NMR buffer pH 6 (50mM MES, 100mM NaCl, 5% D_2O) and in NMR buffer pH 7.5 (50mM Tris, 50mM NaCl, 5% D_2O) at a protein concentration of 0.8mM at a ^{1}H Larmor frequency of 600. The following experiments were performed: 2D ^{15}N-^{1}H BEST HSQC, 3D BEST HNCO, 3D BEST-TROSY HNcaCO, 3D BEST HNCA, 3D BEST HNcoCA, 3D BEST HNcoCACB and 3D BEST HNCACB (43).

The same experimental conditions were used for the assignment of ClpC1 NTD in the presence of cyclomarin, except for a lower protein concentration due to low solubility of cyclomarin (0.2 mM ClpC1 NTD, 0.22 mM cyclomarin). DMSO controls were also measured.

For the assignment of ClpC1 NTD in the presence of ArgP and ArgP plus cyclomarin, ^{15}N-^{1}H BEST HSQC, BEST HNCO and BEST HNCA spectra were

recorded at a ^{1}H Larmor frequency of 850 MHz. Assignment was performed by following the chemical shifts of backbone amide, C_α and CO peaks in apo and ligand bound spectra. The sample conditions used were: 0.2 mM ClpC1 NTD and 2 mM ArgP or 0.2 mM ClpC1 NTD, 2 mM ArgP and 0.22 mM cyclomarin in NMR buffer pH 6.

Data processing and analysis were performed using the NMRPipe software package (44) and CCPN software (45).

Titration of cyclomarin into *Mtb*ClpC1 NTD:

For the cyclomarin titration four titration points were measured with 0.2 mM *Mtb*ClpC1 NTD each. The DMSO content in all samples was 2.2 %. The measured ratios between the NTD and cyclomarin were 1:0, 1:1.1, 1:1.5 and 1:2. No chemical shift changes were observed after a ratio of 1:1.1 *Mtb*ClpC1 NTD:cyclomarin. Ratio used for experiments 1:1.1.

Titration of ArgP into *Mtb*ClpC1 NTD:

Five titration points were measured with 0.3 mM *Mtb*ClpC1 NTD each. The DMSO content in all samples was 0.6 %. The measured ratios between the NTD and ArgP were 1:0, 1:0.5, 1:1, 1:2 and 1:10. No intensity changes were observed after a ratio 1:2 *Mtb*ClpC1 NTD : ArgP. Ratio used for experiments 1:10.

Titration of cyclomarin and ArgP into *Mtb*ClpC1 NTD:

Four different samples were measured for the titration of *Mtb*ClpC1 NTD with cyclomarin and ArgP. All samples contained 4.2 % DMSO and 0.2 mM *Mtb*ClpC1 NTD. We measured one reference sample, one sample with 0.22 mM cyclomarin added, one sample with 0.22 mM cyclomarin and 0.2 mM ArgP and one sample containing 0.22 mM cyclomarin and 2 mM ArgP.

Titration of Ecumicin into MtbClpC1 NTD:

Three titration points were measured with 0.2 mM *Mtb*ClpC1 NTD each. The DMSO content in all samples was 4 %. The measured ratios between the NTD and Ecumicin were 1:0, 1:1, and 1:2. Due to the insolubility of Ecumicin no higher concentrations of Ecumicin could be measured.

Titration of Pyrazinamide into MtbClpC1 NTD:

Three titration points were measured with 0.2 mM *Mtb*ClpC1 NTD each. The DMSO content in all samples was 2 %. The measured ratios between the NTD and ArgP were 1:0, 1:1 and 1:10. No intensity or chemical shift changes were observed at any measured Pyrazinamide concentration.

BEST-TROSY type (43) ^{15}N CPMG relaxation dispersion experiments were performed with the pulse scheme described by Franco et al (46) at static magnetic field strengths of 700 and 850 MHz, at a sample temperature of 37 °C. Effective relaxation rate constants, R_{2eff}, were measured at 11 (700 MHz) and 13 (850 MHz) different CPMG frequencies and derived from the commonly employed two-point measurement scheme (47), $R_{2eff}=-1/T\cdot\ln(I/I_0)$, where I is the peak intensities in the experiment with the CPMG pulse train and I_0 the one in a reference experiment without relaxation delay. T is the total relaxation delay, which was chosen as 60 ms and 40 ms in the experiments performed at 700 and 850 MHz, respectively. Peak heights and error margins were extracted in the software nmrView (OneMoon Scientific). A two-state exchange model was fitted jointly to the dispersion data of 24 residues using the program ChemEx (48). Briefly, the program involves the integration of the Bloch-McConnell equations throughout the explicit train of CPMG pulses, taking into account offset effects and finite pulse lengths. Error estimates were obtained from Monte Carlo simulations. The fit curves from the joint fit are shown in Fig S7, and a table of residue-wise chemical-shift differences is provided in table S5.

Acknowledgments:

Hugo Fraga is a COFUND fellowship recipient co-funded by the European Union and the Tres Cantos Open Lab Foundation (TC189). This work used the platforms of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). Special thanks to Dr Caroline Mas for valuable advises. Dr. Goldberg's lab has received grants from the Tres Cantos Open Lab Foundation and National Institute of General Medical Sciences. We thank the ESRF for beamtime at BM29.

Author contributions:

KW, MB and HF conceived and performed experiments, KW, MB, UK, JL, LB, AG, LB, PS and HF analyzed the data, KW and HF wrote the manuscript.

Declaration of Interests:

The authors declare no competing interests.

References:

1. Zumla, A., Nahid, P., and Cole, S. T. (2013) Advances in the development of new tuberculosis drugs and treatment regimens. *Nat Rev Drug Discov*. **12**, 388–404
2. Akopian, T., Kandror, O., Raju, R. M., Unnikrishnan, M., Rubin, E. J., and Goldberg, A. L. (2012) The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *The EMBO Journal*. **31**, 1529–1541
3. Andersson, F. I., Blakytny, R., Kirstein, J., Turgay, K., Bukau, B., Mogk, A., and Clarke, A. K. (2006) Cyanobacterial ClpC/HSP100 protein displays intrinsic chaperone activity. *J Biol Chem*. **281**, 5468–5475
4. Kirstein, J., Schlothauer, T., Dougan, D. A., Lilie, H., Tischendorf, G., Mogk, A., Bukau, B., and Turgay, K. (2006) Adaptor protein controlled oligomerization activates the AAA+ protein ClpC. *The EMBO Journal*. **25**, 1481–1491
5. Trentini, D. B., Suskiewicz, M. J., Heuck, A., Kurzbauer, R., Deszcz, L., Mechtler, K., and Clausen, T. (2016) Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature*. 10.1038/nature20122
6. Yee, M., Gopal, P., and Dick, T. (2017) Missense Mutations in the Unfoldase ClpC1 of the Caseinolytic Protease Complex Are Associated with Pyrazinamide Resistance in Mycobacterium tuberculosis. *Antimicrob. Agents Chemother.* **61**, AAC.02342–16–6
7. Zhang, S., Chen, J., Shi, W., Cui, P., Zhang, J., Cho, S., Zhang, W., and Zhang, Y. (2017) Mutation in clpC1 encoding an ATP-dependent ATPase involved in protein degradation is associated with pyrazinamide resistance in Mycobacterium tuberculosis. *Nature Publishing Group*. **6**, e8
8. Vasudevan, D., Rao, S. P. S., and Noble, C. G. (2013) Structural basis of mycobacterial inhibition by cyclomarin A. *J Biol Chem*. **288**, 30883–30891
9. Schmitt, E. K., Riwanto, M., Sambandamurthy, V., Roggo, S., Miault, C., Zwingelstein, C., Krastel, P., Noble, C., Beer, D., Rao, S. P. S., Au, M., Niyomrattanakit, P., Lim, V., Zheng, J., Jeffery, D., Pethe, K., and Camacho, L. R. (2011) The Natural Product Cyclomarin Kills Mycobacterium

Tuberculosis by Targeting the ClpC1 Subunit of the Caseinolytic Protease. *Angew. Chem. Int. Ed.* **50**, 5889–5891

10. Gao, W., Kim, J.-Y., Anderson, J. R., Akopian, T., Hong, S., Jin, Y.-Y., Kandror, O., Kim, J.-W., Lee, I.-A., Lee, S.-Y., McAlpine, J. B., Mulugeta, S., Sunoqrot, S., Wang, Y., Yang, S.-H., Yoon, T.-M., Goldberg, A. L., Pauli, G. F., Suh, J.-W., Franzblau, S. G., and Cho, S. (2015) The cyclic peptide ecumicin targeting ClpC1 is active against Mycobacterium tuberculosis in vivo. *Antimicrob. Agents Chemother.* **59**, 880–889
11. Gavrish, E., Sit, C. S., Cao, S., Kandror, O., Spoering, A., Peoples, A., Ling, L., Fetterman, A., Hughes, D., Bissell, A., Torrey, H., Akopian, T., Mueller, A., Epstein, S., Goldberg, A., Clardy, J., and Lewis, K. (2014) Lassomycin, a Ribosomally Synthesized Cyclic Peptide, Kills Mycobacterium tuberculosis by Targeting the ATP-Dependent Protease ClpC1P1P2. *Chem Biol*. **21**, 509–518
12. Jung, I.-P., Ha, N.-R., Kim, A.-R., Kim, S.-H., and Yoon, M.-Y. (2017) Mutation analysis of the interactions between Mycobacterium tuberculosis caseinolytic protease C1 (ClpC1) and ecumicin. *Int. J. Biol. Macromol.* **101**, 348–357
13. Wang, F., Mei, Z., Qi, Y., Yan, C., Hu, Q., Wang, J., and Shi, Y. (2011) Structure and mechanism of the hexameric MecA-ClpC molecular machine. *Nature*. **471**, 331–335
14. Hart, K. M., Ho, C. M. W., Dutta, S., Gross, M. L., and Bowman, G. R. (2016) Modelling proteins' hidden conformations to predict antibiotic resistance. *Nat Commun*. **7**, 12965
15. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR*. **44**, 213–223
16. Barbie, P., and Kazmaier, U. (2016) Total synthesis of desoxycyclomarin C and the cyclomarazines A and B. *Org. Biomol. Chem.* **14**, 6055–6064
17. Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim Biophys Acta*. **1814**, 942–968
18. Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E., and Kern, D. (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature*. **438**, 117–121
19. Ishibashi, M., Tsumoto, K., Tokunaga, M., Ejima, D., Kita, Y., and Arakawa, T. (2005) Is arginine a protein-denaturant? *Protein Expr. Purif.* **42**, 1–6
20. Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., and Lian, L.-Y. (2004) A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc.* **126**, 8933–8939
21. Shukla, D., and Trout, B. L. (2010) Interaction of arginine with proteins and the mechanism by which it inhibits aggregation. *J. Phys. Chem. B*. **114**, 13426–13438
22. Vagenende, V., Han, A. X., Mueller, M., and Trout, B. L. (2013) Protein-associated cation clusters in aqueous arginine solutions and their effects on protein stability and size. *ACS Chem. Biol.* **8**, 416–422
23. Tamiola, K., Acar, B., and Mulder, F. A. A. (2010) Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.* **132**, 18000–18003
24. Kar, N. P., Sikriwal, D., Rath, P., Choudhary, R. K., and Batra, J. K. (2008) Mycobacterium tuberculosis ClpC1. *FEBS J*. **275**, 6149–6158
25. Kirstein, J., Dougan, D. A., Gerth, U., Hecker, M., and Turgay, K. (2007) The tyrosine kinase McsB is a regulated adaptor protein for ClpCP. *The EMBO Journal*. **26**, 2061–2070
26. Svergun, D. I., and Koch, M. H. J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics*. **66**, 1735–1782
27. Carroni, M., Franke, K. B., Maurer, M., Jäger, J., Hantke, I., Gloge, F., Linder, D., Gremer, S., Turgay, K., Bukau, B., and Mogk, A. (2017) Regulatory coiled-coil domains promote head-to-head assemblies of AAA+ chaperones essential for tunable activity control. *elife*. **6**, 741–24
28. Zhang, Y., Shi, W., Zhang, W., and Mitchison, D. (2014) Mechanisms of Pyrazinamide Action and

Resistance. *Microbiol Spectr*. **2**, MGM2–0023–2013
29. He, L., Sharpe, T., Mazur, A., and Hiller, S. (2016) A molecular mechanism of chaperone-client recognition. *Sci Adv*. **2**, e1601625–e1601625
30. Schuetz, A. K., and Kay, L. E. (2016) A Dynamic molecular basis for malfunction in disease mutants of p97/VCP. *elife*. **5**, 653
31. Lo, J. H., Baker, T. A., and Sauer, R. T. (2001) Characterization of the N-terminal repeat domain of Escherichia coli ClpA-A class I Clp/HSP100 ATPase. *Protein Sci*. **10**, 551–559
32. Ishikawa, T., Maurizi, M. R., and Steven, A. C. (2004) The N-terminal substrate-binding domain of ClpA unfoldase is highly mobile and extends axially from the distal surface of ClpAP protease. *Journal of Structural Biology*. **146**, 180–188
33. Barbie, P., and Kazmaier, U. (2016) Total Synthesis of Cyclomarin A, a Marine Cycloheptapeptide with Anti-Tuberculosis and Anti-Malaria Activity. *Org. Lett.* **18**, 204–207
34. Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., Gobbo, A., Gordon, E., Huet, J., Kieffer, J., Lentini, M., Mattenet, M., Morawe, C., Mueller-Dieckmann, C., Ohlsson, S., Schmid, W., Surr, J., Theveneau, P., Zerrad, L., and McSweeney, S. (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Radiat*. **20**, 660–664
35. Brennich, M. E., Round, A. R., and Hutin, S. (2017) Online Size-exclusion and Ion-exchange Chromatography on a SAXS Beamline. *J Vis Exp*. 10.3791/54861
36. Orthaber, D., Bergmann, A., and Glatter, O. (2000) SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J Appl Crystallogr*. **33**, 218–225
37. Ashiotis, G., Deschildre, A., Nawaz, Z., Wright, J. P., Karkoulis, D., Picca, F. E., and Kieffer, J. (2015) The fast azimuthal integration Python library: pyFAI. *J Appl Crystallogr*. **48**, 510–519
38. Brennich, M. E., Kieffer, J., Bonamis, G., De Maria Antolinos, A., Hutin, S., Pernot, P., and Round, A. (2016) Online data analysis at the ESRF bioSAXS beamline, BM29. *J Appl Crystallogr*. **49**, 203–212
39. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr*. **25**, 495–503
40. Franke, D., and Svergun, D. I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr*. **42**, 342–346
41. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initioshape determination in small-angle scattering. *J Appl Crystallogr*. **36**, 860–864
42. Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M., and Svergun, D. I. (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr*. **50**, 1212–1225
43. Favier, A., and Brutscher, B. (2011) Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J Biomol NMR*. **49**, 9–15
44. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. **6**, 277–293
45. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Structure, Function, and Bioinformatics*. **59**, 687–696
46. Franco, R., Gil-Caballero, S., Ayala, I., Favier, A., and Brutscher, B. (2017) Probing Conformational Exchange Dynamics in a Short-Lived Protein Folding Intermediate by Real-Time Relaxation-Dispersion NMR. *J. Am. Chem. Soc.* **139**, 1065–1068
47. Mulder, F. A., Mittermaier, A., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Studying excited states of proteins by NMR spectroscopy. *Nat Struct Biol*. **8**, 932–935
48. Bouvignies, G., and Kay, L. E. (2012) Measurement of proton chemical shifts in invisible states of

slowly exchanging protein systems by chemical exchange saturation transfer. *J. Phys. Chem. B*. **116**, 14311–14317

Figure Legends:

Figure 1. Structural model of *Mtb*ClpC1 showing drug-resistant mutations in the N-terminal domain.
(a) Side and top view of *Mtb*ClpC1 structural model based on *Bs*ClpC in complex with MecA. In gray NTD and linker region, in blue D1 domain and in green D2 domain. NTD and linker are predicted to be mobile. (b) *Mtb*ClpC1 NTD is the assumed target of the natural product antibiotics ecumicin, cyclomarin and lassomycin. Mutations in pyrazinamide-resistant strains have also been mapped to this domain. Depicted as pink spheres are ecumicin-resistant mutations (L92S, L96P, L92F) (10), in yellow, lassomycin-resistant (Q17R, Q17H, R21S, P79T) (11) and in green pyrazinamide-resistant (G99D (7) and L88V, G99D, I108T, R114L mutations (6)). The crystal structure of cyclomarin-bound *Mtb*ClpC1 NTD is shown in Figure 2d.c) Binding site of cyclomarin (yellow) and ArgP (red) on the *Mtb*ClpC1 NTD.

Figure 2. NMR assignment and cyclomarin-binding site of *Mtb*ClpC1 NTD.
(a) ^{1}H-^{15}N correlated backbone amide spectrum of apo (black) and cyclomarin bound (blue) *Mtb*ClpC1 NTD. 95% of amide resonances of the apo and 79% of the cyclomarin bound protein spectrum have been assigned. (b) TALOS+ predicted helix propensity derived from NMR assignments of apo *Mtb*ClpC1 NTD. In dark blue predicted helix, in white predicted loop and as a gray background helical parts in the X-ray structure of *Mtb*ClpC1 NTD (PDB: 3wdb). The secondary structure in solution and in the crystal, seem to be identical. (c) Combined chemical shift difference between apo and cyclomarin-bound *Mtb*ClpC1 NTD ^{1}H-^{15}N HSQC spectra. Chemical shift differences are mapped on the structure in Figure 2d. (d) Chemical shift differences from Figure 2c plotted on the structure of cyclomarin bound *Mtb*ClpC1 NTD (PDB: 3wdc). Assigned backbone amides are shown as spheres, and unassigned residues as gray cartoon. Chemical shift differences are plotted in a spectrum from blue to white, whereby blue indicates a strong effect. Cyclomarin is shown as yellow sticks

Figure 3. The effect of ArgP binding on *Mtb*ClpC1 NTD.
(a) Sequence alignment of *Mtb*ClpC1 NTD and *Bs*ClpC NTD. Identical residues are highlighted in black and similar residues in grey. The binding site of ArgP in *Bs*ClpC NTD is circled in red. (b) Representative Isothermal Calorimetry Titration of ArgP binding to *Mtb*ClpC1 NTD (N =1.99 $\pm$ 0.02; K_d 5.2 $\pm$ 0.5 μM; ΔH -4066 $\pm$ 59 cal/mol; ΔS 10.3 cal/mol/deg). (c) NTD is able to pull down lysozyme phosphorylated by McsB kinase but not non-treated lysozyme (lane 1 and 4). Cyclomarin (50 μM) is unable to block substrate binding but a reduction is observed with ArgP (1 mM). (d) ^{1}H-^{15}N correlated backbone amide spectrum of apo (black), ArgP (green), cyclomarin (blue) and ArgP/cyclomarin (red) bound *Mtb*ClpC1 NTD. 95% of apo, 55% ArgP, 85% of cyclomarin and 75% of CymA/ArgP bound *Mtb*ClpC1 NTD amide resonances are NMR visible. (e) Loss in peak intensity of resonances in ^{1}H-^{15}N HSQC spectra upon ArgP binding with (red) or without (green) cyclomarin added. (f) Peak height ratio of ArgP bound *Mtb*ClpC1 NTD (Figure 3d, left) plotted on its structure (PDB: 3wdb). Assigned residues are shown as spheres, unassigned residues as gray cartoon, residues that disappear upon ArgP binding as white cartoon. Peak height ratio is drawn as a spectrum from green to white. Whereby white indicates loss in intensity. Two arginine-phosphate molecules (red sticks) are placed at the putative ArgP binding site, identical to the X-ray structure of *Bs*ClpC NTD (PDB: 5hbn).

Figure 4. ArgP induces millisecond dynamics in *Mtb*ClpC1 NTD.

(a) Residues that exhibit millisecond dynamics plotted on the *Mtb*ClpC1 NTD structure (PDB: 3wdb). All assigned residues are shown as spheres. Residues that have a $\Delta R_{2,eff}$ of 5 are in yellow, of 15 orange and of 30 or more red. (b) Examples of CPMG curves for residues with a $\Delta R_{2,eff}$ of 5, 15 or 30. Apo *Mtb*ClpC1 NTD has no millisecond dynamics (lower row), 63% of all NMR visible residues experience ms dynamics when ArgP is bound (upper row), if cyclomarin is added before ArgP no more ms dynamics can be observed, resulting in flat dispersion curves (middle row).

Figure 5. Effect of ArgP on *Mtb*ClpC1 NTD secondary structure and *Mtb*ClpC1 substrate degradation.
(a) CD spectrum of apo (black) and ArgP bound (green) *Mtb*ClpC1 NTD. ArgP binding does not influence the secondary structure of *Mtb*ClpC1 NTD. (b) ArgP bound (green) is more stable than Apo *Mtb*ClpC1 NTD. Intrinsic tyrosine fluorescence was measured as a function of temperature. (c) ArgP (green) does not change exposure of hydrophobic regions in *Mtb*ClpC1 NTD (black). Shown in gray the fluorescence of ANS in buffer. (d) ArgP inhibits FITC-casein degradation by *Mtb*ClpC1P1P2 (black curve). Cyclomarin (20µM) is able to block this inhibition (blue curve).

Figure 6. ClpC1 forms high oligomeric species in solution.

(a) SEC-SAX chromatogram of *Mtb*ClpC1, showing the radius of gyration of the eluted species (pink) with the respective absorbance at 280 nm (black). The first peak immediately after the column void volume corresponds to protein aggregates. (b) DAMMIF models obtained from *Mtb*ClpC1 scattering curve in the presence of different ligands ecumicin (pink), cyclomarin (blue) and ArgP (green). With the exception of the appendix, an artifact derived from sample heterogeneity, the obtained model fits well to the structure obtained previously by Cryo-EM. (c) The scattering curve of the resting state *Sa*ClpC decamer was predicted and fitted using Crysol 3 to the experimental curve obtained from the apo*Mtb*ClpC1 tail of the peak .

