TECHNISCHE UNIVERSITÄT MÜNCHEN
TUM SCHOOL OF NATURAL SCIENCES

NMR structural and biochemical studies of the chloroplast proteins OEP21 and Tic40

Kai Philipp Klöpfer, M.Sc.

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Priv.-Doz. Dr. Gerd Gemmecker

Prüfende der Dissertation:

1. Prof. Dr. Franz Hagn
2. Prof. Dr. Bernd Reif

Die Dissertation wurde am 16.02.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 15.03.2023 angenommen.

LIST OF PUBLICATIONS

I hereby declare that parts of this thesis have been submitted to scientific international journals or have been already published:

Günsel, U., Klöpfer, K., Häusler, E., Hitzenberger, M., Bölter, B., Zacharias, M., Soll, J. & Hagn, F. Structural basis of metabolite transport by the chloroplast outer envelope channel OEP21. *The manuscript has been accepted for publication by Nature Structural & Molecular Biology.*

Schuhmann, P., Engstler, C., Klöpfer, K., Gügel, I. L., Abbadi, A., Dreyer, F., Leckband, G., Bölter, B., Hagn, F., Soll, J. & Carrie, C. 2022. Article Two wrongs make a right: Heat stress reversion of a male-sterile *Brassica napus* line. *J Exp Bot*. (DOI: 10.1093/jxb/erac082) [1]

Klöpfer, K. & Hagn, F. 2019. Beyond detergent micelles: The advantages and applications of non-micellar and lipid-based membrane mimetics for solution-state NMR. *Prog Nucl Magn Reson Spectrosc,* 114-115**,** 271-283. (DOI: 10.1016/j.pnmrs.2019.08.001) [2]

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADAM	a disintegrin and metalloprotease
e.g. ADAM17	a disintegrin and metalloprotease 17
ADP	adenosine diphosphate
AMP	adenosine monophosphate
APS	ammonium persulfate
Asna1	arsenical pump-driving ATPase protein 1
ATP	adenosine triphosphate
ATPase	adenosine triphosphatase
BME	β-Mercaptoethanol
Bna	*Brassica napus*
CANDIS	conserved ADAM17 dynamic interaction sequence
CD	circular dichroism
CTP	cytidine triphosphate
CV	column volume
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleoside triphosphate
DPC	dodecylphosphocholine
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
ER	endoplasmic reticulum
FP	fluorescence polarization
GAP	glyceraldehyde 3-phosphate
GB1	protein G B1 domain
GdmCl	guanidinium chloride
GDP	guanosine diphosphate
Get	guided entry of tail-anchored proteins
e.g. Get3	guided entry of tail-anchored proteins factor 3
GMPPNP	5'-guanylyl imidodiphosphate
GTP	guanosine triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hip	Hsp70 interacting protein
Hop	Hsp70-Hsp90 organizing protein
Hsc70	heat shock cognate protein 70
Hsp	heat shock protein
HSQC	heteronuclear single quantum coherence
IC_{50}	half maximal inhibitory concentration
IPTG	isopropyl β-D-1-thiogalactopyranoside
IRHD	iRhom homology domain
iRhom	inactive rhomboid protein
e.g. iRhom2	inactive rhomboid protein 2