Figure 7. *Mtb*ClpC1 exists in equilibrium between a resting state and a functional hexamer.

*Mtb*ClpC1 can form a resting state in equilibrium with the active hexameric form. Phosphorylation of arginines marks proteins for degradation by the ClpCP machinery. Phosphorylated arginines bind to the NTD where they induce millisecond dynamics that could either facilitate contact between different NTDs or the transfer of the substrate to the D1 domain pore. Although, cyclomarin binding does not change the structure of the NTD or substrate binding it restricts ArgP induced dynamics.

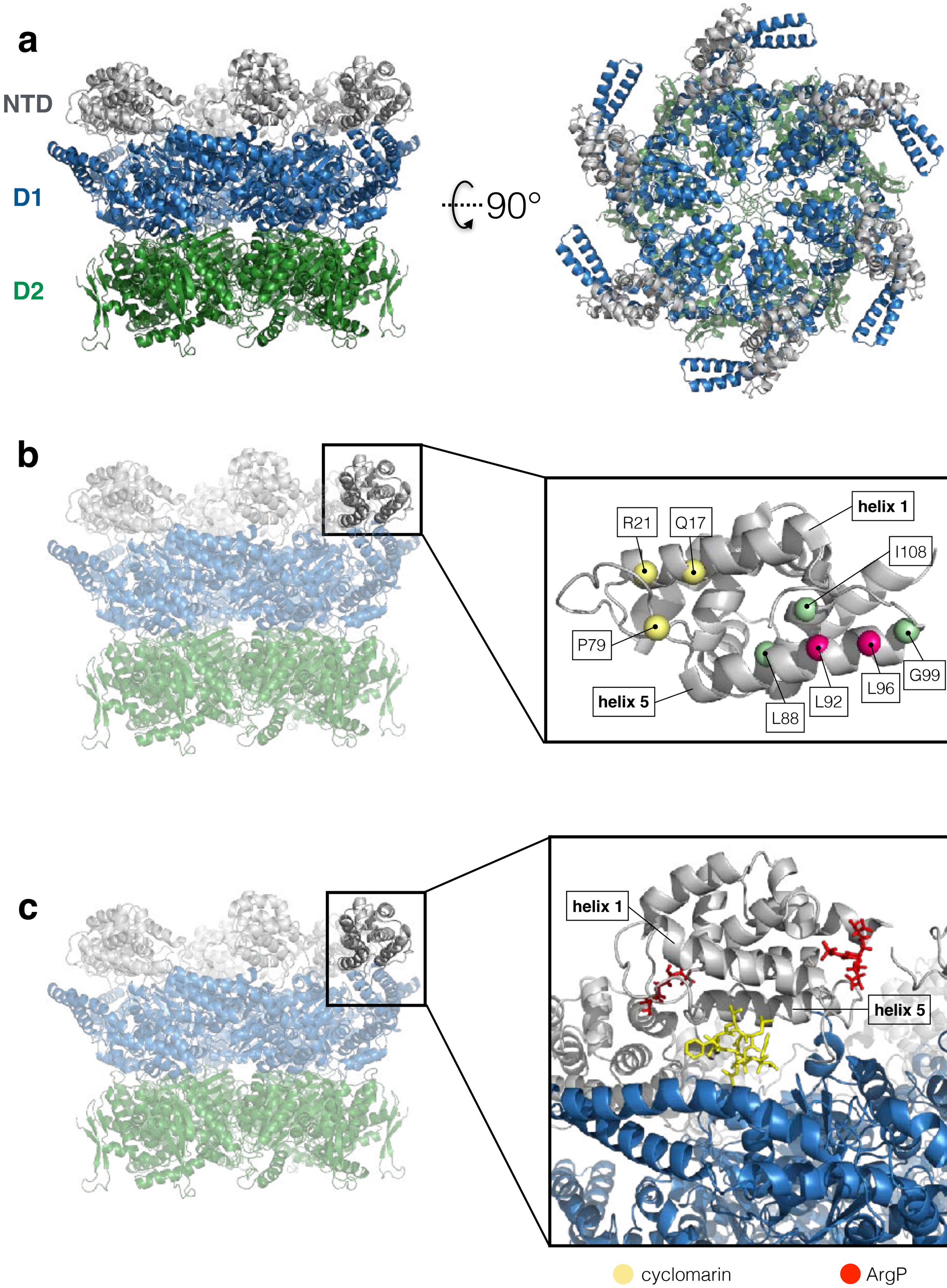

a
NTD
D1
D2
90°
b
R21
Q17
helix 1
I108
P79
helix 5
L88
L92
L96
G99
c
helix 1
helix 5
cyclomarin
ArgP

Figure 1

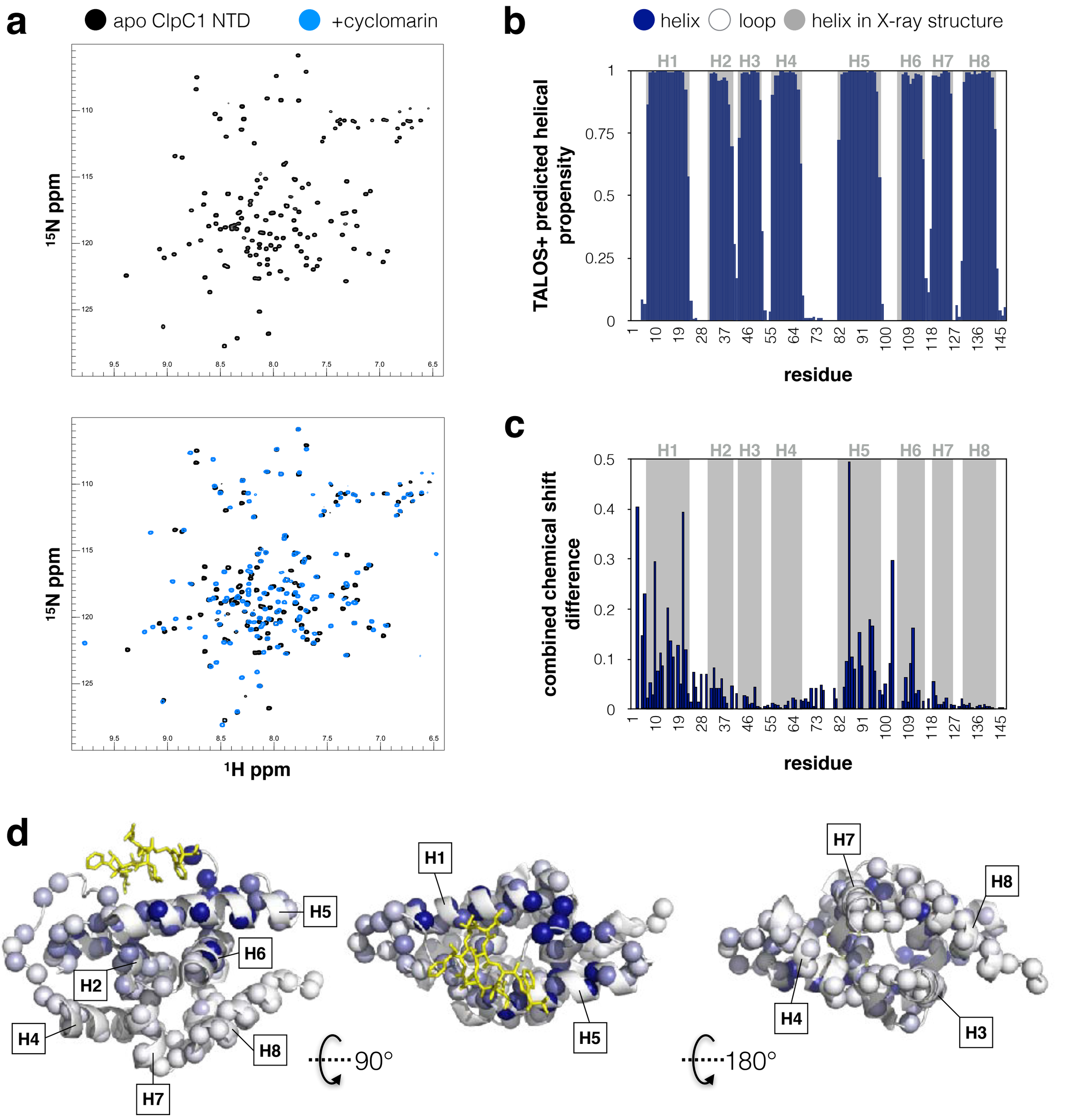
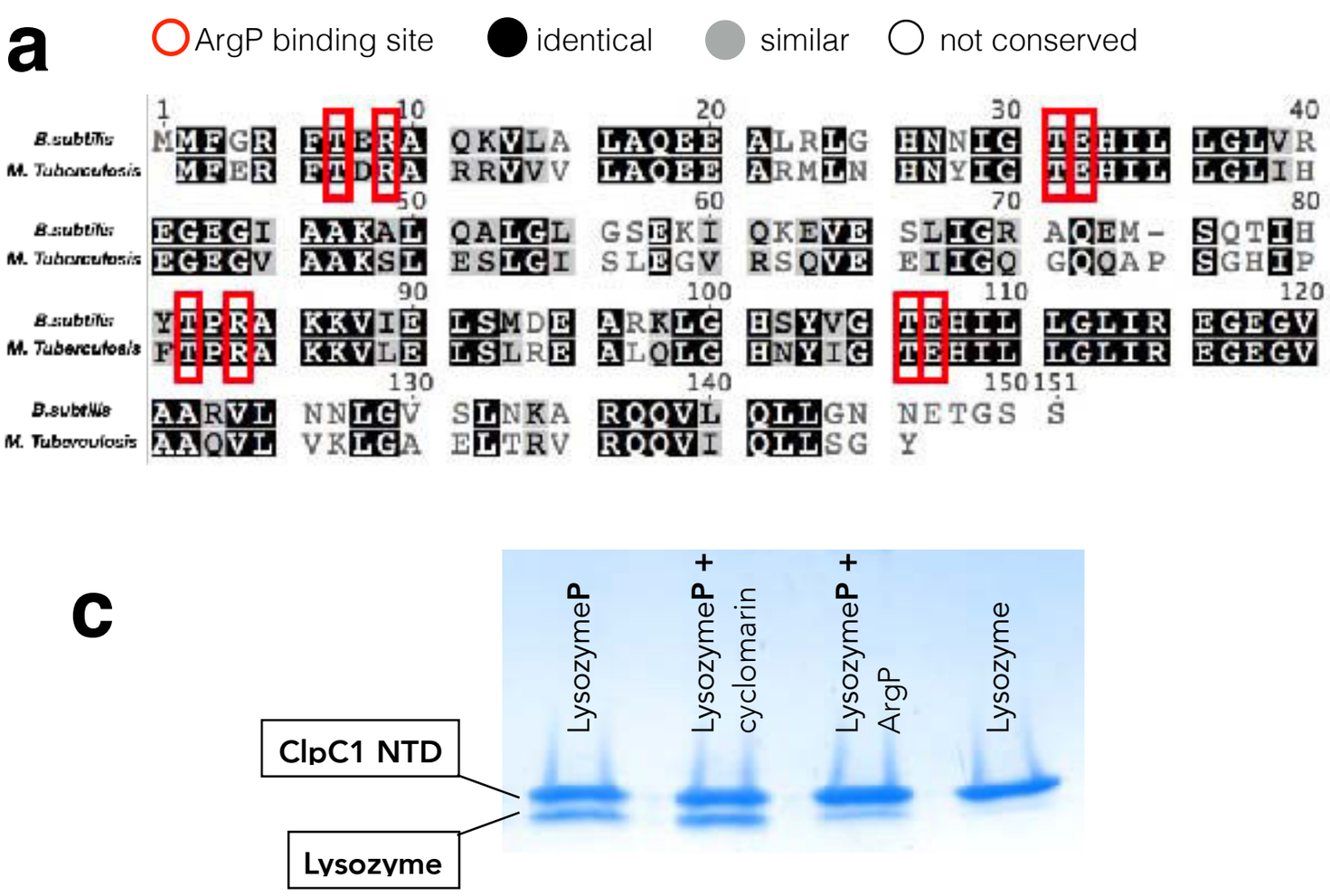
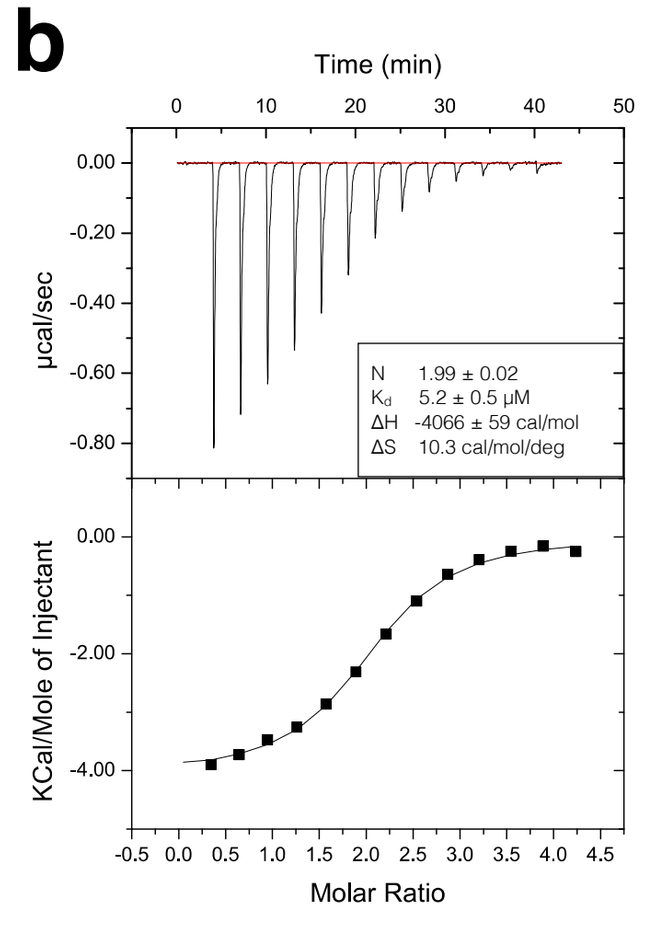
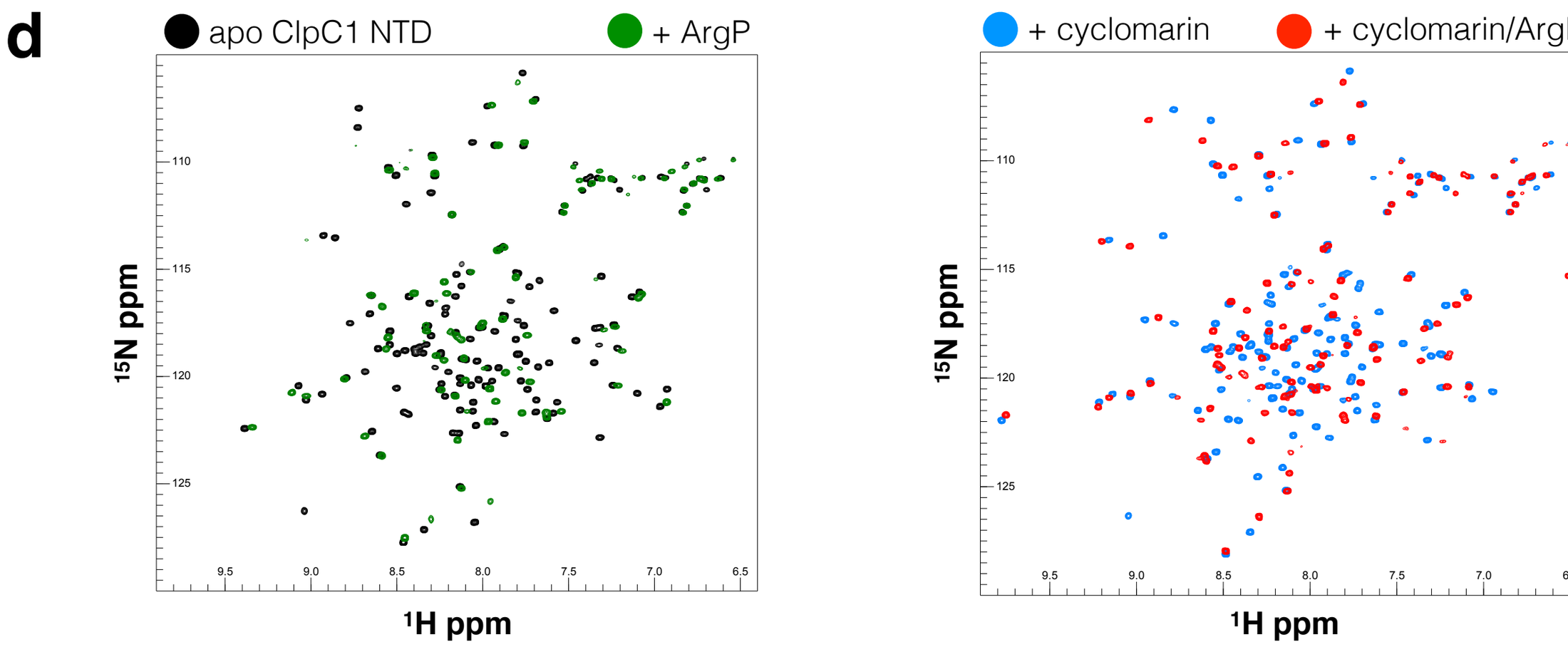

a
apo ClpC1 NTD
+cyclomarin
15N ppm
1H ppm
b
helix
loop
helix in X-ray structure
TALOS+ predicted helical propensity
residue
c
combined chemical shift difference
residue
d
H1
H2
H3
H4
H5
H6
H7
H8
90°
180°
cyclomarin
peak shift upon cyclomarin binding

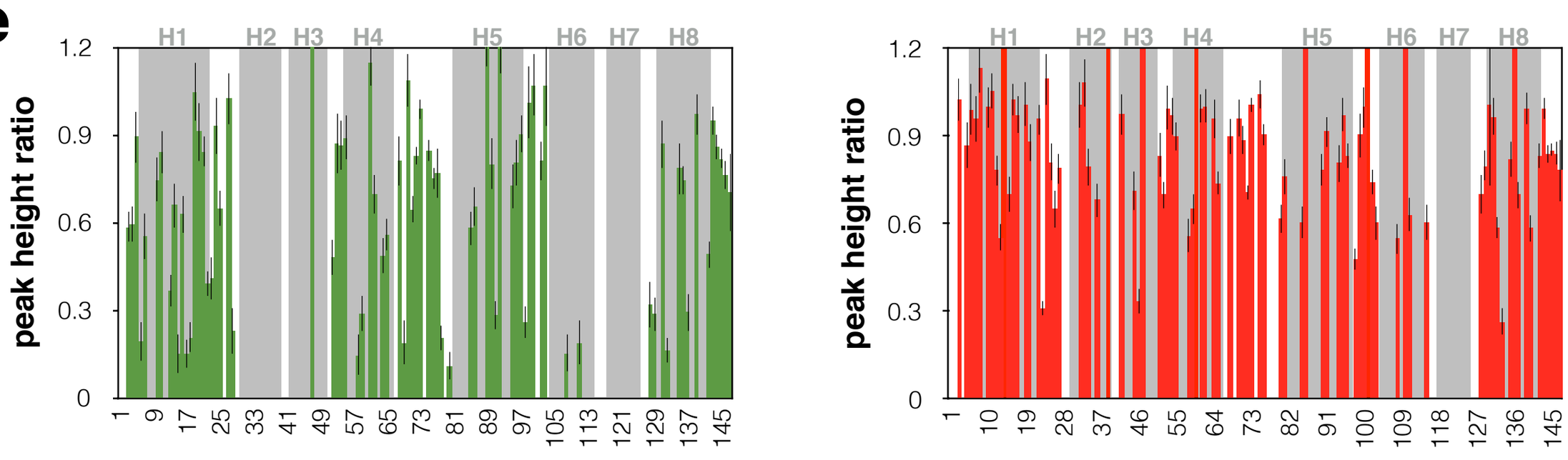
Figure 2


a
ArgP binding site
identical
similar
not conserved
B.subtilis
M. Tuberculosis
1 10 20 30 40
MMFGR FTERA QKVLA LAQEE ALRLG HNNIG TEHIL LGLVR
MFER FTDRA RRVVV LAQEE ARMLN HNYIG TEHIL LGLIH
50 60 70 80
EGEGI AAKAL QALGL GSEKI QKEVE SLIGR AQEM- SQTIH
EGEGV AAKSL ESLGI SLEGV RSQVE EIIGQ GQQAP SGHIP
90 100 110 120
YTPRA KKVIE LSMDE ARKLG HSYVG TEHIL LGLIR EGEGV
FTPRA KKVLE LSLRE ALQLG HNYIG TEHIL LGLIR EGEGV
130 140 150 151
AARVL NNLGV SLNKA RQQVL QLLGN NETGS S
AAQVL VKLGA ELTRV RQQVI QLLSG Y


b
Time (min)
0 10 20 30 40 50
µcal/sec
0.00
-0.20
-0.40
-0.60
-0.80
N 1.99 ± 0.02
Kd 5.2 ± 0.5 µM
ΔH -4066 ± 59 cal/mol
ΔS 10.3 cal/mol/deg
KCal/Mole of Injectant
0.00
-2.00
-4.00
-0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5
Molar Ratio

c
ClpC1 NTD
Lysozyme
LysozymeP
LysozymeP + cyclomarin
LysozymeP + ArgP
Lysozyme


d
apo ClpC1 NTD
+ ArgP
+ cyclomarin
+ cyclomarin/ArgP
15N ppm
110
115
120
125
9.5 9.0 8.5 8.0 7.5 7.0 6.5
1H ppm


e
H1 H2 H3 H4 H5 H6 H7 H8
peak height ratio
1.2
0.9
0.6
0.3
0
1 9 17 25 33 41 49 57 65 73 81 89 97 105 113 121 129 137 145
residue

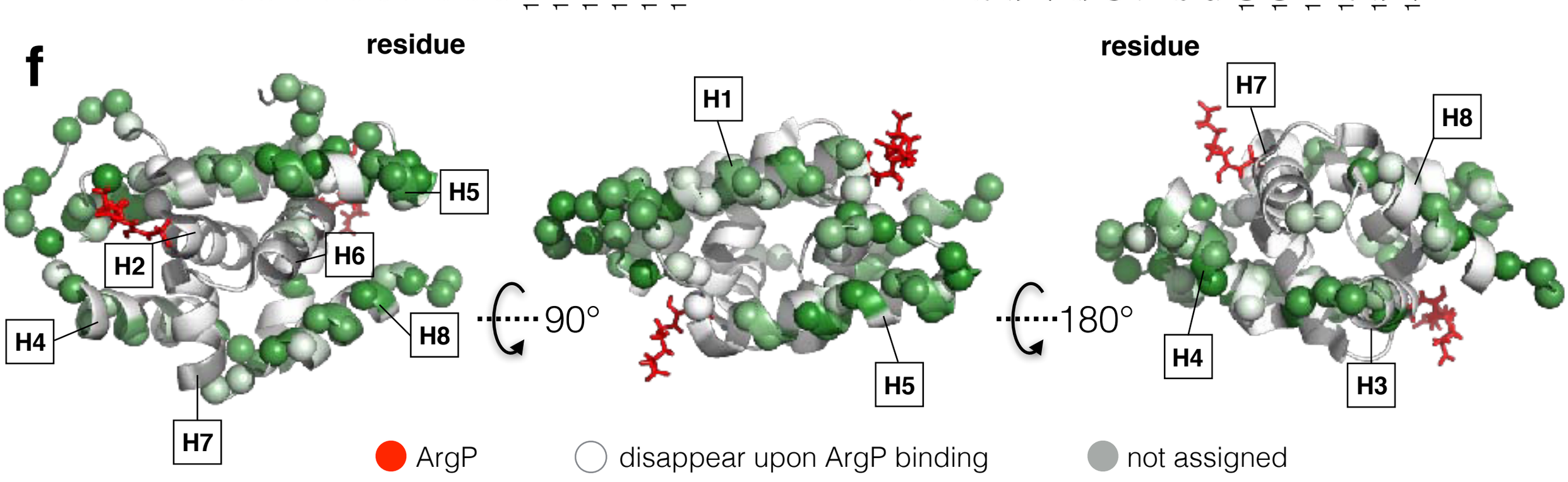

f
H1
H2
H4
H5
H6
H7
H8
90°
180°
ArgP
disappear upon ArgP binding
not assigned

Figure 3

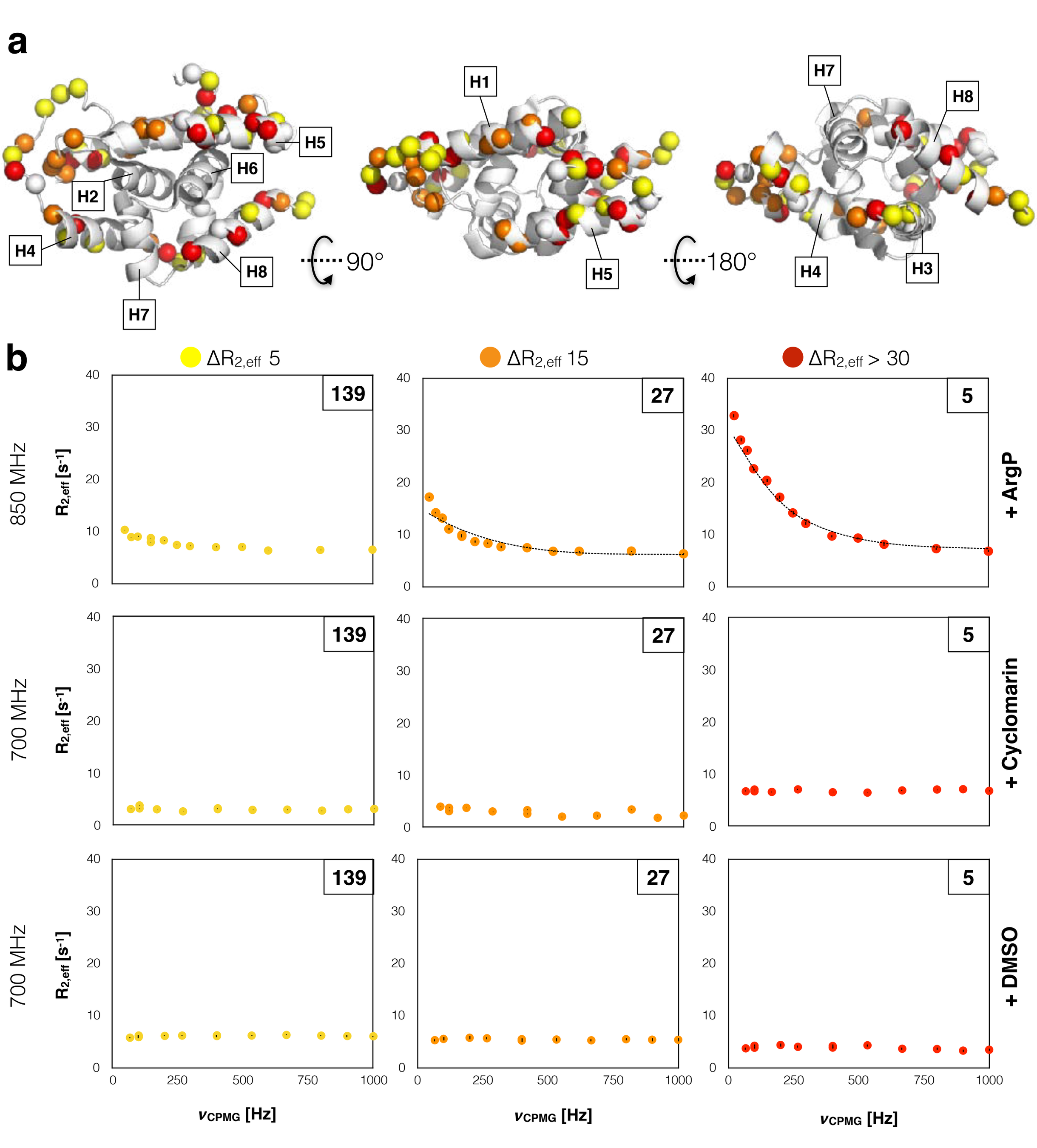
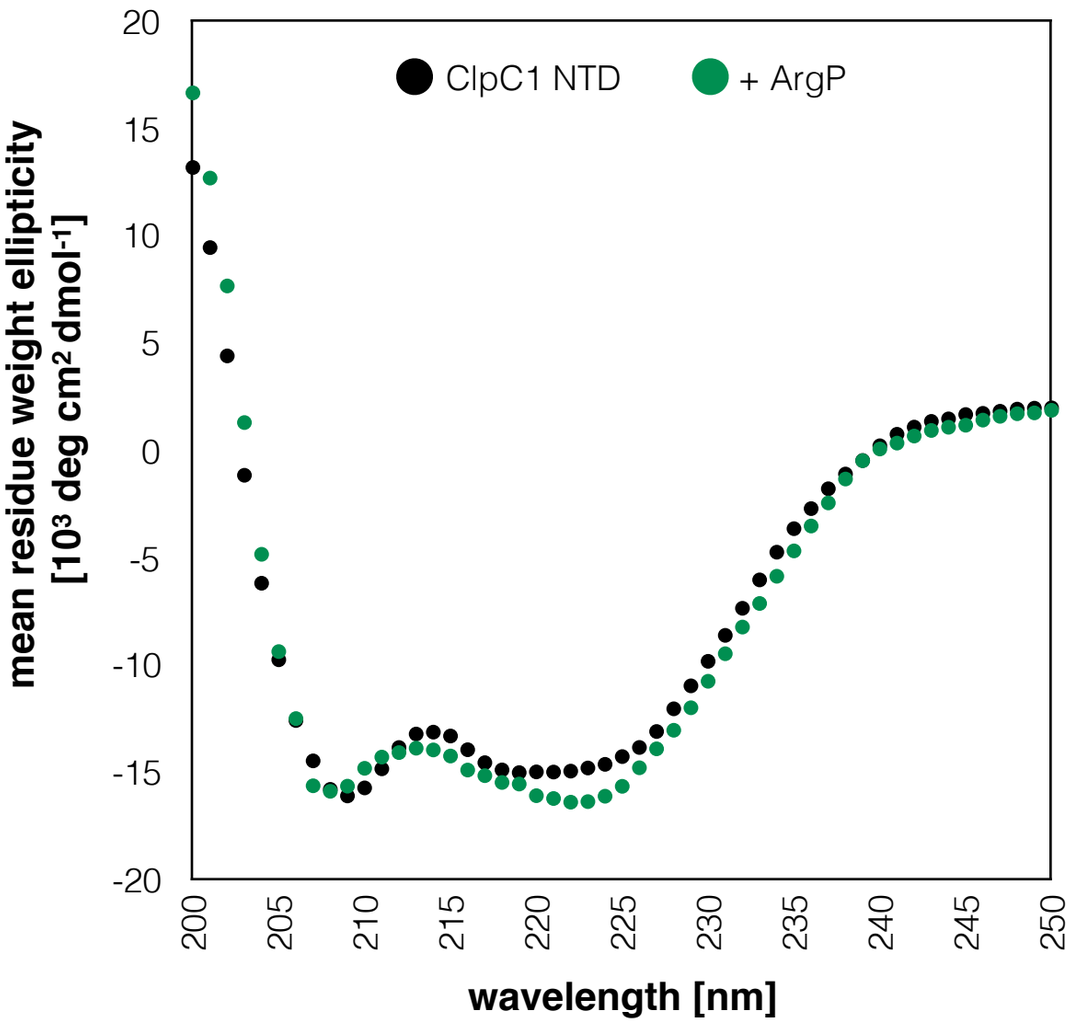
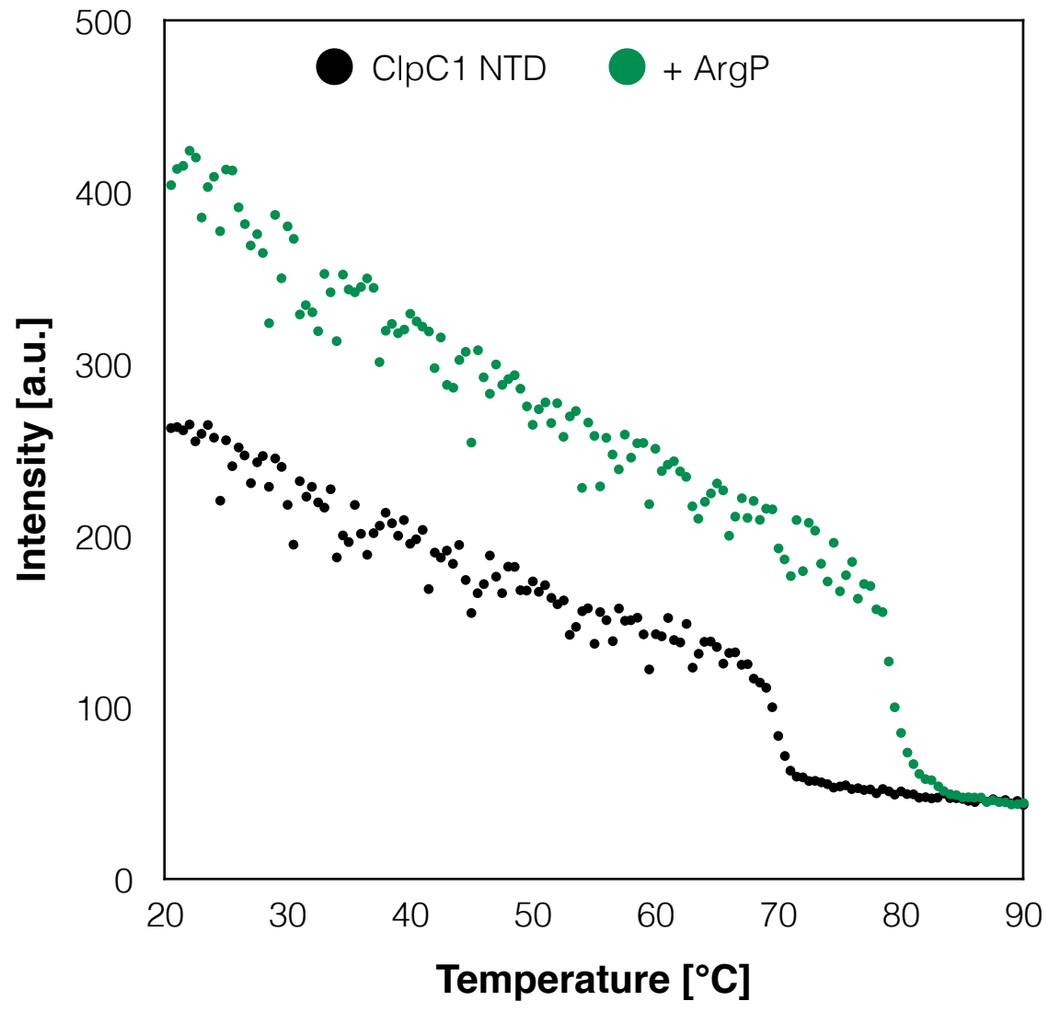
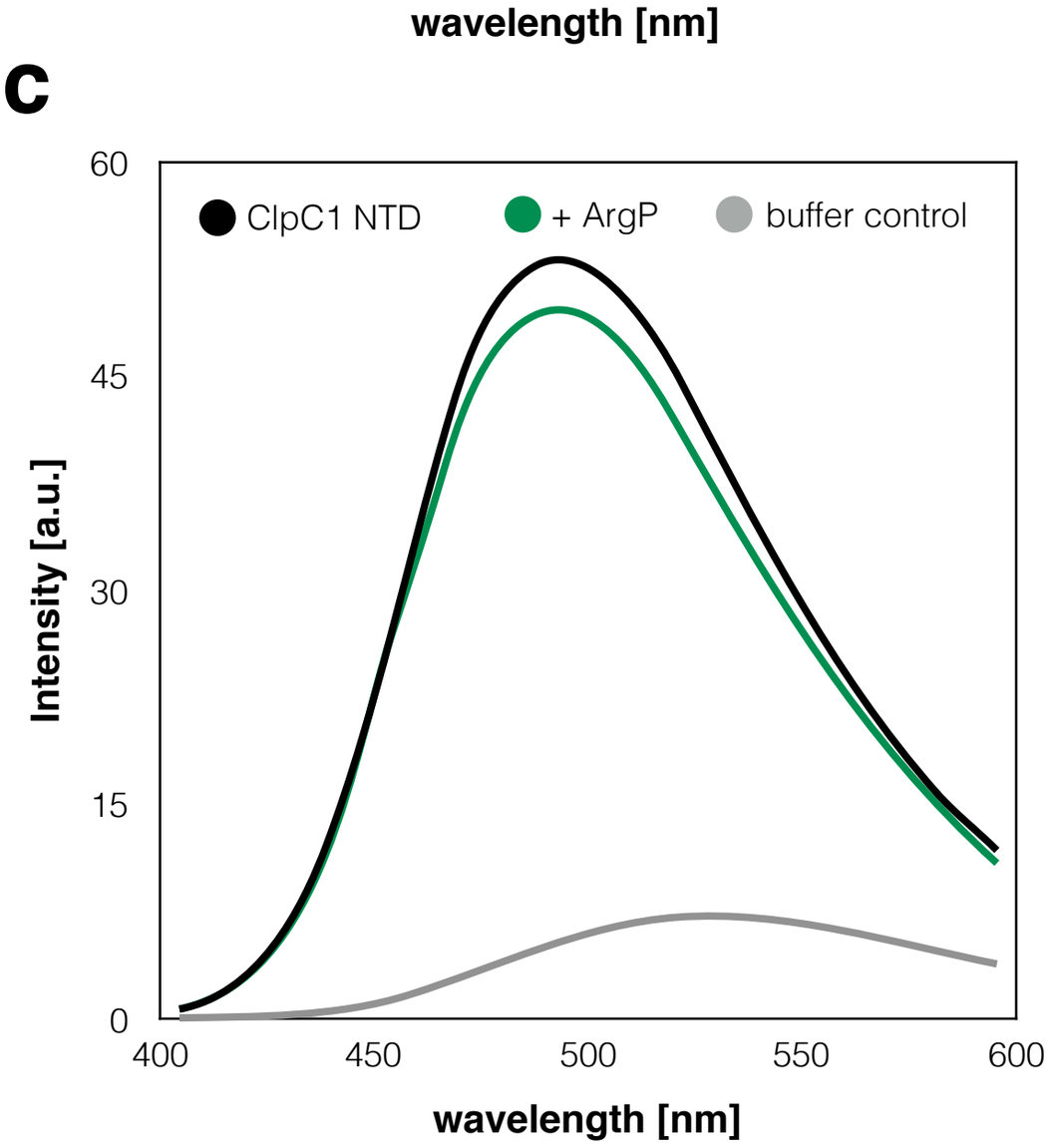

a
H1
H2
H3
H4
H5
H6
H7
H8
90°
180°
b
ΔR2,eff 5
ΔR2,eff 15
ΔR2,eff > 30
850 MHz
700 MHz
R2,eff [s-1]
139
27
5
+ ArgP
+ Cyclomarin
+ ArgP
+ DMSO
νCPMG [Hz]
0
10
20
30
40
0
250
500
750
1000

Figure 4


a
ClpC1 NTD
+ ArgP
mean residue weight ellipticity [10³ deg cm² dmol⁻¹]
20
15
10
5
0
-5
-10
-15
-20
200
205
210
215
220
225
230
235
240
245
250
wavelength [nm]


b
ClpC1 NTD
+ ArgP
Intensity [a.u.]
500
400
300
200
100
0
20
30
40
50
60
70
80
90
Temperature [°C]


c
ClpC1 NTD
+ ArgP
buffer control
Intensity [a.u.]
60
45
30
15
0
400
450
500
550
600
wavelength [nm]