ITC	isothermal titration calorimetry
K_D	dissociation constant
LDAO	lauryldimethylamine oxide
MANT	2'-(or-3')-*O*-(*N*-Methylanthraniloyl)
MBP	maltose-binding protein
MES	2-(*N*-morpholino)ethanesulfonic acid
MPD	membrane proximal domain
MSP	membrane scaffolding protein
MWCO	molecular weight cut-off
NADH	nicotinamide adenine dinucleotide (reduced form)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
NaPi	sodium phosphate (NaH_2PO_4 / Na_2HPO_4)
Ni-NTA	nickel^{2+} ions coupled to nitrilotriacetic acid
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
OD_{600}	optical density measured at a wavelength of 600 nm
OEP	outer envelope protein
e.g. OEP21	outer envelope protein 21
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank (https://www.rcsb.org)
PEP	phosphoenolpyruvate
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
SEC	size-exclusion chromatography
Sgt2	small, glutamine-rich, tetratricopeptide repeat protein 2
TACE	TNFα converting enzyme
TCEP	tris(2-carboxyethyl)phosphine
TEMED	tetramethylethylenediamine
TEV	tobacco etch virus protease
TFE	2,2,2-Trifluoroethanol
Tic	translocon at the inner envelope of chloroplasts
e.g. Tic40	translocon at the inner envelope of chloroplasts, subunit 40
TNFα	tumor necrosis factor α
Toc	translocon at the outer envelope of chloroplasts
TPR	tetratricopeptide repeat
TRC	transmembrane domain recognition complex 40
Tris	tris(hydroxymethyl)aminomethane
TROSY	transverse relaxation optimized spectroscopy
UTP	uridine triphosphate
UV	ultraviolet
VDAC	voltage-dependent anion channel
3-PGA	3-phophoglyceric acid

SUMMARY

This work provides comprehensive structural and functions insights into the chloroplast protein called outer envelope protein 21 (OEP21). In addition, structural data on subunit 40 of the translocon at the inner envelope of chloroplasts (Tic40) are presented.

Chloroplasts are the primary sites of photosynthesis in plant cells, with triosephosphates such as glyceraldehyde 3-phosphate (GAP) being the main products of photosynthetic CO_2 fixation. Thereby, these triosephosphates must be transported across the inner and outer envelope membranes into the cytosol for further metabolic processing and energy supply. While transport across the inner envelope membrane is well understood, a detailed picture of the transport proteins in the outer envelope remains elusive. One of these transport proteins is the β-barrel porin OEP21, which forms an anion-selective channel, has an overall cone-shaped structure, and serves as the major exit pore for triosephosphates in C3 plants. In this work, the binding mechanism of OEP21 is studied in detail to provide the basis for studying the functionality of OEP21. Therefore, CD, NMR, ITC, and fluorescence polarization experiments were conducted and demonstrated that negatively charged metabolites can bind to highly positively charged inner surface of OEP21, with the binding affinity dependent on the number of negative charges in the metabolite. Thus, nucleoside triphosphates such as ATP, which is the most relevant for metabolism, bind with significant higher affinity to OEP21 compared to monophosphorylated metabolites such as GAP. In addition, various OEP21 variants were generated in which single positively charged lysine or arginine residues were mutated to alanine, resulting in comparable binding affinity for ATP as for the wildtype. Since the removal of individual positive charges was compensated by other positively charged residues in the proximity, the binding mechanism appears to be rather promiscuous and depends mainly on the bulk electrostatic properties of the OEP21 pore. This was also confirmed by the decreased binding affinity in the presence of increased sodium chloride concentrations and by competition experiments in which ATP binding was studied in the presence of other phosphorylated compounds, particularly GAP. Furthermore, it was shown that the binding affinity of ATP to OEP21 is significantly reduced in the presence of magnesium, which is present in increased concentrations in the intermembrane space, for example. These data suggest that phosphorylated metabolites use the same positively charged binding surface of OEP21 and that dissociation of the high-affinity binder ATP can be achieved at cellular solute concentration levels. Based on the results presented here on the binding mechanism and interplay of different metabolites, metabolite transport through OEP21 was experimentally investigated in further work, which provided a comprehensive understanding of the functionality of OEP21.

Moreover, OEP21 forms larger oligomers in the outer envelope, and it could be shown that oligomerization can be controlled by LDAO concentration *in vitro*. In addition, OEP21 can appear as dimeric form meditated by a cysteine residue that is oriented toward the membrane and well-positioned to form a disulfide bridge under oxidizing conditions. This oxidation-dependent oligomerization of OEP21 in response to oxidative stress may be a critical switch for plant cells in executing the hypersensitive response, a type of programmed cell death to defend against pathogen infections in plants.

The second protein studied here is Tic40, which is a component of the protein import apparatus of the inner envelope of chloroplasts. In this work, structural differences between the fertile BnaA10 and the non-fertile BnaC9 Tic40 variant were identified. Therefore, BnaA10 was investigated by NMR spectroscopy, where ^{13}C chemical shift information revealed eleven α-helical secondary structure elements in BnaA10. In addition, a structural model of BnaA10 was derived by a combined approach using structure prediction and known NMR structural data. This structural model agreed very well with the α-helical secondary structural elements determined using ^{13}C chemical shift information. However, the structural model showed an elongated shape, lacking larger hydrophobic clusters that would be expected for a compactly folded protein. In conclusion, based on the structural model, two mutated positions in the Tic40 variant BnaC9 were identified as a possible reason for the difference in fertility, where BnaC9 appears to exhibit a higher degree of protein flexibility and a weakening of the folding state, causing its infertility.

ZUSAMMENFASSUNG

Diese Arbeit liefert umfassende strukturelle und funktionelle Erkenntnisse über das Chloroplasten-Protein OEP21 (engl. *outer envelope protein 21*). Außerdem werden strukturelle Daten zu Tic40 (engl. *translocon at the inner envelope of chloroplasts*, Untereinheit 40) gezeigt.

Chloroplasten sind die Hauptorte der Photosynthese in Pflanzenzellen, wobei Triosephosphate wie Glycerinaldehyd-3-phosphat (GAP) die wesentlichen Produkte der photosynthetischen CO_2-Fixierung darstellen. Dabei müssen diese Triosephosphate durch die innere und äußere Hüllmembran in das Cytosol transportiert werden, wo sie weiterverarbeitet werden und Energie liefern. Während der Transport durch die innere Hüllmembran gut verstanden ist, ist über die Transportproteine der äußeren Hüllmembran bisher wenig bekannt. Eines dieser Transportproteine ist das β-Fass-Porin OEP21, welches einen Anionen-selektiven Kanal bildet, eine insgesamt kegelförmige Struktur aufweist und als primäre Ausgangspore für Triosephosphate in C3-Pflanzen dient. In dieser Arbeit wird der Bindungsmechanismus von OEP21 im Detail untersucht, um die Grundlage für die Untersuchung der Funktionalität von OEP21 zu schaffen. Dazu wurden CD-, NMR-, ITC- und Fluoreszenzpolarisations-Experimente durchgeführt, die zeigten, dass negativ geladene Metaboliten an die stark positiv geladene innere Oberfläche von OEP21 binden können, wobei die Bindungsaffinität von der Anzahl der negativen Ladungen im Metaboliten abhängig ist. So binden Nukleosidtriphosphate wie ATP, das für den Stoffwechsel am wichtigsten ist, mit deutlich höherer Affinität an OEP21 als mono-phosphorylierte Metabolite wie GAP. Darüber hinaus wurden verschiedene OEP21-Varianten erzeugt, bei denen einzelne positiv geladene Lysin- oder Arginin-Reste zu Alanin mutiert waren, was zu einer vergleichbaren Bindungsaffinität für ATP führte wie beim Wildtyp. Da die Entfernung einzelner positiver Ladungen durch andere positiv geladene Reste in der Nähe kompensiert wurde, scheint der Bindungsmechanismus eher vielschichtig zu sein und hauptsächlich von den elektrostatischen Eigenschaften der OEP21-Pore abzuhängen. Dies wurde auch durch die verringerte Bindungsaffinität in Gegenwart erhöhter Natriumchlorid-Konzentrationen und durch Konkurrenzexperimente bestätigt, bei denen die ATP-Bindung in Gegenwart anderer phosphorylierter Verbindungen, insbesondere GAP, untersucht wurde. Außerdem konnte gezeigt werden, dass die Bindungsaffinität von ATP an OEP21 in Gegenwart von Magnesium, das z. B. im Intermembranraum in erhöhter Konzentration vorhanden ist, deutlich reduziert ist. Diese Daten deuten darauf hin, dass phosphorylierte Metaboliten die gleiche positiv geladene Bindungsoberfläche von OEP21 nutzen und dass die Dissoziation des hochaffinen Binders ATP bei zellulären Konzentrationen erreicht werden kann.