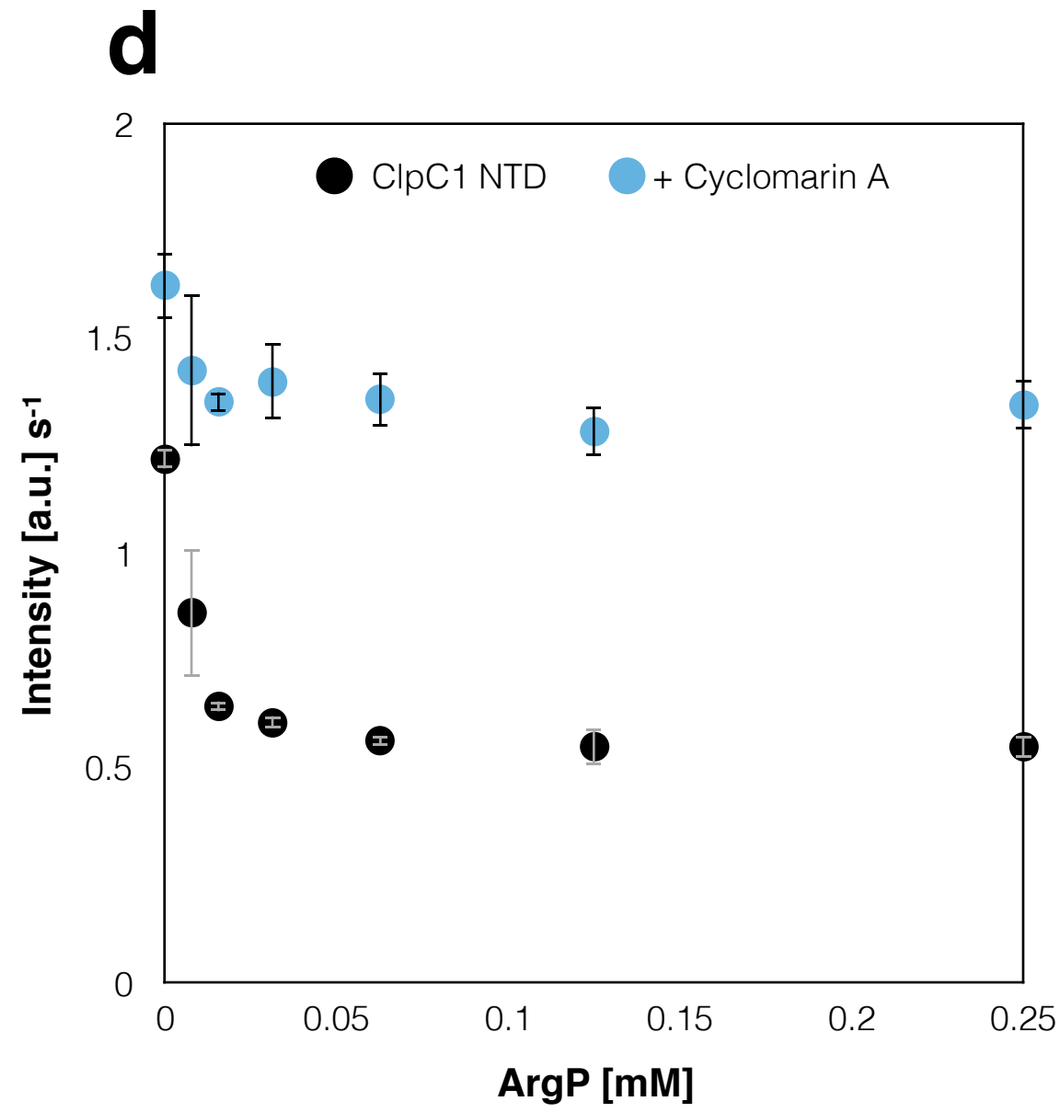
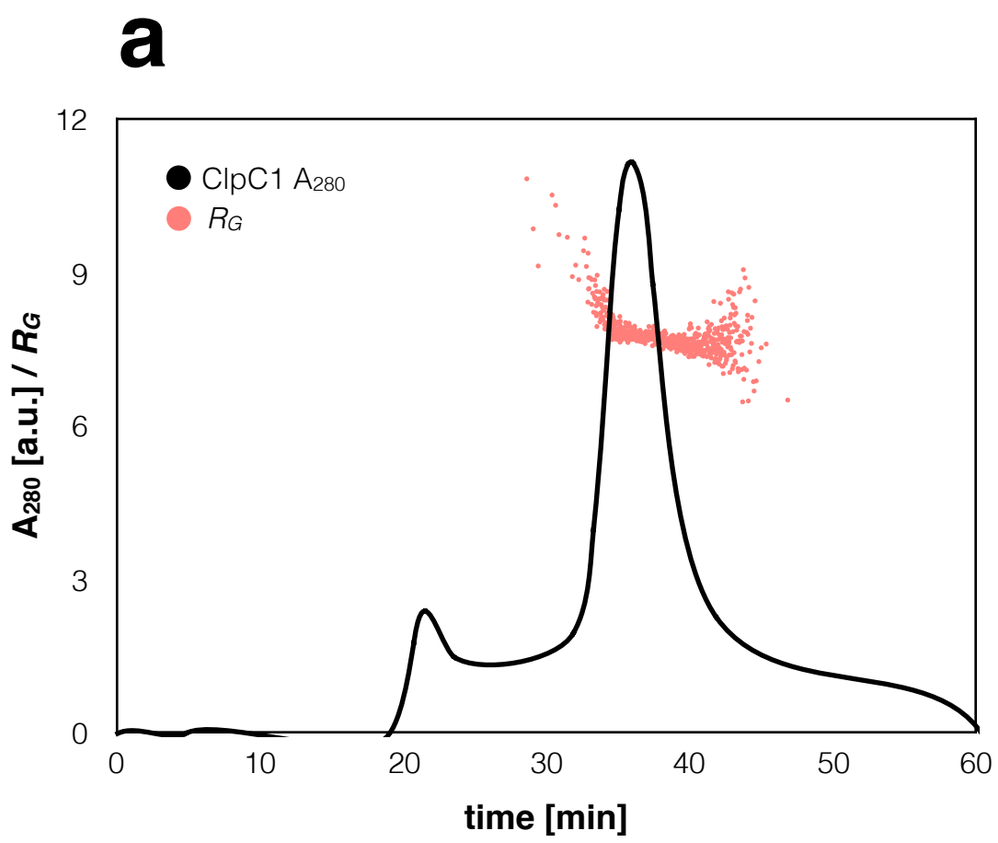
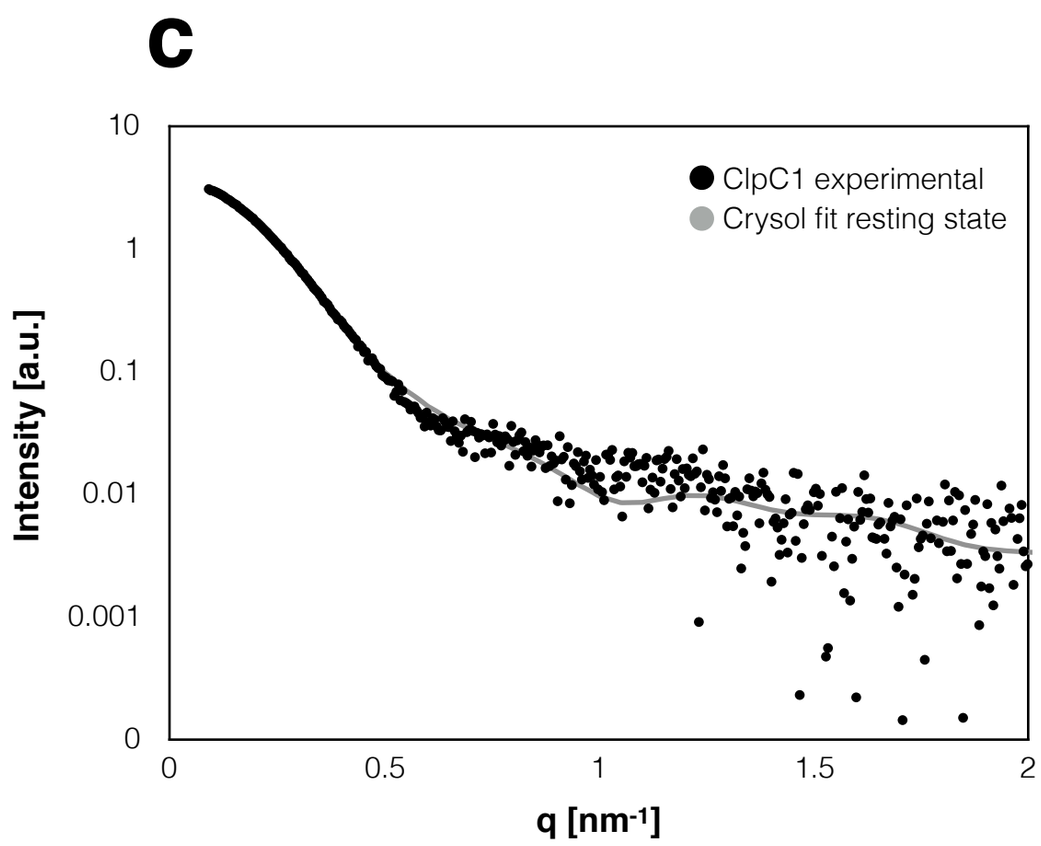

d
ClpC1 NTD
+ Cyclomarin A
Intensity [a.u.] s⁻¹
2
1.5
1
0.5
0
0
0.05
0.1
0.15
0.2
0.25
ArgP [mM]

Figure 5


a
ClpC1 A280
RG
A280 [a.u.] / RG
time [min]
0
3
6
9
12
0
10
20
30
40
50
60


c
ClpC1 experimental
Crysol fit resting state
Intensity [a.u.]
q [nm-1]
10
1
0.1
0.01
0.001
0
0
0.5
1
1.5
2

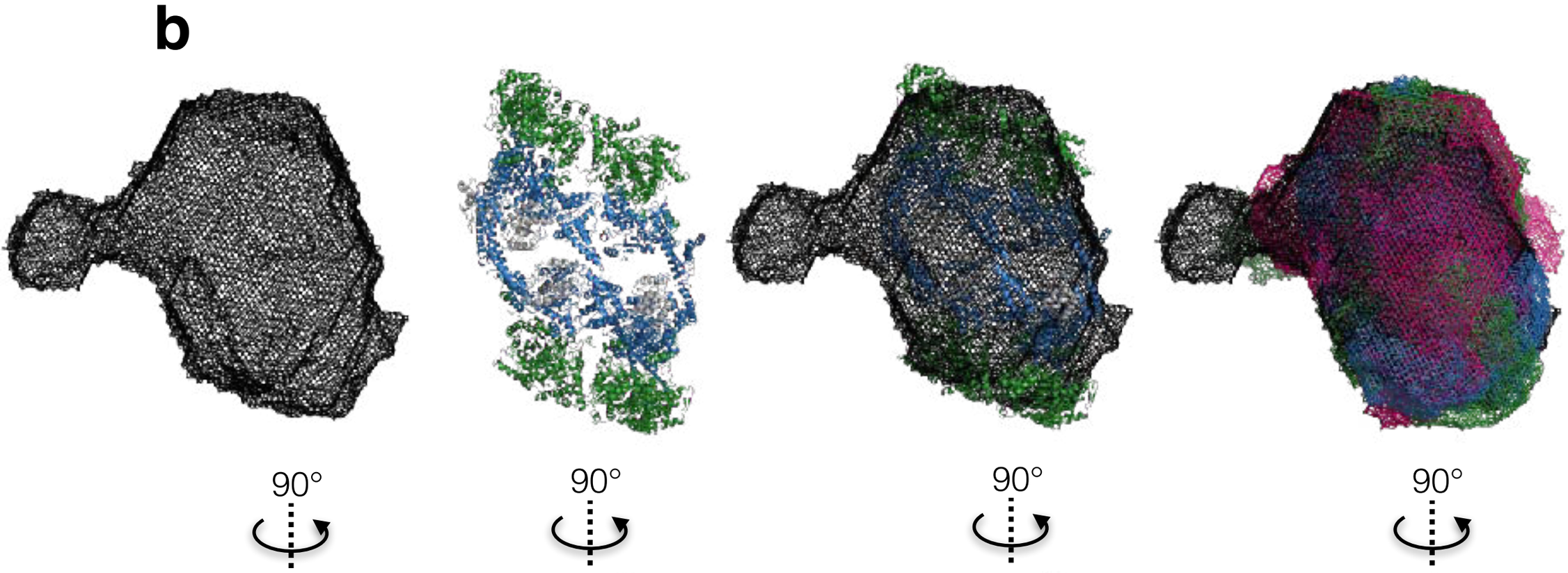

b
90°
90°
90°
90°
apo ClpC1 NTD
+ ArgP
+ cyclomarin
+ ecumicin

Figure 6

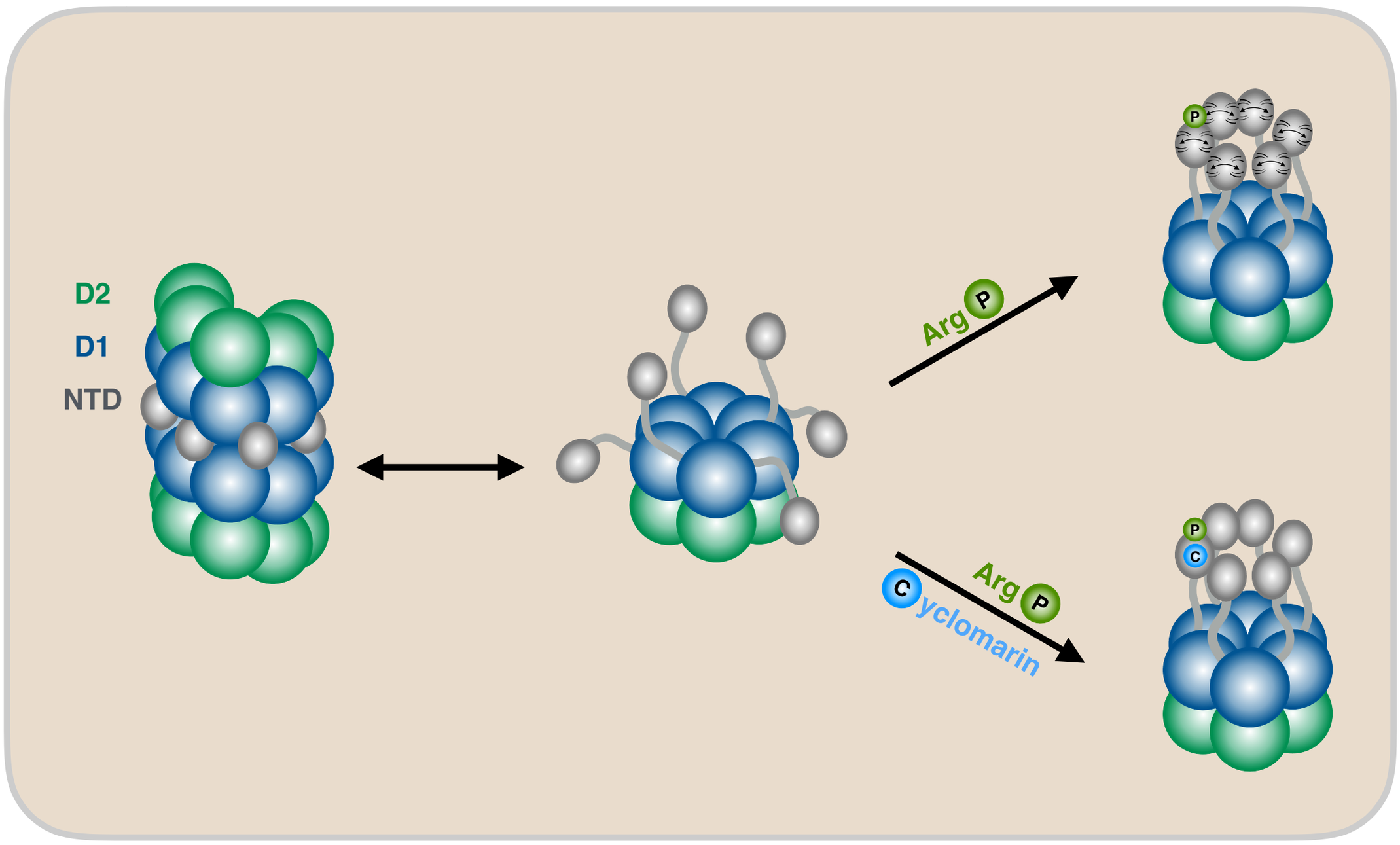


Figure 7

Table S1- Calculated SAXS parameters.

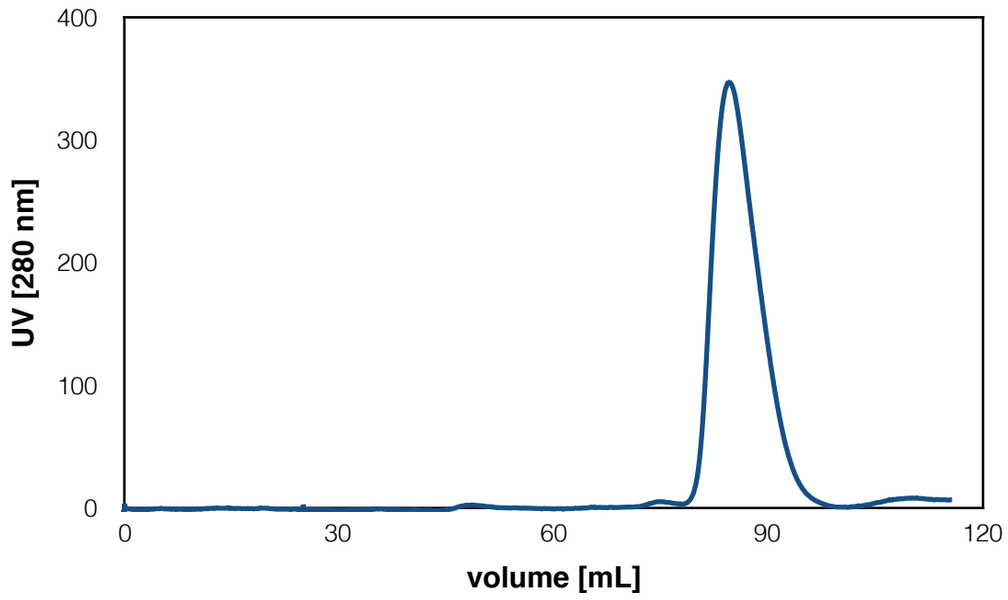
Structural parameters	ClpC1		ClpC1 + ecumicin	ClpC1 + cyclomarin	ClpC1 + ArgP
SEC Peaks[a]	peak	tail	peak	peak	peak
I_0 (cm^{-1}) [from Guinier]	0.121	0.037	0.079	0.109	0.159
R_g (Å) [from Guinier]	78.9 ± 1.5	76.4 ± 2.2	82.1 ± 1.8	81.6 ± 2.7	80.2 ± 1.2
$q_{min}R_g$ – $q_{max}R_g$ used for Guinier	0.75 – 1.27	0.69-1.26	0.86-1.28	0.87-1.30	0.75-1.27
D_{max} (Å)	251	250	280	272.6	275
I_0 (cm^{-1}) [from p(r)]	0.122	0.0358	0.077	0.112	0.159
R_g (Å) [from p(r)]]	79.8	77.0	81.2	84.8	81.4
Porod exponent [from Scåtter]	4.0	4.0	4.0	4.0	4.0
Porod volume V_p (10^3 $Å^3$) [from Scåtter]	2416 ± 100	2156± 20	2351± 20	2408 ± 20	2176 ± 20
Molecular mass M_r (kDa) [from V_p],	1400	1250	1382	1416	1088-1451
Correlated volume V_c ($Å^2$)	3326	2851	3655	2891	2893
Molecular mass M_r (kDa) [from V_c] , assuming only protein	1100	860	1200	840	840
χ^2, crysol3	n.a.	3.34[b]	n.a.	n.a.	n.a.
χ^2, Dammif	2.23[c]	1.59[c]	2.11[c]	3.73[c]	2.71[c]

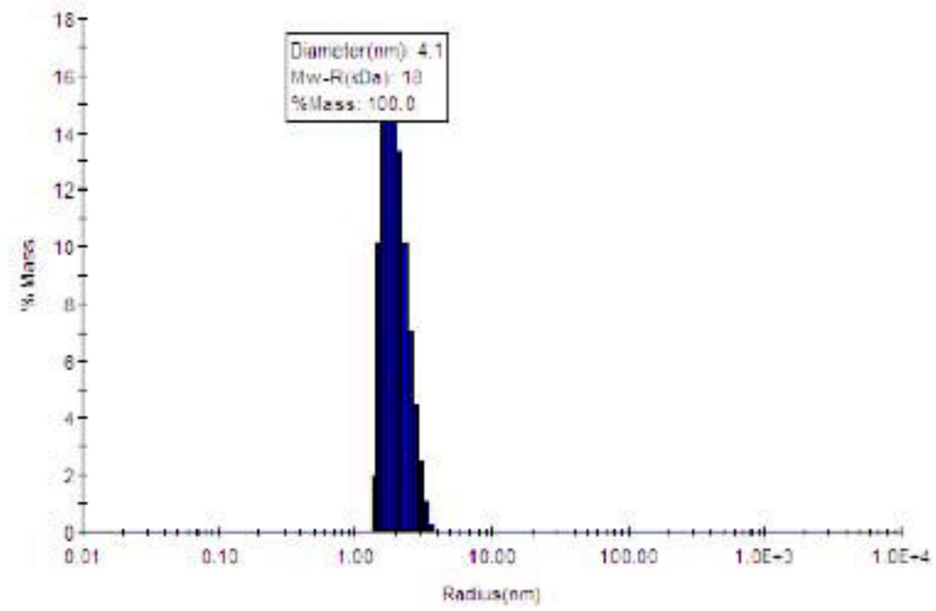
[a]- For apo ClpC1 we analysed the top and the tail of ClpC1 main peak. For cyclomarin, ecumicin and ArgP samples due to baseline correction issues in the tail of the peak only the top of ClpC1 peak was analyzed.

[b]- The predicted scattering curve (Crysol) of the atomistic model of *Bs*ClpC resting state was fitted to the experimental data.

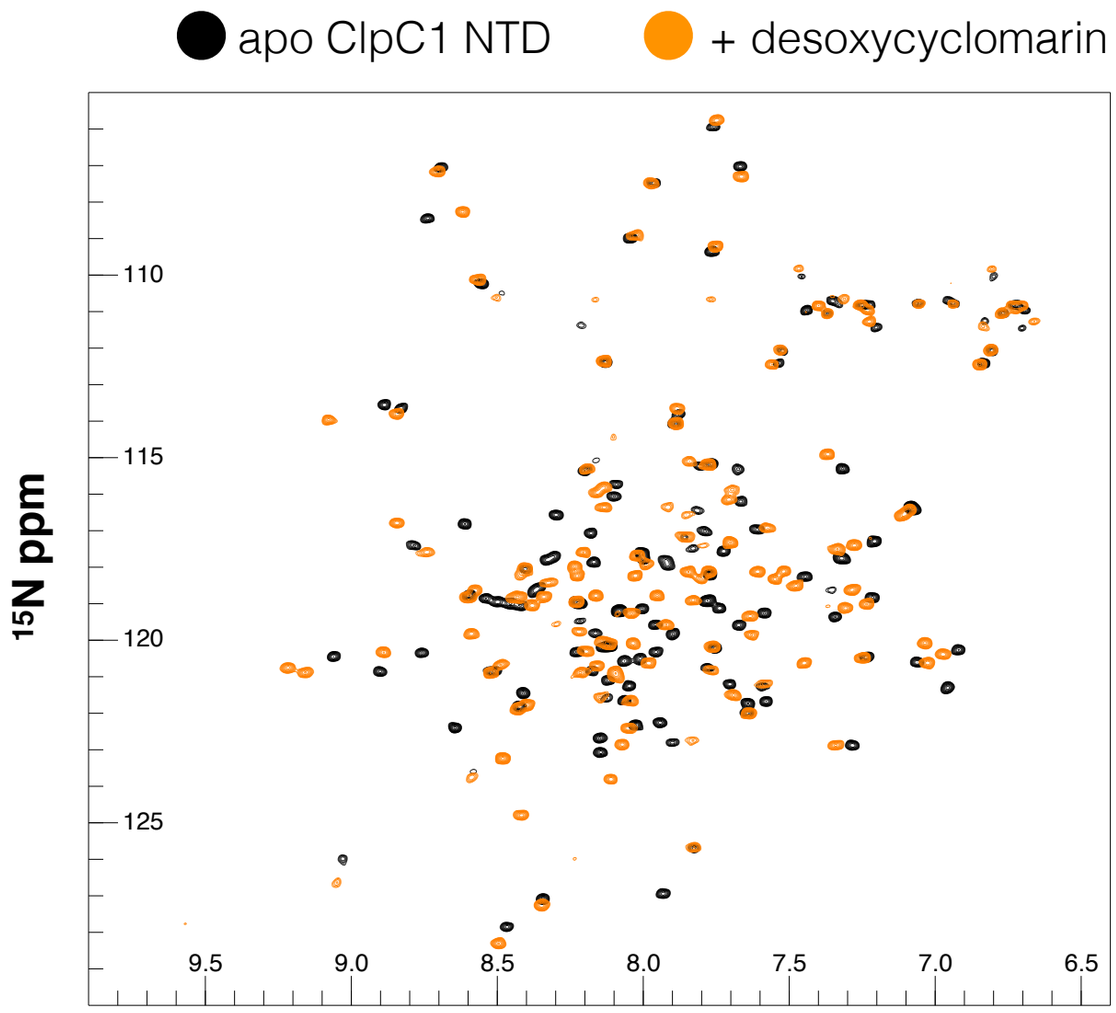
[c]- The scattering curves obtained from the DAMMIF models were fitted to the experimental scattering curves.