Auf der Grundlage der hier vorgestellten Ergebnisse über den Bindungsmechanismus und das Zusammenspiel verschiedener Metaboliten wurde der Transport von Metaboliten durch OEP21 in weiteren Arbeiten experimentell untersucht, was ein umfassendes Verständnis der Funktionalität von OEP21 ermöglichte. Außerdem bildet OEP21 größere Oligomere in der äußeren Hüllmembran, und es konnte gezeigt werden, dass die Oligomerisierung *in vitro* durch die LDAO-Konzentration gesteuert werden kann. Dabei kann OEP21 als dimere Form auftreten, die durch einen Cysteinrest vermittelt wird, der zur Membran hin orientiert und gut positioniert ist, um unter oxidierenden Bedingungen eine Disulfidbrücke zu bilden. Diese oxidationsabhängige Oligomerisierung von OEP21 als Reaktion auf oxidativen Stress könnte ein kritischer Schalter für Pflanzenzellen bei der Ausführung der hypersensitiven Reaktion sein, einer Art von programmiertem Zelltod zur Abwehr von Pathogeninfektionen in Pflanzen.

Das zweite hier untersuchte Protein ist Tic40, das eine Komponente des Proteinimportapparats der inneren Hüllmembran von Chloroplasten ist. In dieser Arbeit wurden strukturelle Unterschiede zwischen der fruchtbaren BnaA10-Variante und der nicht-fruchtbaren BnaC9-Variante festgestellt. Dabei wurde BnaA10 mittels NMR-Spektroskopie untersucht, wobei die Informationen über die chemische Verschiebung von ^{13}C elf α-Helix-Sekundärstrukturelemente in BnaA10 offenbarten. Außerdem wurde ein Strukturmodell von BnaA10 durch einen kombinierten Ansatz aus Strukturvorhersage und bekannten NMR-Strukturdaten abgeleitet. Dieses Strukturmodell stimmte sehr gut mit den α-Helix-Sekundärstrukturelementen überein, die anhand der Informationen über die chemische Verschiebung von ^{13}C bestimmt wurden. Das Strukturmodell wies jedoch eine längliche Form auf und ließ größere hydrophobe Ansammlungen vermissen, die für ein kompakt gefaltetes Protein zu erwarten wären. Auf der Grundlage des Strukturmodells wurden zwei mutierte Positionen in der Tic40-Variante BnaC9 als möglicher Grund für den Unterschied in der Fruchtbarkeit identifiziert, wobei BnaC9 einen höheren Grad an Proteinflexibilität und eine schwächere Faltung aufzuweisen scheint, was seine Unfruchtbarkeit verursacht.

1. INTRODUCTION

1.1. Biological background

1.1.1. Membrane proteins

Membrane proteins are located at the interface between two compartments, e.g. between extracellular space and cytoplasm or between intermembrane space and mitochondrial matrix, or else they account for most of the mass of small vesicles involved in endocytosis, exocytosis, or intracellular trafficking [3]. The amount of membrane proteins as part of membranes varies depending on the membrane function. In the case of the internal membranes of chloroplasts and mitochondria, which are involved in adenosine triphosphate (ATP) production, membrane proteins can account for up to 75 % of the membrane mass [4]. In contrast, membrane proteins make up only up to 25 % of the mass of the myelin membrane, which mainly serves as electrical insulation for nerve-cell axons [4]. For most of their lives, membrane proteins interact closely with lipids and water in their environment, where, like other proteins, they must be synthesized by the ribosome and then make their way to various membrane locations within a cell [3]. This results in unique and sometimes conflicting requirements for membrane protein folding, translocation, and stability [3]. Membrane proteins have a wide range of important functions, including transport of molecules and ions, recognition of immune system molecules, energy and signal transduction [5], which are critical for normal development and physiology and, when disrupted, lead to a variety of diseases [6]. Due to these diverse biological functions, membrane proteins are of utmost importance for drug design and represent around 70 % of all current drug targets [2, 7]. Therefore, a comprehensive understanding of their structure and structure-function relationships is essential [5]. Although membrane proteins make up about 30 % of the proteome, they are grossly underrepresented in solved atomic structures and account for only about 1 % of the structures in the protein data bank (PDB) [8]. This is due to technical challenges associated with expressing and purifying membrane proteins in a functional form and sufficient quantity for structural analysis, as well as the difficulties of analyzing membrane proteins in their native-like environment as they are inserted into the lipid membrane [2, 5, 8]. Selecting an appropriate membrane mimetic that supports membrane protein stability and function can be challenging because several membrane mimetics are available [2].