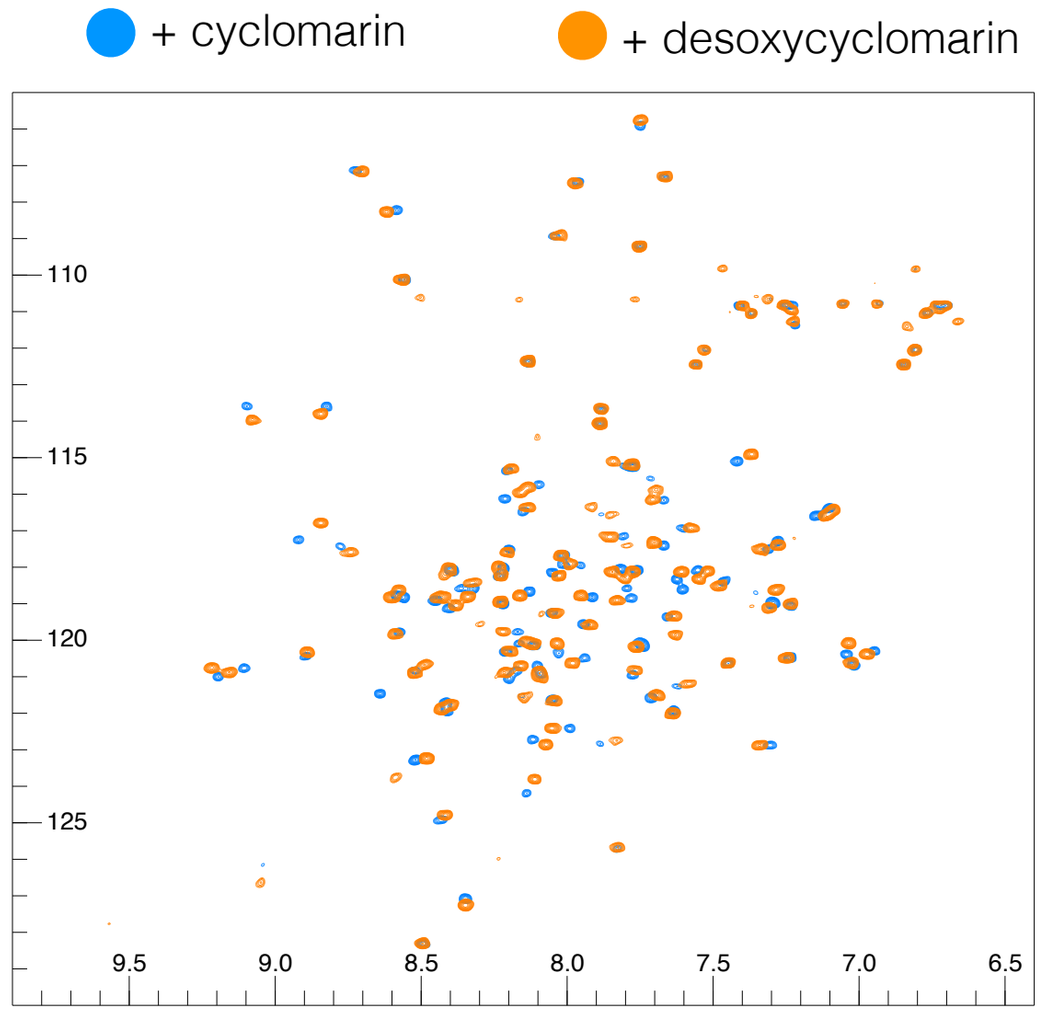
a



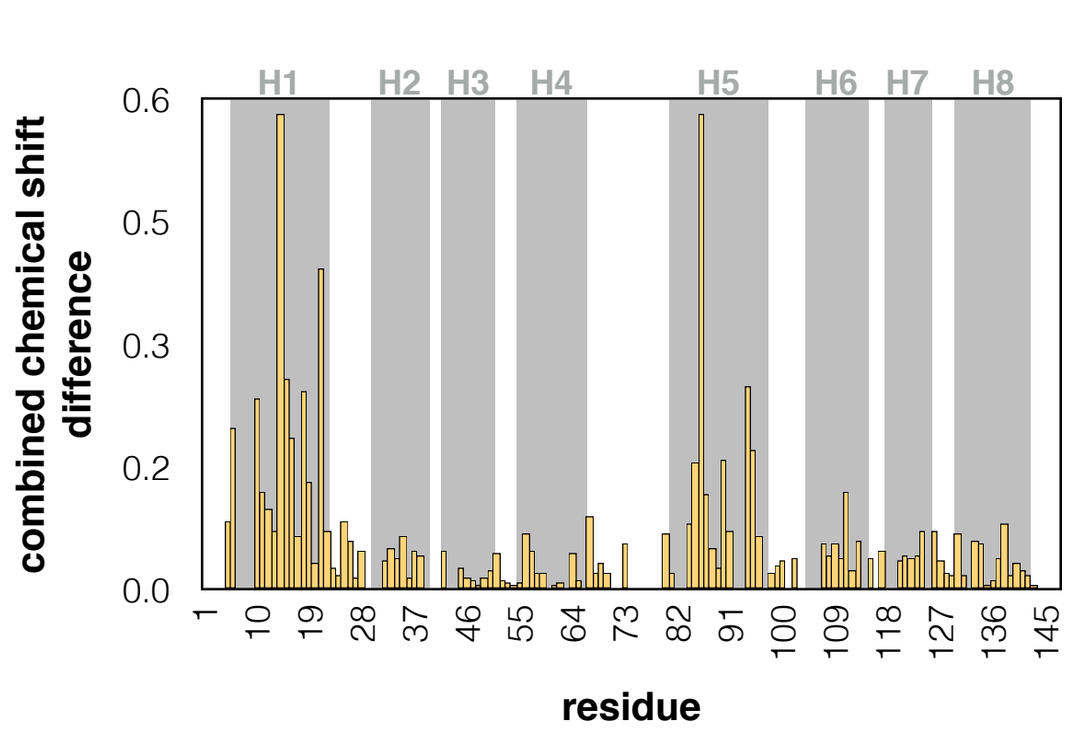


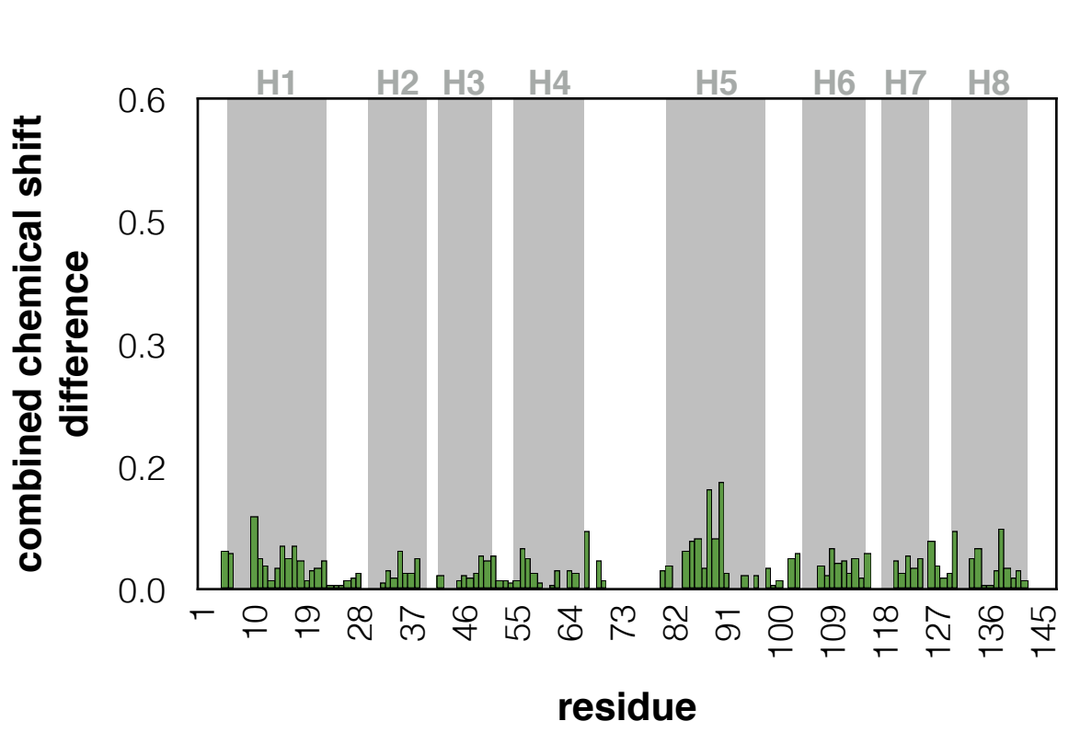
b



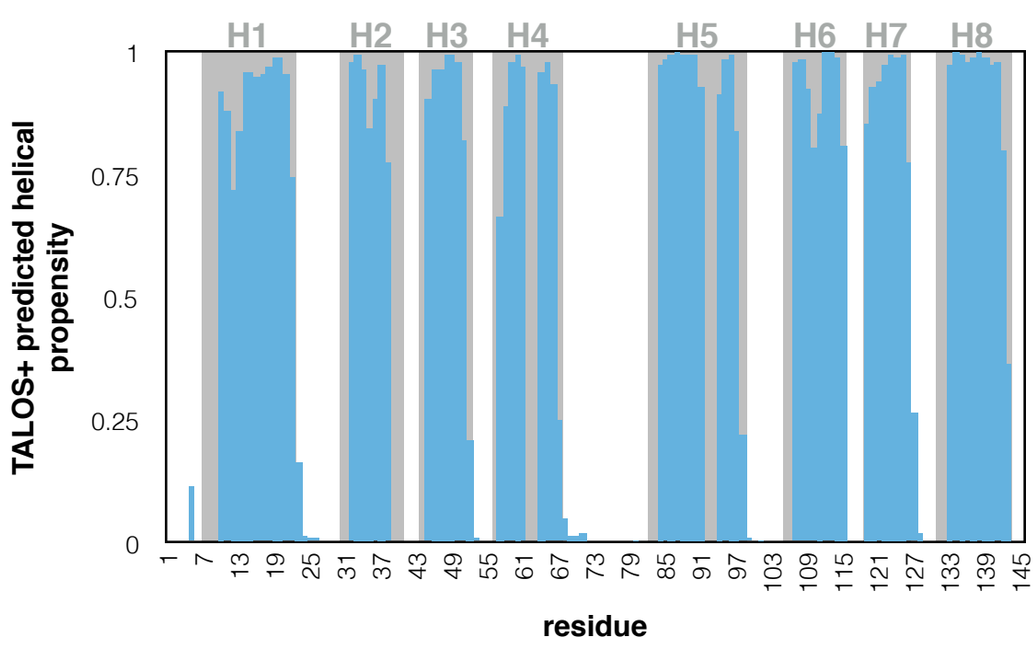


c

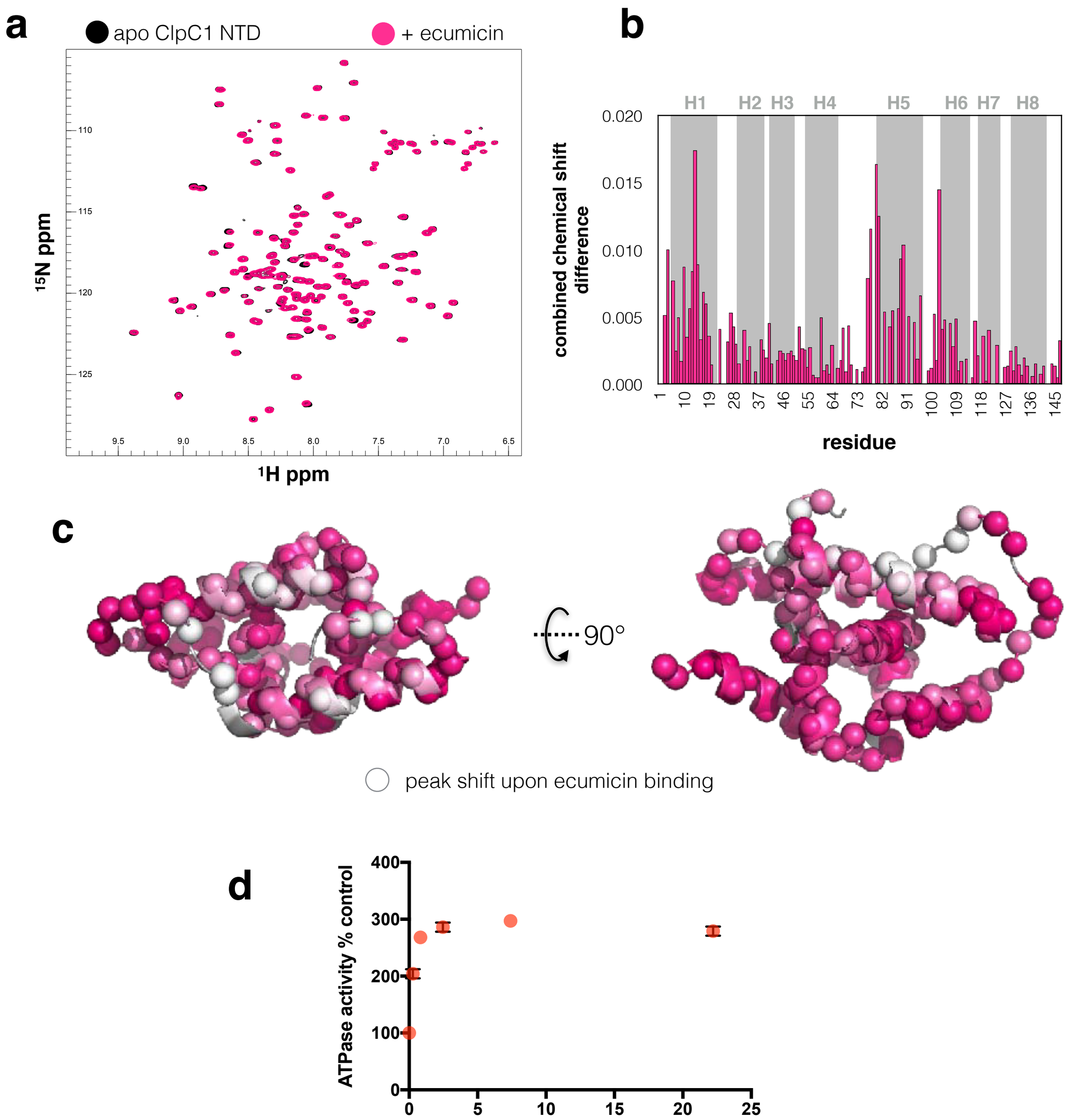




d

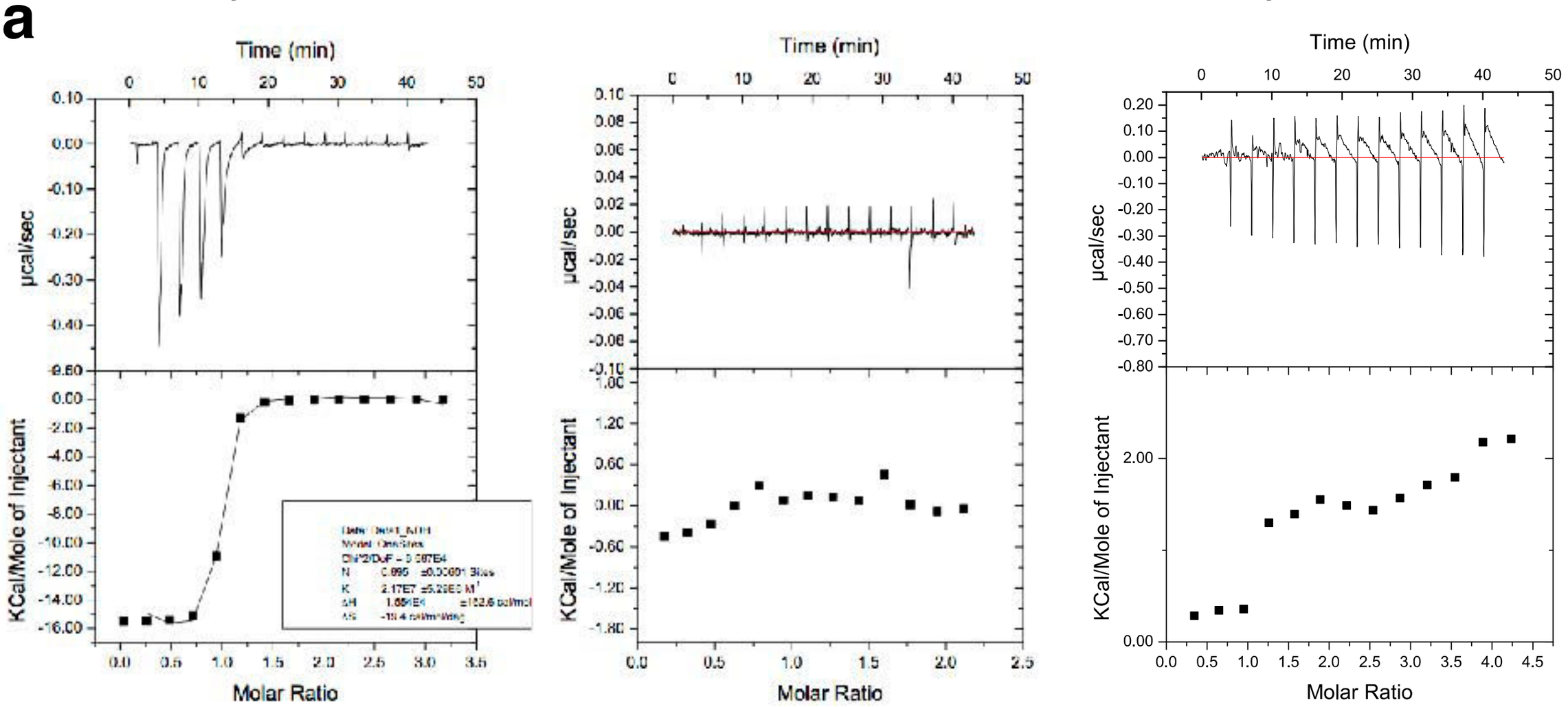


Supporting Figure 1: a) Left - SEC of purified NTD. Right - Dynamic Light Scattering profile of *Mtb*ClpC1 NTD 1 mg/mL. No differences were observed with the addition of ArgP 1 mM. b) Left - ^{1}H-^{15}N correlated backbone NTD amide spectrum for apo (black) and desoxycyclomarin bound (orange). Right - ^{1}H-^{15}N correlated backbone NTD amide spectrum for desoxyclyclomarin (orange) and cyclomarin bound (blue). c) Left - CSP induced by desoxycyclomarin versus apo NTD. Right - CSP induced by desoxycyclomarin versus cyclomarin bound NTD. d) TALOS+ predicted secondary structure derived from NMR assignments of cyclomarin bound *Mtb*ClpC1 NTD amide, CO, Cα and Cβ shifts.