The major difficulty in studying membrane proteins is obtaining the membrane protein of interest, since they are usually present only in small quantities in biological membranes and a single protein species is rarely the major component of the membrane [9]. Therefore, most membrane proteins cannot be readily obtained in sufficient amounts from their native environments, so attempts are made to overexpress them [9]. The main problem in working with heterologous expression systems is the aggregation of the membrane protein in the cytoplasm, so that a high yield of stable and functional protein is rarely obtained [9, 10]. In addition, mammalian proteins often require post-translational modifications that are not possible in bacterial hosts [9].

Nevertheless, in recent years, many 3D structures of miscellaneous membrane proteins have been solved by experimental methods such as nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography, and cryo-electron microscopy [5]. Therefore, knowledge of membrane protein's structure is a rich source of biological information, such as their orientation with respect to the lipid bilayer that often provides clues to their function [3].

1.1.2. Constitution of biological membranes

Biological membranes make life possible by forming compartments that provide separation between the outside and inside of an organism, separating different cell organelles, and controlling which substances enter and leave through their selective permeability [11]. In addition, biological membranes enable organisms to generate energy by allowing ions gradients to form across the membrane, and they control the flow of messages between cells by sending, receiving, and processing information in the form of electrical or signals [11]. Interestingly, the properties of biological membranes are remarkably similar from simple unicellular prokaryotes to complex multicellular eukaryotes, such as humans [11]. Biological membranes consist of a bilayer of lipid molecules, and three types of lipids are found in membranes, namely phospholipids, glycolipids, and sterols [11], as illustrated in FIGURE 1. Due to their composition, all membrane lipids are amphipathic, which results in the hydrophilic heads facing outward towards the aqueous environment and the hydrophobic tails facing inward toward each other [11]. In addition to lipids, membrane proteins and sugars are also important components of biological membranes [11].

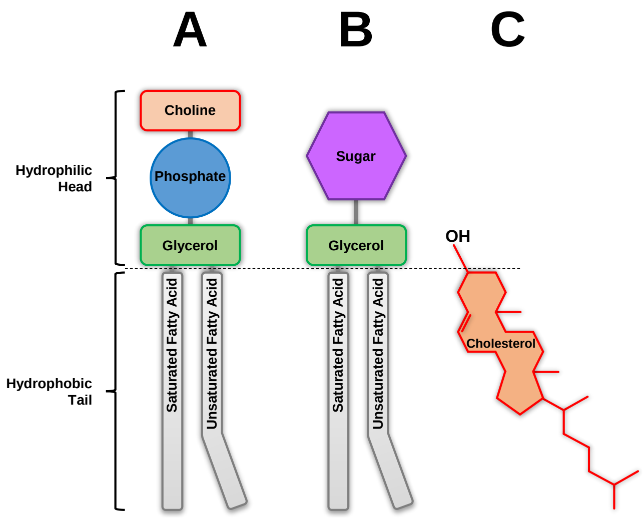


FIGURE 1: Types of membrane lipids

The three types of membrane lipids are shown schematically, namely the glycerophospholipid phosphatidylcholine (A), a glycolipid (B), and the sterol cholesterol (C). This figure is based on WATSON (2015) [11].