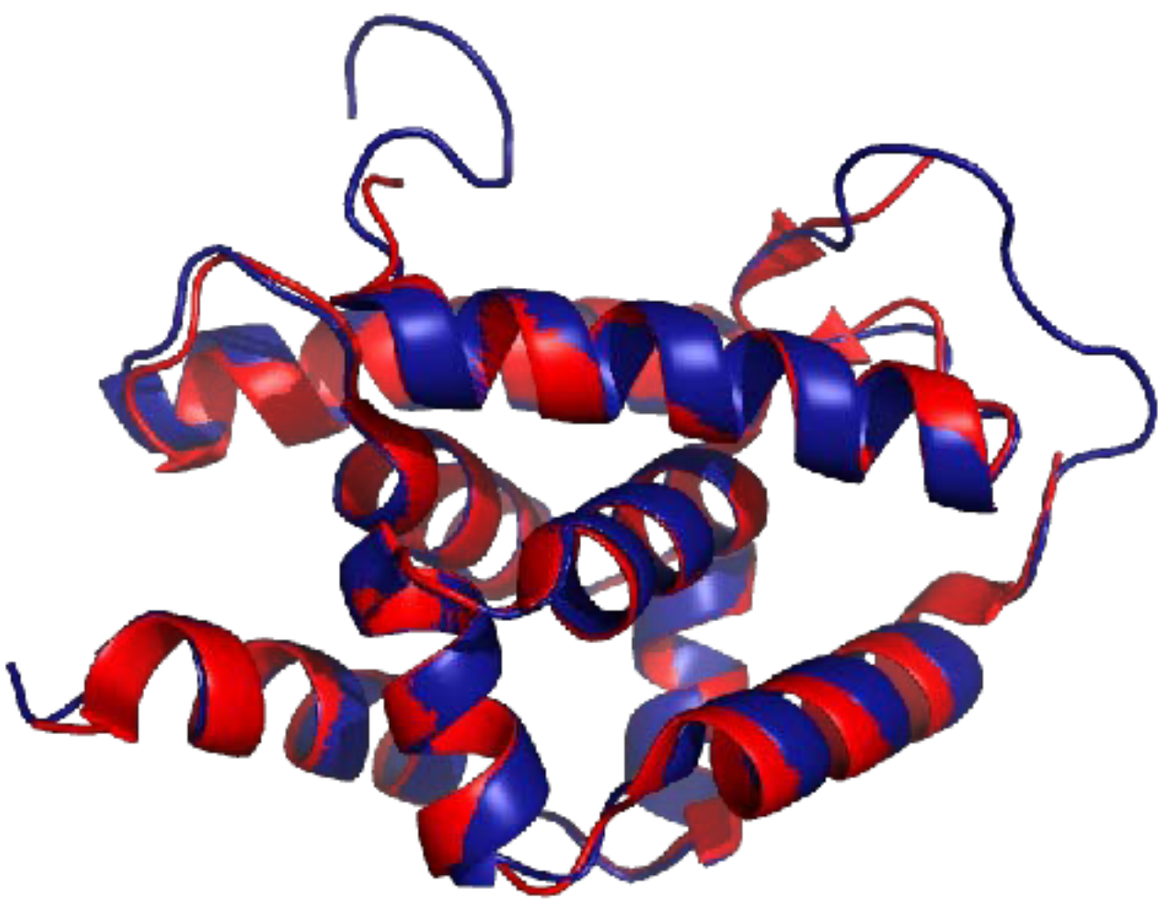


Supporting Figure 2: Ecumicin binding results in only modest changes in the backbone spectra of *Mtb*ClpC1 NTD compared with cyclomarin (see Fig2a). a) ^{1}H-^{15}N correlated backbone amide spectrum for apo (black) and ecumicin-bound (pink) *Mtb*ClpC1 NTD. b) CSP induced by ecumicin, compare with Fig2c for cyclomarin induced CSP. c) Ecumicin induced CSP are mapped onto the NTD structure. d) Ecumicin is able to activate *Mtb*ClpC1 ATPase activity. Results expressed as a ratio to the control experiment.

a



b

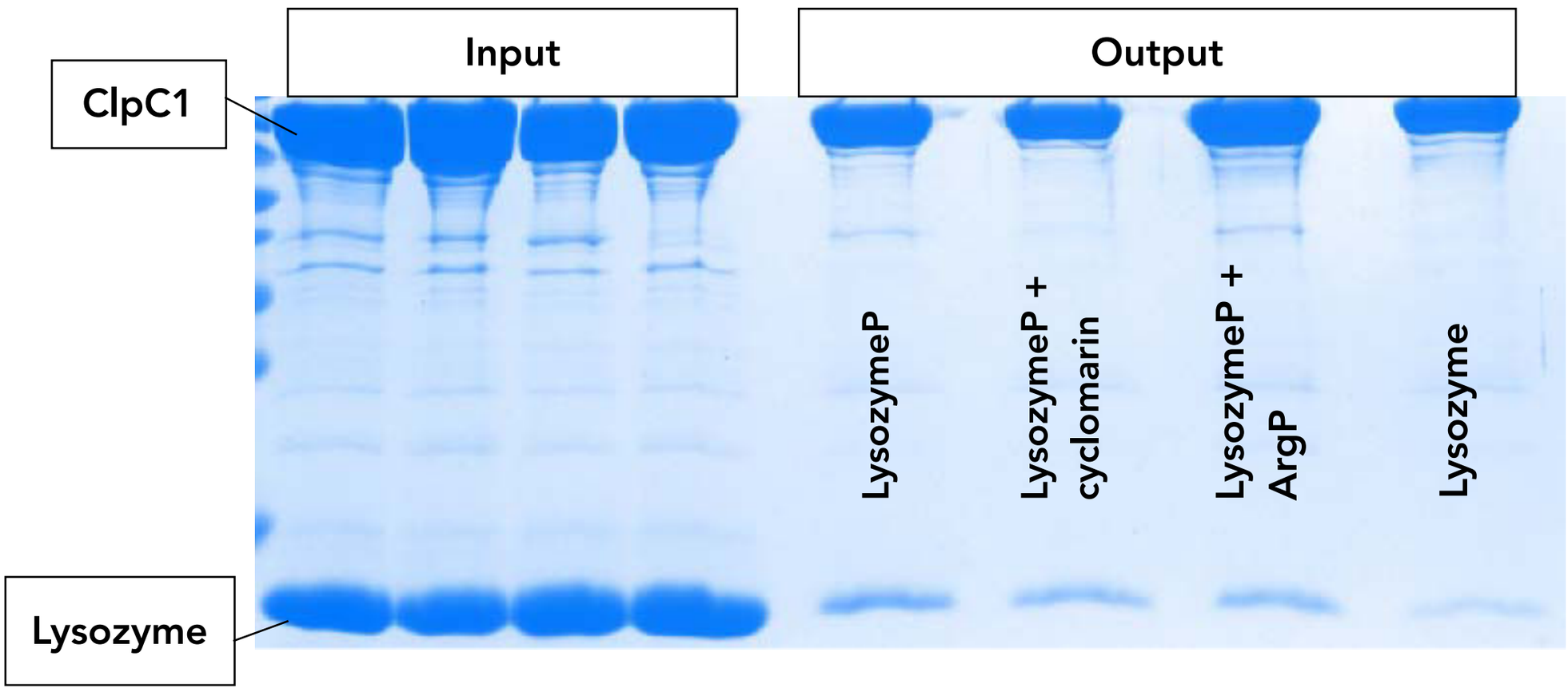


ClpC NTD *Bacillus subtilis*

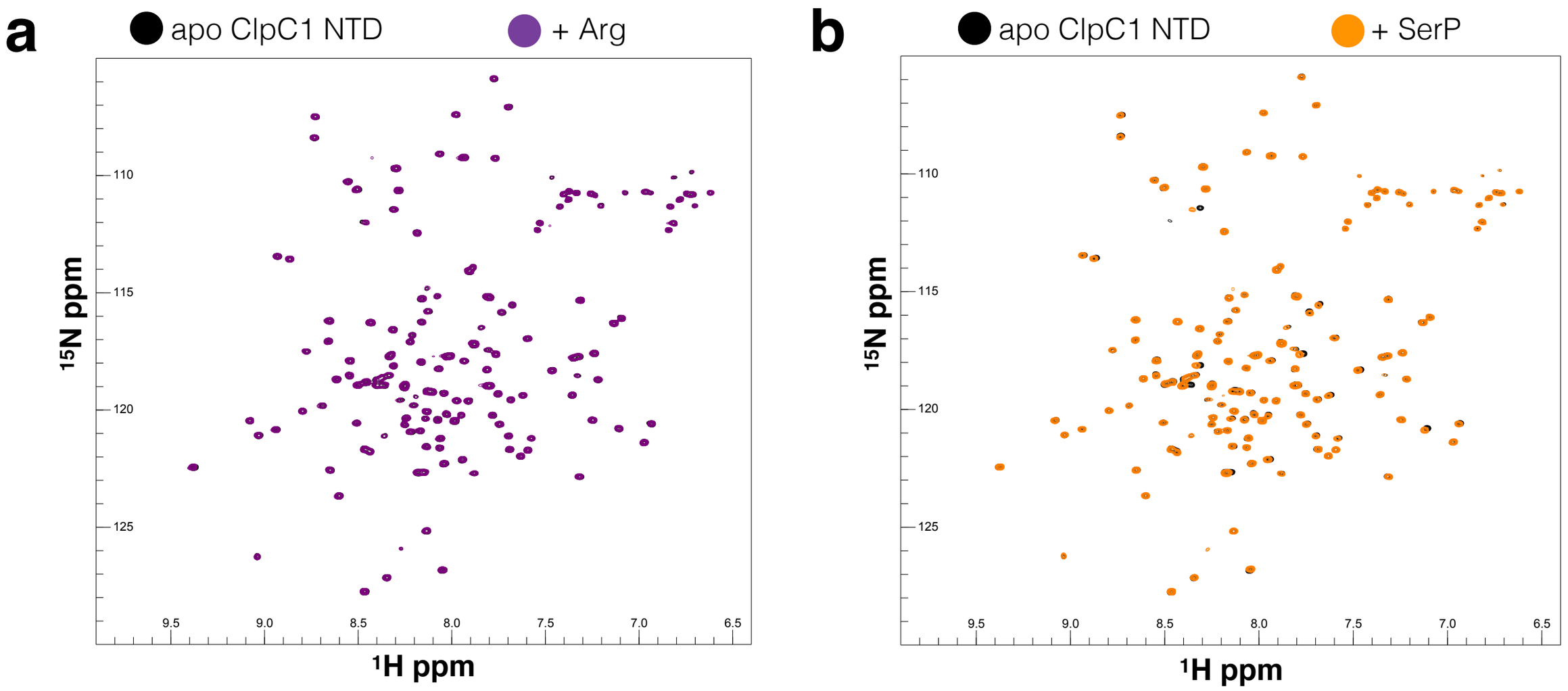
ClpC1 NTD *Mycobacterium tuberculosis*

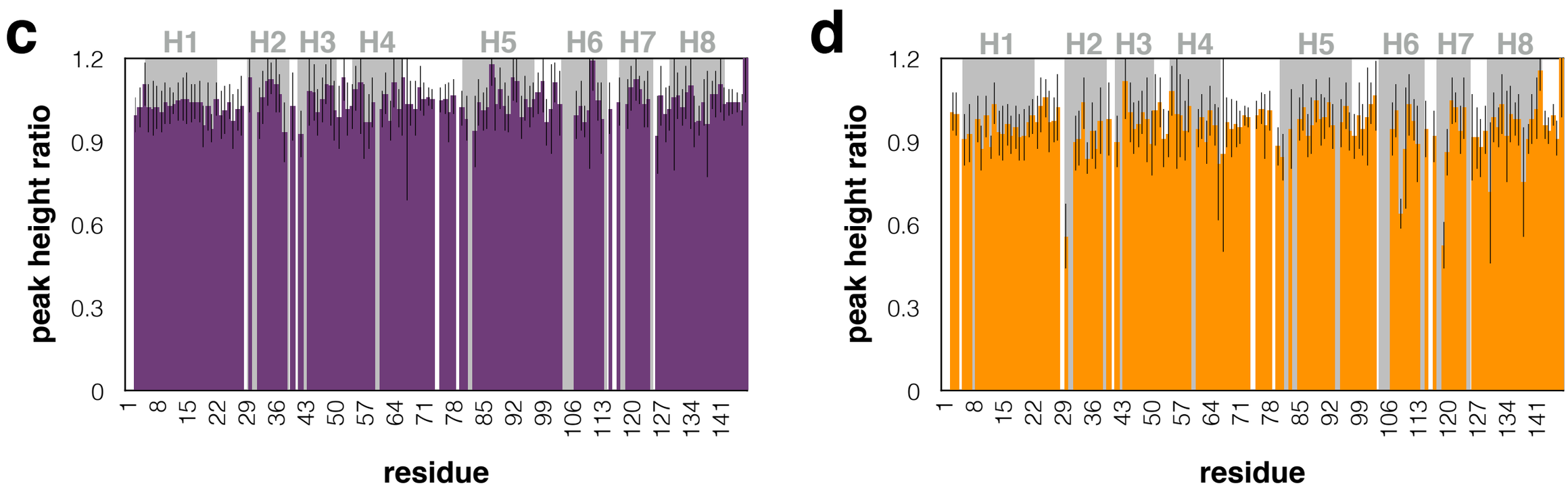
RMSD = 1.7 Å

c



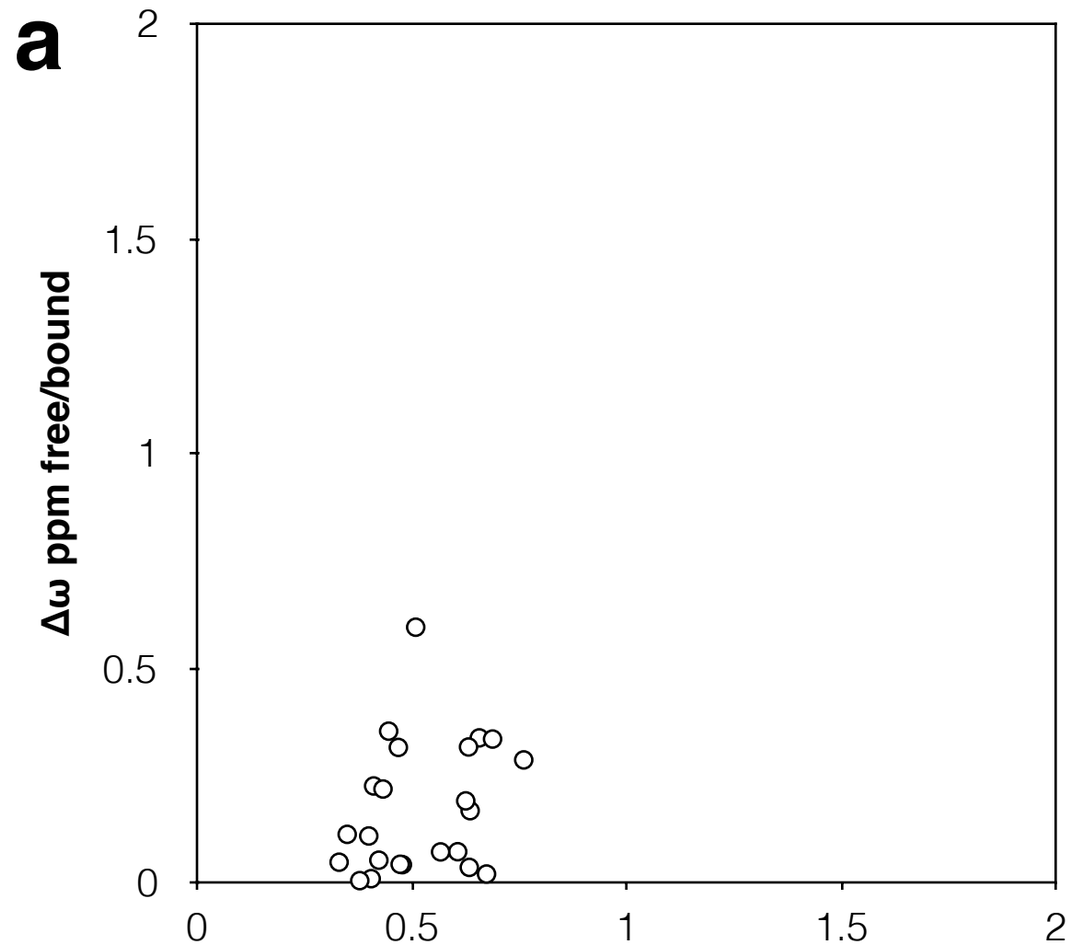
Supporting Figure 3: a) Representative ITC curves of cyclomarin, ecumicin and pyrazinamide binding to *Mtb*ClpC1 NTD. b) Alignment of *B. subtilis* (red) and *M. tuberculosis* (blue) ClpC1 NTD domains. Of note the high structural similarity with a RMSD of 1.7 Å. c) ClpC1 is able to pull down lysozyme phosphorylated by McsB kinase. Cyclomarin was used at 100 μM and ArgP at 2 mM.



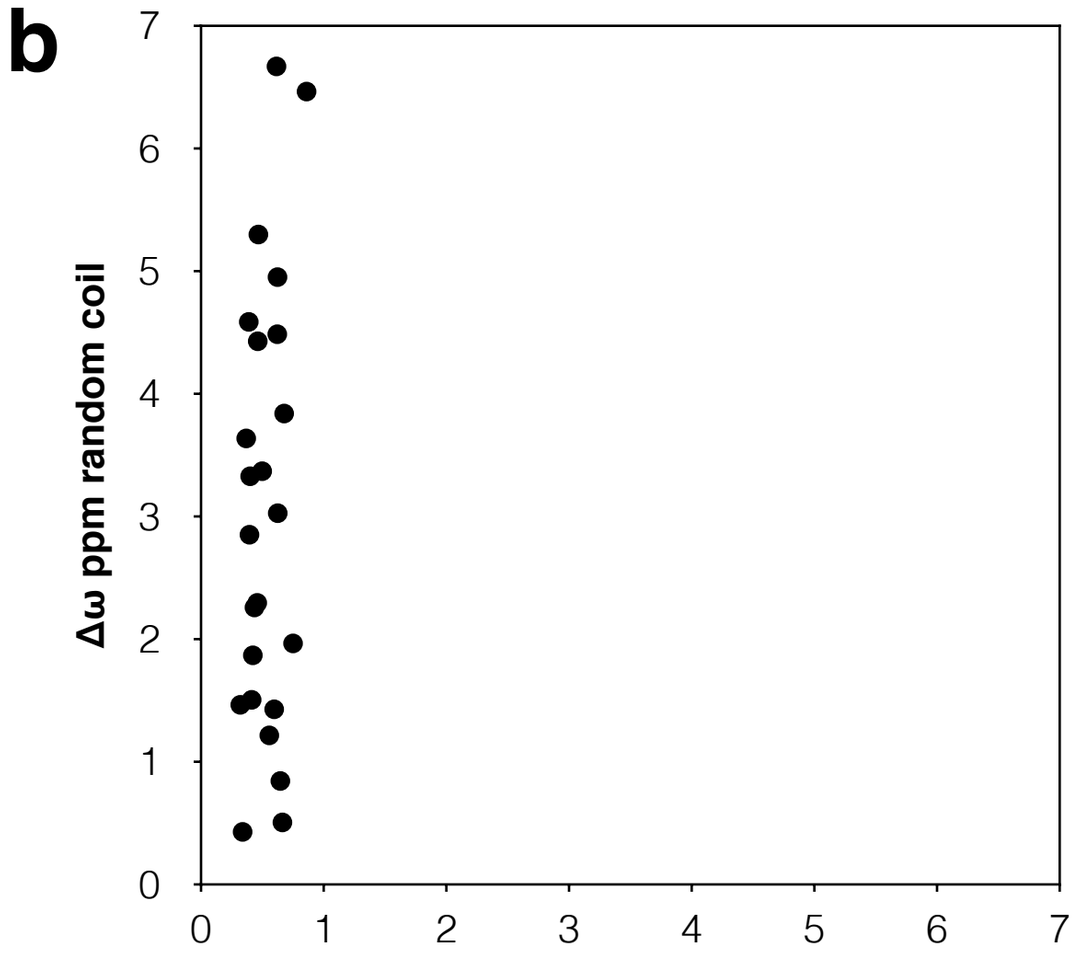


Supporting Figure 4: Arginine and SerinePhosphate do not bind to *Mtb*ClpC1 NTD. a) ^{1}H-^{15}N correlated backbone amide spectrum for apo (black) and arginine (purple). b) ^{1}H-^{15}N correlated backbone amide spectrum for apo (black) and serine-phosphate (orange) bound *Mtb*ClpC1 NTD. c) Intensity changes induced by arginine. d) Intensity changes induced by serine-phosphate.

a



b



c

residue	Δω chemex fit [ppm]	error
5	0.66	0.04
10	0.69	0.04
11	0.41	0.03
14	0.40	0.03
16	0.43	0.03
20	0.51	0.04
22	0.45	0.07
24	0.48	0.03
27	0.40	0.03
54	0.61	0.03
55	0.35	0.02
64	0.87	0.07
71	0.76	0.04
86	0.33	0.04
90	0.47	0.04
91	0.57	0.10
96	0.63	0.04
97	0.42	0.03
102	0.47	0.02
103	0.63	0.03
131	0.67	0.04
132	0.63	0.14
136	0.62	0.03
142	0.38	0.02

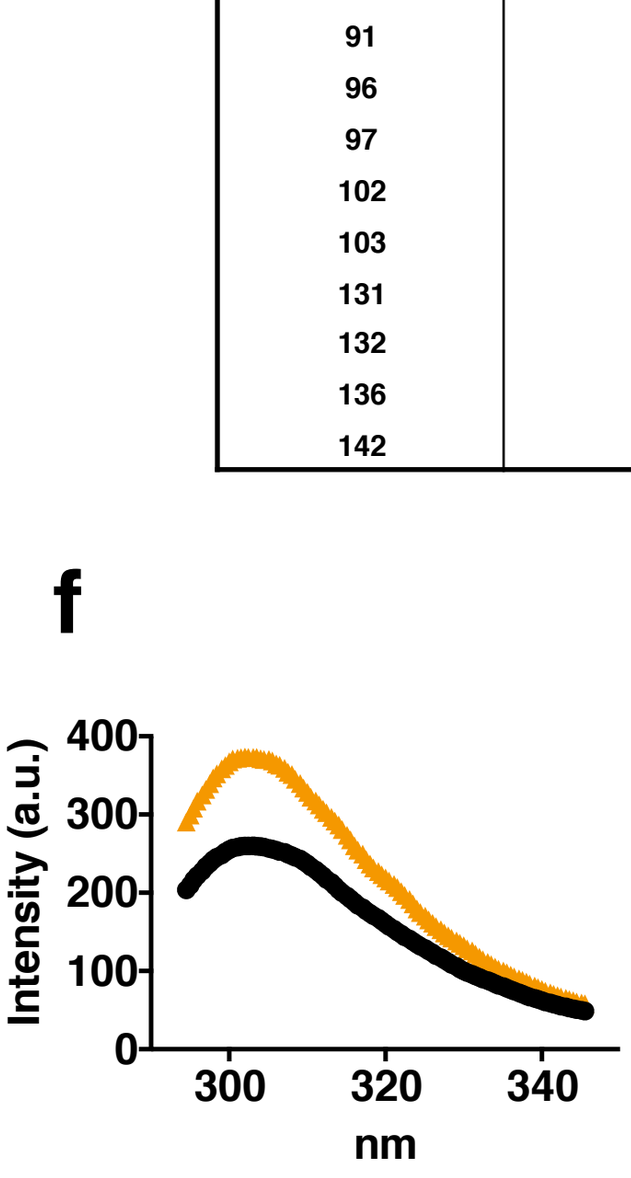
d

k_{ex}	1300 ± 50 s^{-1}
p_B	5 ± 3 %
p_{free}	0.3%

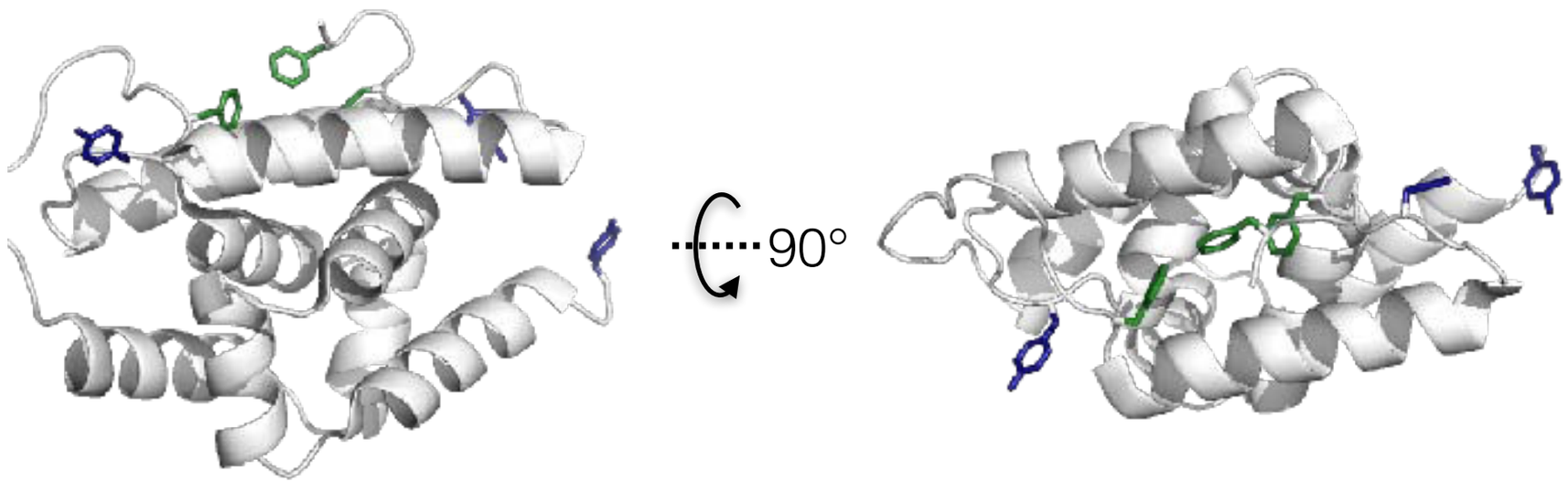
e

compound	diffusion coefficient [cm^2s^{-1}]
DMSO	1.358E-06
ArgP	1.326E-06
Cyclomarin	1.307E-06
ArgP/Cyclomarin	1.325E-06

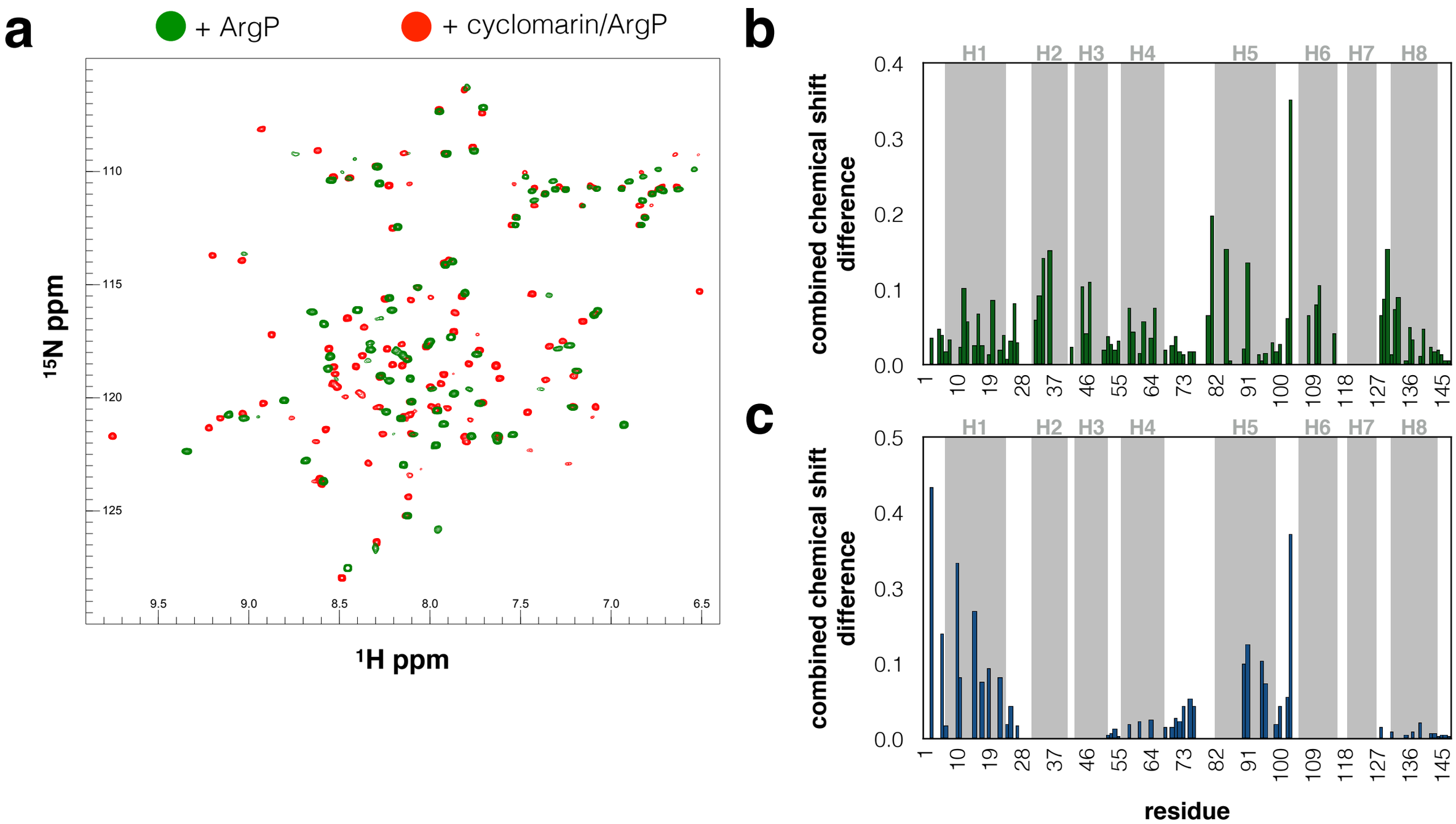
f



g

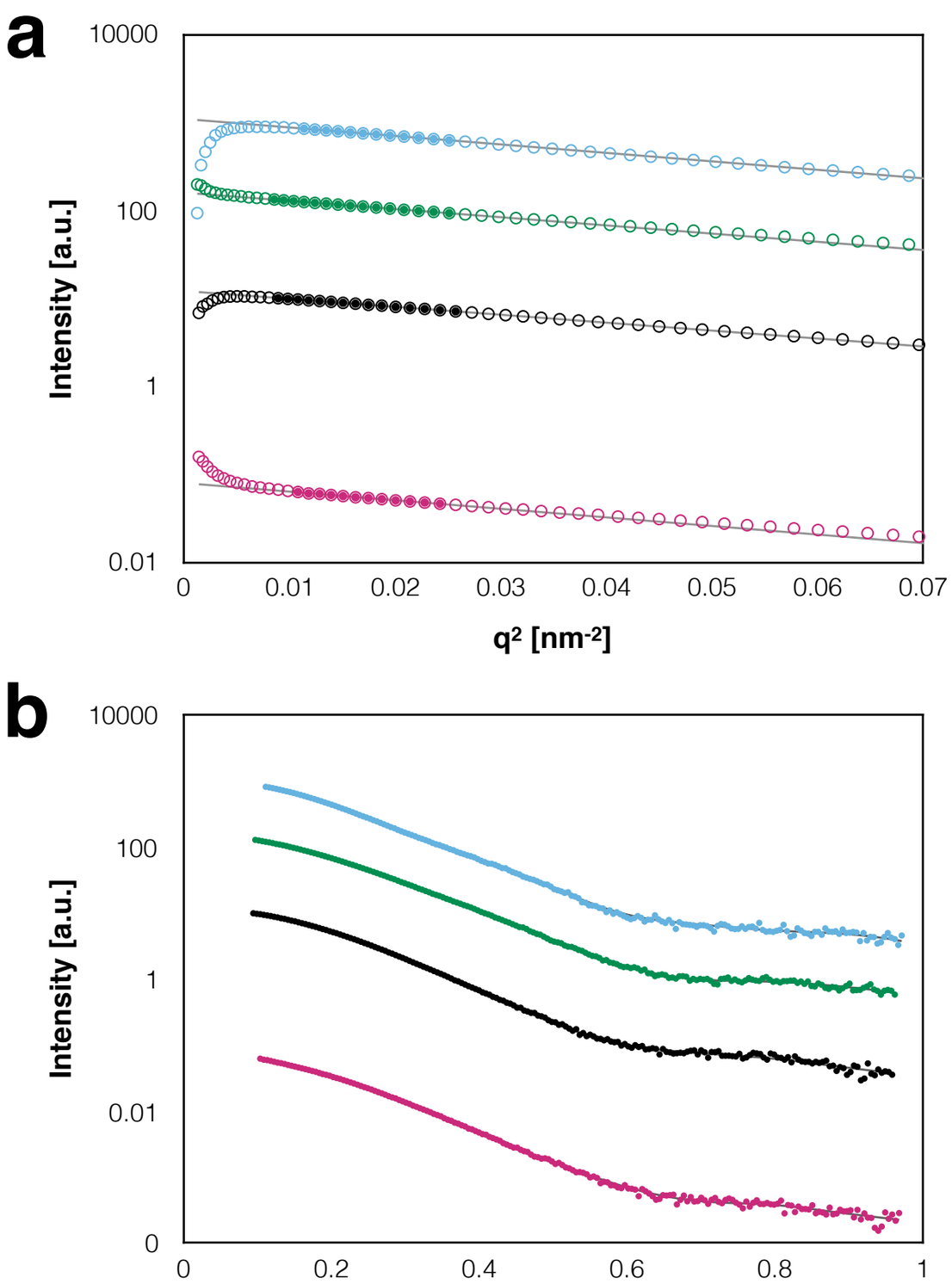


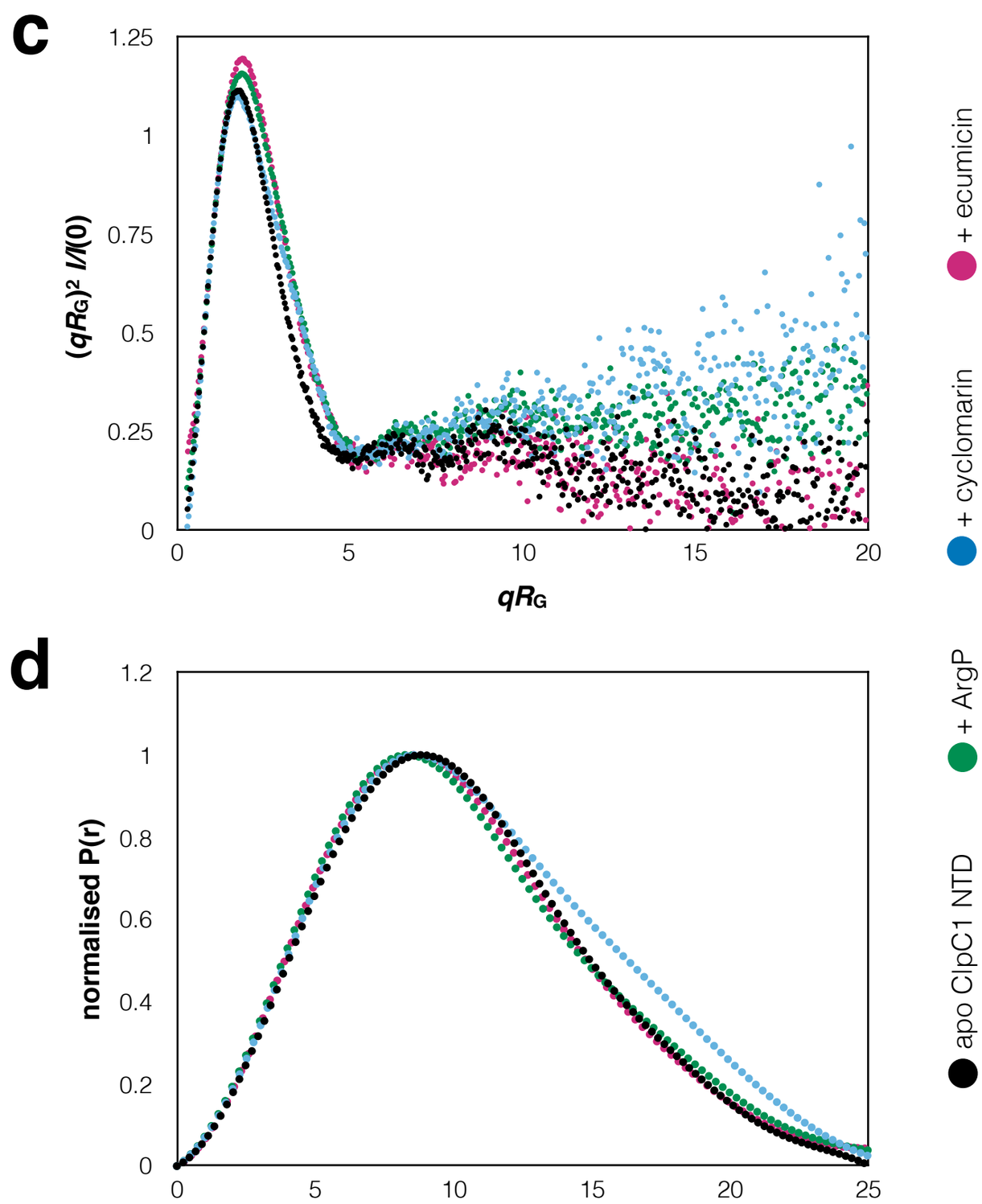
Supporting figure 5: a) Correlation between the theoretical chemical shift differences of the excited state derived from fitting of the CPMG RD data and experimental chemical shift differences between apo and bound state. b) Correlation between the chemical shift differences derived from Chemex fits of the CPMG RD data and chemical shift differences between the folded NTD and calculated random coil shifts derived from ncIDP. c) Table of chemical shift differences of all fitted residues from Chemex CPMG RD data fit. d) Parameters of the Chemex CPMG RD fit. e) Diffusion coefficients of *Mtb*ClpC1 NTD in the presence of different ligands, derived from diffusion ordered spectroscopy (DOSY) NMR measurements. f) Intrinsic fluorescence spectrum of apo (black) and ArgP bound *Mtb*ClpC1 NTD (orange). g) Aromatic residues distribution in *Mtb*ClpC1 NTD - tyrosine side chains in blue, phenylalanine side chains in green.



Supporting Figure 6: a) ^{1}H-^{15}N correlated backbone amide spectrum for ArgP (green) and ArgP/cyclomarin bound NTD (red). b) CSP between cyclomarin bound and cyclomarin/ArgP bound spectrum (for the respective ^{1}H-^{15}N correlated backbone amide spectra see Fig3c). c) CSP between an ArgP and ArgP/cyclomarin bound NTD spectrum. The CSPs coincide with the binding site of cyclomarin (see Fig2).

Supporting Figure 7: CPMG RD fit curves of all ArgP bound *Mtb*ClpC1 NTD residues used for Chemex fitting. The residue number is shown in the upper right corner.





Supporting Figure 8: a) Guinier plots, b) scattering curves, c) Kratky plots and d) pair-distance distribution functions of *Mtb*ClpC1 in the different conditions tested: apo (black), cyclomarin (blue), ecumicin (pink) and ArgP (green).

Appendix C

Summary in french

Le génome mitochondrial ne code que pour environ 1% de toutes les protéines mitochondriales. Les protéines restantes sont traduites dans le cytosol et doivent ensuite être importées dans les mitochondries. Dans ce but, les mitochondries ont mis au point un mécanisme d'importation complexe composé de cinq voies d'importation différentes : 1) la voie préséquence, 2) la voie des transporteurs mitochondriaux ou “mitochondrial carriers” (MC), 3) la voie Mia40, 4) la voie des tonneaux béta ou “beta barrels” et 5) la voie MIM. Les voies 2 et 4 dépendent du chaperon TIM910. La porte d'entrée générale pour l'importation des protéines mitochondriales est le complexe translocase de la membrane externe (TOM). Les voies 1), 2), 3) et 4) ainsi qu’un sous ensemble de la voie 5) utilisent ce complexe afin de transloquer des protéines à travers la membrane externe. Le complexe TOM est constitué de la protéine Tom40 formant le canal, du récepteur des préséquences Tom20, du récepteur central Tom22, du récepteur de MC Tom70 et de trois petites protéines Tom nécessaires à la stabilité et à l'assemblage du complexe.
Un premier modèle de l'importation des précurseurs de MC dans les mitochondries, divisé en cinq étapes consécutives, a été proposé par Köhler *et al.* Lors de la première étape, le précurseur mitochondrial se lie au récepteur TOM dans le cytoplasme. Dans un second temps, le précurseur MC traverse le pore Tom40, qui dépend de la présence de Tim9 et Tim10. Le précurseur MC est ensuite entièrement importé dans l'espace intermembranaire mitochondrial (IMS), puis inséré dans la membrane par le complexe TIM22, un potentiel électrique étant nécessaire pour l'insertion. Le MC atteint finalement son état mature lors de la cinquième et dernière étape. Ce modèle a été étendu en raison de nouvelles découvertes et représente maintenant une fonction encore plus large et plus polyvalente du chaperon TIM910 dans ce réseau d'importation complexe.
Tim9 et Tim10 sont des membres de la famille des petits Tims, comprenant les protéines essentielles Tim9, Tim10 et Tim12 et les protéines non-essentielles Tim8 et Tim13. La structure des petits Tims est hautement conservée et comprend un motif twin CX3C commun qui stabilise un pli hélice-boucle-hélice. Les cinq membres des petits Tims sont exclusivement situés dans l'IMS. Le nombre de petits Tims est variable, tandis que la structure de leur pli est conservée. En effet, certains eucaryotes ne possédant qu’un seul petit Tim ont été découverts, conduisant à la conclusion que tous les petits Tim sont issus d’une seule et même protéine ancestrale, ayant été modifiée et dupliquée au cours de l’évolution. Aucun ancêtre de Tim n'ayant été trouvé chez les procaryotes, Tim a dû être mis en place par la cellule hôte eucaryote afin de faciliter l'importation des protéines membranaires dans les mitochondries.
Le but principal des petits Tims est de transporter des précurseurs de protéines membranaires hydrophobes à travers l'IMS aqueux. Tim9 et Tim10 sont présents dans trois complexes cellulaires différents : le complexe soluble TIM910, le complexe TIM91012 encré à la membrane et à TIM22 et une forme soluble du complexe TIM91012. Le complexe TIM910 soluble est le plus abondant, tandis que le complexe TIM91012 attaché à la membrane est exprimé trois à quatre fois moins. Le complexe TIM91012 soluble est probablement une forme de transport grâce à laquelle Tim12 est transporté de Mia40 vers le complexe TIM22 par TIM910. La nature plutôt dynamique du complexe TIM910 avec ses extrémités Nter et Cter flexibles, caractérisées par des facteurs B élevés, a longtemps empêché la cristallisation du complexe. Ainsi, les premiers aperçus structuraux ont dû être obtenus par SAXS, RMN et dichroïsme circulaire (CD). La première structure cristalline d'un complexe TIM910 humain a été résolue en 2006 par Webb et al. Il a révélé une architecture hexamérique avec une alternance de sous-unités Tim9 et Tim10, comprenant un pli à hélice avec des tentacules

hélicoïdaux. Chaque sous unité contient deux ponts disulfure dans le motif CX3C, formant un pont disulfure interne et externe. Le repliement des deux sous-unités est un motif hélice-boucle-hélice avec les 10-15 derniers résidus de l'extrémité Nter et Cter vraisemblablement dépliés. Les hélices forment un anneau interne (extrémité Nter) et externe (extrémité Cter) autour d'une cavité d'approximativement 15 Â de diamètre. La cavité est recouverte par des chaînes latérales hydrophiles et les hélices des deux sous-unités sont maintenues ensemble par des interactions hydrophobes et deux ponts salins. Cette structure ne présente apparemment aucun site d'interaction préférentiel, ce qui peut amener à s'interroger sur l'éventuelle nécessité d'un réarrangement structural afin de pouvoir lier des substrats.

Les substrats principaux du chaperon TIM910 sont les MC, une grande famille composée de 53 membres chez l'homme et 35 chez la levure, dédiés au transport de métabolites comme les acides aminés, les nucléotides ou les céto-acides à travers la membrane mitochondriale interne. Le transport sélectif des métabolites dans et hors des mitochondries est un processus crucial la survie cellulaire, afin d'assurer l'approvisionnement en énergie et en composants pour la synthèse des acides aminés et des nucléotides. Tous les MC possèdent des motifs structuraux communs et conservés constitués de six hélices transmembranaires, chaque seconde hélice étant pliée autour d'un résidu proline. Ces hélices transmembranaires sont reliées par des boucles et contiennent trois hélices supplémentaires du côté de la matrice.

Les protéines en tonneaux béta sont les plus courantes au sein de la membrane externe et sont dérivées de protéines bactériennes. Trois types de ces protéines sont présentes dans les mitochondries : Tom40, le formateur de canal du complexe TOM; Sam50, le pore de la machinerie d'insertion SAM et VDAC, la protéine la plus abondante de la membrane externe mitochondriale et canal pour petites molécules et ions.

Lorsque Webb et al. ont résolu la première structure cristalline d'un complexe TIM910 en 2006, trois scénarios de liaison possibles ont été suggérés. Ils sont notamment basés sur la nécessité d'un désassemblage du complexe à cause du caractère trop hydrophile de la cavité de TIM910 et de sa taille, jugée trop petite pour pouvoir accueillir une protéine substrat. De plus, aucun site de liaison apparent n'a pu être mis en lumière dans la structure de TIM910. Dans le premier modèle, les hélices de la protéine substrat et le complexe TIM910 se dimérisent au cours du transport. Dans un second modèle, les sous unités Tim9 sont déplacées par des hélices de substrat. Enfin, dans le troisième modèle, les "tentacules" (Nter de Tim9 et Tim10) servent de substituts lipidiques afin de transporter le substrat. Plusieurs études des différentes fonctions des terminaisons Nter et Cter des petites protéines Tim existent. Toutes les expériences convergent à l'idée que l'extrémité Cter est importante pour la stabilité du complexe et l'extrémité Nter est nécessaire à la liaison au substrat. Cela a d'ailleurs été démontré pour Tim9, Tim10 et Tim12. Il convient cependant de mentionner que l'effet d'une troncature de l'extrémité Nter ou Cter ne peut être observé que lorsqu'une partie plutôt importante de l'extrémité est retirée (environ un tiers de la protéine complète).

Le but de ce projet de recherche était d'élucider le mécanisme de liaison du chaperon TIM910 et de ses substrats protéiques membranaires. Comme décrit ci-dessus, l'information structurale et les données biochimiques sur ces complexes sont rares, en raison de la difficulté à travailler avec des substrats hydrophobes susceptibles d'interagir de manière hautement dynamique.

Jusqu'à présent, aucune information sur un substrat lié à TIM910 n'était disponible, les modèles de liaison n'ayant été obtenus qu'avec des tests peptidiques utilisant de courtes séquences de substrat et des hypothèses dérivées de la structure du complexe apo TIM910. Pour ces raisons, la formation du complexe seul, préalable à toute étude biochimique ou structurale, semble déjà poser une défi considérable. Néanmoins, nous avons osé étudier le comportement de ce système insaisissable. Nos objectifs, dans l'ordre, étaient de :

- 1) Décrire la dynamique du chaperon apo hexamérique et de ses sous-unités isolées par RMN
- 2) Former un complexe stable entre le chaperon et le substrat
- 3) Déduire des informations sur la structure et la dynamique du complexe chaperon/ substrat par RMN, idéalement sur le chaperon et le substrat lié
- 4) Déterminer le site de liaison du substrat sur TIM910
- 5) Obtenir des informations sur la stoechiométrie et la structure globale du complexe en utilisant des techniques de plus faible résolution comme le SAXS ou l'AUC
- 6) Comparer substrats protéiques membranaires externes et internes de TIM910, dont la structure est significativement différente lorsqu'ils sont insérés dans la membrane
- 7) Élucider le mécanisme de transfert de substrat à la membrane mitochondriale interne en étudiant la sous-unité Tim12 associée à la membrane.

Pour répondre à tous les points mentionnés ci-dessus, nous nous sommes appuyés principalement sur la RMN, une technique idéale pour l'étude des systèmes dynamiques. Nous avons également combiné la RMN avec différentes techniques biochimiques et biophysiques afin de valider nos résultats et d'obtenir des informations supplémentaires inaccessibles par la RMN.

Le comportement des sous-unités isolées Tim9 et Tim10 observé est en plein accord avec les données publiées. Tim9 forme de préférence des dimères et des monomères tandis que Tim10 est principalement présent sous forme de dimères et trimères (selon les données de filtration sur gel). Les deux sous-unités sont totalement dépliées lorsque les deux liaisons disulfure du motif CX3C sont brisées et se replient partiellement dès que les ponts disulfures sont reformés, faisant du motif CX3C le seul déterminant de leur structure.

Un état rigide complètement replié n'est atteint que lors de la formation du complexe hexamérique, impossible lorsque les sous-unités comprenant un motif CX3C sont réduites. Fait intéressant, les résidus situés à proximité du motif CX3C ne sont pas visibles dans les spectres RMN de Tim9 et Tim10 oxydés, cet effet étant plus prononcé dans Tim9 que dans Tim10. Les résidus proches de cette région invisible présentent une dynamique de l'ordre de la milliseconde, correspondant à des courbes de dispersion CPMG RD non plates.

La détection d'un telle dynamique suggère que les pics correspondant aux résidus non observables sont élargis par le même processus d'échange conformationnel, dont nous n'avons pas pu déterminer la nature. Plusieurs scénarios peuvent expliquer cette observation : l'échange entre différentes espèces d'oligomères, le repliement/dépliement ou encore l'ouverture et la fermeture des ponts disulfure.

A l'instar de ses sous-unités, le comportement du complexe hexamère TIM910 est identique à celui déduit à partir de la structure obtenue par cristallographie des rayons X. Les hélices sont formées dans les régions prévues, la boucle reliant les deux motifs CX3C est plutôt rigide et les parties Nter et Cter invisibles dans la structure cristallographique sont dépliées et flexibles. La partie la plus structurée du complexe est la région centrale comprenant le motif CX3C, la boucle de connexion et la partie hélicoïdale adjacente au motif CX3C.

L'existence simultanée d'espèces hexamères et monomères à également été détectée, alors qu'elle ne l'avait pas été précédemment. Bien que l'espèce principale soit toujours hexamérique, une petite population de monomères est toujours présente et ne peut pas être éliminée par filtration sur gel.
Ceci peut probablement être expliqué par un équilibre entre l'hexamère et le monomère, qui se manifeste également par une légère asymétrie du pic de filtration sur gel du côté correspondant aux espèces plus petites. Dans les spectres RMN, cet équilibre est traduit par un dédoublement des 5-10 derniers pics N-terminaux et C-terminaux. Ce dédoublement n'est pas dû à la présence d'une seconde conformation du complexe mais correspond au signal d'une petite sous-population de monomères.
L'élément clé de nos études était la formation d'un complexe entre TIM910 et une protéine substrat. Nous experiences d'expression de la protéine substrate GGC *in vitro* en présence de différentes quantités de TIM910 montrent que le chaperon est fonctionnel et capable de maintenir ses substrats en solution d'une manière dépendante de la concentration. Nous avons décidé de produire le substrat dans des corps d'inclusion et de le replier en présence du chaperon TIM910 afin d'optimiser nos rendements. La principale difficulté dans la préparation de ces échantillons est la forte propension du substrat à l'agrégation.
Lors de tous les tests impliquant un repliement par dialyse ou dilution éclair, le substrat protéique membranaire s'est agrégé avant que TIM910 ne puisse s'y lier, même à de très faibles concentrations de substrat et de chaperon élevé.
La cause de ce problème réside dans la nature du chaperon TIM910. En effet, bien que le chaperon TIM910 soit très stable et puisse être facilement replié lorsque les ponts disulfure sont formés, l'intégrité de l'hexamère est facilement influencée par la présence de dénaturants. De faibles quantités de dénaturants (au-dessus de 50 mM de guanidine-HCl) sont capables de dissocier l'hexamère en sous-unités individuelles. Comme notre étude le montre, l'hexamère, plutôt que le monomère, est l'espèce active et donc aucun substrat ne peut être lié lorsque la concentration en dénaturant est trop élevée.

Pour la plupart des substrats, la concentration en guanidine-HCl nécessaire afin d'empêcher l'agrégation est beaucoup plus élevée que le seuil de guanidine-HCl nécessaire au repliement de Tim910.
Ainsi, l'objectif principal de la méthode que nous avons développée pour la formation de complexes était la prévention de l'agrégation du substrat, tout en réduisant le dénaturant à une concentration où la formation d'hexamère de TIM910 est possible. À cette fin, nous avons mis au point un protocole comprenant la liaison du substrat à une colonne NiNTA dans des conditions dénaturantes, suivi par une élimination progressive du dénaturant en présence de TIM910. Avec cette configuration, l'agrégation du substrat est retardée en présence de TIM910 afin qu'il puisse lier le substrat dès qu'il atteint le seuil de formation de l'hexamère. Bien que nous n'ayons pas réussi à effectuer une chromatographie d'exclusion stérique sur des complexes substrat / TIM910, nous avons analysé les échantillons par AUC et constaté qu'ils sont stables et homogènes.
Nous n'avons pas d'explication claire quand à l'absence la destruction du complexe, pourtant très stable, lors de la filtration sur gel de type Superdex.. Une possibilité envisagée pourrait être que le substrat interagit avec la résine de filtration sur gel et favorise ainsi la dissociation du complexe. Ce problème ne se pose pourtant pas lors de l'utilisation de colonnes de type

Sec5 Agilent mais leur résolution est trop faible à de tels poids moléculaire. Il n'est donc pas possible de les utiliser afin déterminer l'homogénéité des échantillons.
La structure secondaire dominante dans l'apo TIM910 est l'hélice, avec quelques parties courtes non structurées aux extrémités Nter et Cter. En conséquence, les spectres CD de l'apo TIM910 montrent des propriétés alpha-hélicoïdales. A 25 ° C, les spectres CD d'apo TIM910 et TIM910 liés à AAC ou GGC sont presque identiques. Ce n'est que lorsque les spectres sont enregistrés à 35 ° C que la partie non structurée de l'apo TIM910 augmente en accord avec les données RMN, montrant que la population d'hexamères repliés diminue avec l'augmentation de la température jusqu'à ce que le complexe se déploie complètement. Les caractéristiques des spectres CD de TIM910 avec GGC ou AAC lié à 25 °C ou à 35 ° C sont d'autre part identiques, impliquant que la liaison des substrats stabilise une conformation plus repliée de TIM910. Bien qu'il ne soit pas possible de distinguer l'état du substrat de celui du chaperon par CD, nous supposons que le substrat est au moins en conformation partiellement hélicoïdale, car le spectre ne montre pas de caractéristiques correspondant à des protéines dépliées ou en feuillet-beta . Naturellement, tous les spectres CD des complexes TIM910 / substrat sont dominés par la contribution de TIM910 en raison des différences de taille du chaperon et du substrat, il est donc difficile d'émettre des hypothèses sur la structure secondaire du substrat.
Pour déterminer la localisation du site par lequel TIM910 lie ses substrats, nous avons produit des échantillons protéiques marqués isotopiquement pour l'étude par RMN, détectables respectivement du côté chaperon ou du côté substrat. Les spectres RMN de corrélation 1H-15N de TIM910 lié à la protéine substrat GGC non marquée ne contenaient qu'environ la moitié des pics observés sur les spectres de apo TIM910, tandis que les spectres de GGC liée à TIM910 non marqué ne présentaient qu'un dixième des pics prédis.
Comme nous nous attendions à ce que la taille du complexe représente un défi pour la RMN, nous avons essayé de le surmonter en deutérant nos protéines, mais la qualité des spectre restait très mauvaise et ne pouvait donc pas seulement être expliquée par cette taille importante.
De plus, les différentes parties du complexe ne semblaient pas affectées par l'élargissement des pics d'une manière uniforme. L'obtention d'informations résolues sur les résidus individuels présents dans ces échantillons était impossible et nous avons donc décidé d'utiliser un marquage sélectif des chaînes méthyle latérales afin déterminer le site de liaison. Nous avons choisi de marquer les résidus de leucine et de valine de TIM910, en raison de leur distribution sur la structure cristallographique de TIM910, et les résidus isoleucine et alanine sur GGC car leurs fréquences ne chevauchaient pas celles des autres résidus marqués méthyles dans les spectres RMN. Par la suite, nous avons également préparé des échantillons TIM910 qui ont été marqués sur les alanines, les leucines et les valines pour augmenter le nombre de sondes pour les expériences de RMN. Pour TIM910, un seul ensemble de pics a été observé dans les spectres d'apo TIM910 et de TIM910 lié à GGC, prouvant que la symétrie du chaperon TIM910 n'est pas rompue lors de la liaison au substrat. Les spectres de GGC ne montrent pas de pics résolus individuellement, mais plutôt un mélange de pics élargis, indiquant l'existence de différentes conformations ou états.
Les mesures de relaxation ont révélé que plusieurs des résidus dans TIM910 et GGC subissent un échange à l'échelle de la milliseconde. Le fait que ces deux partenaires de liaison, TIM910 et GGC, subissent ces dynamiques soutient l'idée que c'est le même processus qui peut être expérimenté des deux côtés. Ceci est corroboré par le fitting des données de CPMG,

montrant que la vitesse d'échange dans les deux protéines est la même. Ainsi, la faible qualité des spectres RMN pourrait être expliquée par la liaison dynamique du substrat, qui peut être expérimentée en tant que dynamique d'échange milliseconde sur le chaperon et le substrat. Nous avons utilisé les mesures de type CPMG, PRE, sPRE et de transfert de magnétisation (HETSOFAST) pour valider l'emplacement du site de liaison. En bref, toutes les méthodes convergent vers un même site de liaison au substrat, bien que la précision des différentes méthodes soit assez différente.
Les résidus hydrophobes conservés dans la séquence de TIM910 correspondent à ceux situées au niveau du site de liaison au substrat déterminé par RMN.

De plus, la fente formée entre l'hélice interne et externe de TIM910 est hautement hydrophobe, alors que le pore interne de TIM910 est bordé de résidus hydrophiles. Ainsi, la conservation et l'hydrophobicité du site de liaison supportent notre hypothèse selon laquelle TIM910 lie ses substrats à l'extérieur dans une fente hydrophobe formée par l'hélice N- et C-terminale.

Une des problématiques principales de ce projet était la détermination de la structure d'un complexe entre TIM910 et une protéine substrat. Nous avons écarté la cristallographie aux rayons X car nous savions déjà grâce aux données obtenues par RMN que ces complexes sont très dynamiques, ainsi que la cryo-EM car le complexe était trop petit et la RMN car cette fois le complexe était trop gros. Nous avons donc dû nous appuyer sur des données biochimiques et des techniques structurelles à basse résolution pour pouvoir obtenir une idée de l'aspect de ce complexe. La première question à laquelle nous avons essayé de répondre est la stoechiométrie du complexe.
Nous avons montré que TIM910 ne subit pas de changements structuraux majeurs lors de la liaison au substrat et que la longueur du substrat déplié détermine la stoechiométrie du complexe. Les GGC et AAC sauvages contiennent six hélices transmembranaires et ont un poids moléculaire d'environ 30 kDa. AACshort est une version courte d'AAC et ne contient que les hélices transmembranaires 2 et 3, avec un poids moléculaire total d'environ 15 kDa. Les profils de sédimentation AUC de TIM910 lié à GGC sont en accord avec un complexe 2:1 entre TIM910 et GGC, tandis que TIM910 lié à AACshort sédimentent comme un complexe 1:1. Nous avons obtenu le même résultat en utilisant des expériences RMN DOSY, TIM910 lié à AAC ou GGC di utilisé comme un complexe 2:1, tandis que TIM910 lié à AACshort avait un taux de diffusion très similaire à l'apo TIM910. Les taux de diffusion du TIM910 lié à la GGC sauvage sont identiques à ceux mesurées individuellement sur TIM910 ou GGC, soulignant la qualité de nos échantillons RMN. Pour pouvoir construire un modèle de nos complexes TIM910-GGC, nous avons collecté des données SAXS en utilisant un protocole NiNTA en ligne pour éviter l'agrégation du substrat et fournir des données SAXS de haute qualité. L'enveloppe SAXS est caractéristiques d’une stoechiométrie 2:1, en accord avec les données AUC, NMR et SDS PAGE. De plus, les deux molécules TIM910 dans l'enveloppe SAXS contiennent encore des cavités libres qui corroborent notre modèle de liaison au substrat, de sorte que la cavité centrale de TIM910 n'est pas occupée par le substrat. Pour construire le modèle SAXS, deux molécules TIM910 sont fixées autour des cavités de l'enveloppe SAXS tandis que GGC est modélisée dans la densité supplémentaire de l'enveloppe. Pour rendre compte de la nature dynamique du GGC lié et obtenir un ensemble de structures, plusieurs modèles GGC ont été générés. Tous les modèles issus des données

SAXS d'origine correspondent à une localisation de la molécule GGC située à l'extérieur du chaperon TIM910.
Le positionnement des deux molécules TIM910 avec les extrémités N-terminales face au motif CX3C n'a pas donné des résultats satisfaisants en accord avec les données expérimentales.
La rotation des deux molécules TIM910 l'une par rapport à l'autre sur un axe perpendiculaire à la cavité ne change pas la qualité des ajustements et donc n'importe laquelle de ces positions pourrait correspondre aux données mesurées.
La localisation de la chaîne GGC autour de TIM910 dans les modèles SAXS est en parfait accord avec le site de liaison au substrat déterminé par RMN. Nous proposons donc que le TIM910 lie de manière modulaire ses substrats dans une fente hydrophobe à l'extérieur du chaperon. Le nombre de molécules TIM910 nécessaires au transport d'un substrat dépend de la longueur du substrat. Cette modularité offrirait au TIM910 la possibilité d'interagir avec une large gamme de substrats de tailles variables.
Ceci est en accord avec les caractéristiques d'un chaperon devant interagir avec des substrats qui diffèrent nettement dans leur structure et leur taille, insérés dans la membrane, et dont les poids moléculaires peuvent varier de 17 kDa à environ 50 kDa.
Les substrats de membranes externes de TIM910 utilisent le même site de liaison que les substrats de membranes internes malgré leurs différences de structure lorsque insérées dans la membrane. En outre, TIM910 se lie préférentiellement aux substrats qui contiennent déjà une structure secondaire, comme on peut le voir en comparant le comportement du peptide VDAC cyclique à celui du VDAC linéaire.

Notre objectif final était de confirmer la localisation du site de liaison en effectuant des expériences d'import *in vivo* chez la levure. À cette fin, nous avons débuté une collaboration avec Nils Wiedemann de l'Université de Freiburg. Pour valider le site de liaison, nous avons choisi des résidus hydrophobes qui, selon nos données RMN, sont impliqués dans la liaison au substrat et les avons mutés en acides aminés hydrophiles. Ainsi, nous avons déterminé que des mutations ponctuelles dans le site de liaison du substrat peuvent provoquer une létalité ou des défauts d'importation chez la levure. Ceci est plutôt inattendu, puisque nous avions spéculé que le substrat interagissait avec TIM910 grâce à une multitude d'interactions faibles. Ainsi, la mutation d'un seul résidu n'aurait pas dû perturber la liaison de manière significative.En tant que projet parallèle, nous avons cherché à résoudre la structure cristalline du complexe hexamérique TIM91012. Bien qu'un premier essai de cristallisation ait révélé trois conditions de cristallisation possibles, d'autres tests d'optimisation n'ont pas donné de cristaux diffractants. En effet, la première condition, qui diffractait déjà à 2 Å, n'a pas pu être reproduite, la seconde pouvait être reproduite mais ne diffractait pas, tandis que la troisième condition n'a pas non plus pu être reproduite. Un obstacle majeur à la cristallisation de TIM91012 pourrait être la dynamique inhérente du complexe et la présence de l'extension C-terminale dépliée. Une tentative d'élimination de cette extension Cterminale par mutation a entraîné l'incapacité de Tim12 à former des complexes et, par conséquent, n'était pas applicable pour d'autres études.

Le but de cette thèse était d'élucider le mécanisme des chaperon des protéines mitochondriales internes et externes de la membrane par le chaperon TIM910 et de définir la structure de ces complexes substrat/chaperon.

Nous avons pu démontrer ici que le site de liaison au substrat de TIM910 est situé à l'extérieur du chaperon dans une fente hydrophobe conservée formée par les hélices intérieures et extérieures. Les substrats intérieurs et extérieurs à la membrane interagissent au niveau du même site de liaison. De plus, une certaine proportion de structure secondaire au niveau de la protéine de substrat favorise la liaison par le chaperon.
TIM910 lie les substrats de manière modulaire, s'adaptant ainsi à la taille du substrat. Les MC, les principaux clients des chaperons, sont liés par deux hexamères TIM910. Nos données RMN et SAXS sur ces complexes nous ont permis d'en construire un modèle structural détaillé. Les mesures de la dynamique à l'échelle de la milliseconde des résidus individuels ont révélé que le substrat et le chaperon subissent des processus d'échange de l'ordre de la milliseconde lorsque le substrat est lié.
Nous faisons l'hypothèse que cet échange est dû à un comportement de liaison dynamique du substrat ainsi qu'à des échanges entre différents sites de liaison ou à des échanges conformationnels à l'échelle de la milliseconde.
En outre, nous avons fourni des informations précieuses sur la structure et la dynamique du chaperon hexamérique et de ses sous-unités isolées. L'échange entre le complexe TIM910 hexamère et la sous-unité Tim12 attachée à la membrane pourrait aider à expliquer le mécanisme de transfert de substrat à la membrane mitochondriale interne

Acknowledgements

I spent nearly four happy years at the IBS in Grenoble. That I enjoyed my stay is because of several people that I would like to acknowledge here.

First of all I would like to thank my supervisor Paul Schanda, who was very supportive and enthusiastic throughout the whole project, but also gave me the freedom to try different things and thus the possibility to learn (sometimes also by making my own mistakes). Another important factor was the good environment in the lab, which was created by all the young, motivated and bright people around. I would like to especially thank Audrey Hessel and Orso Subrini for advice from the very start of my project and for their personal support and friendship. Moreover, I would have never had the possibility to study this long if my family wouldn't have supported me financially and morally throughout all these years. Finally, I would like to thank Hugo Fraga for helping me face many of the difficulties I encountered during this stay and who encouraged me even when things were not working out as well.