Phospholipids that contain two fatty acid chains linked to glycerol and a phosphate group are called glycerophospholipids [11]. An example of a glycerophospholipid is phosphatidylcholine, which has a choline molecule attached to the phosphate group and is commonly found in biological membranes [11]. Besides choline, serine or ethanolamine can also be found at this site, with the corresponding lipids being referred to as phosphatidylserine and phosphatidylethanolamine, respectively [11]. In addition, phospholipids can be based on sphingosine and are therefore called sphingophospholipids, such as sphingomyelin [11]. Unlike phospholipids, glycolipids always have a sugar, such as glucose, in place of the phosphate head and may contain either glycerol or sphingosine [11]. Sterols are absent from most bacterial membranes but are important components of plant and animal membranes, in which stigmasterol and cholesterol, respectively, may typically be present [11]. Unlike phospholipids or glycolipids, cholesterol has a very different structure consisting of a hydroxyl group, which is the hydrophilic head region, a four-ring steroid structure, and a short hydrocarbon side chain [11].

According to the fluid mosaic model, biological membranes are dynamic because proteins and lipids can diffuse laterally through the membrane [11, 12]. Phospholipids can diffuse relatively rapidly in their lipid bilayer, whereas membrane proteins generally move more slowly laterally in the bilayer [11]. In addition, phospholipids can rotate around their head-to-tail axis and their lipid tails are very flexible, resulting in a dynamic and fluid membrane [11]. In contrast, the vertical movement – also known as flipflop – of proteins and lipids from one leaflet to another is extremely slow due to energetic barriers [11]. Therefore, the inner and outer leaflets of the bilayer have different lipid compositions, and the membrane proteins can be inserted into the bilayer in the correct orientation to function [11].

1.1.3. Membrane mimetics

To create a native environment for membrane proteins and thus ensure their stability and function, various membrane mimetics are available [2]. Detergent micelles are the most used membrane mimetics for the preparation of membrane protein and subsequent structural studies [2, 13]. Detergents are generally amphipathic molecules consisting of a hydrophobic tail and a polar head group that spontaneously form typically spherical micelle structures in aqueous solutions [9]. Detergents can be classified into four main categories according to their structure, namely ionic detergents, bile acid salts, nonionic and zwitterionic detergents [9]. Ionic detergents contain a head group with a net cationic or anionic charge and are extremely effective in solubilizing membrane proteins, but have a denaturing effect to some extent [9]. Bile acid salts differ from ionic detergents in that their backbone consists of rigid steroidal groups, which results in bile acid salts having a polar and an apolar face, rather than a well-defined head group [9]. Therefore, bile acid salts form small kidney-shaped aggregates instead of spherical micelles and are relatively mild detergents [9]. Nonionic detergents contain uncharged hydrophilic head groups constating of either polyoxyethylene or glycosidic groups and are generally considered mild and relatively non-denaturing [9]. Zwitterionic detergents combine the properties of ionic and nonionic detergents and are generally more deactivating than nonionic detergents [9]. Lauryldimethylamine oxide (LDAO) and dodecylphosphocholine (DPC) are examples for zwitterionic detergents. The general advantages of using detergent micelles include ease of use and the availability of many different detergent scaffolds, which provide the opportunity to make an optimal selection for a particular membrane protein of interest [2]. Although detergent micelles are commonly used for solution-state NMR spectroscopy, only a small number have been shown to be suitable for high-resolution work, while they commonly form relatively small micelles that offer the advantage of favorable NMR relaxation properties for inserted membrane proteins [2]. Detergent micelles in many cases lead to a functionally folded membrane protein preparation, especially for membrane proteins that exhibit high thermodynamic stability [2]. In contrast, the use of detergent micelles has some disadvantages, as illustrated in FIGURE 2. They are usually quite harsh and therefore often incompatible with more labile membrane proteins, as the detergent reduces their thermodynamic stability and leads to a misfolded or unfolded membrane protein [2, 9]. In addition, detergent molecules often bind to hydrophobic cavities of membrane proteins that are important for enzymatic activity [2]. Probably the most important problem is that detergent micelles can interfere with the binding of partner proteins, leading to the unfolding or precipitation of soluble protein domains [2, 14, 15]. When studying complexes of two different membrane proteins, detergents optimized individually for each protein may not be compatible with each other, which could lead to structural perturbations or unfolding of one or both membrane proteins in the complex [2].

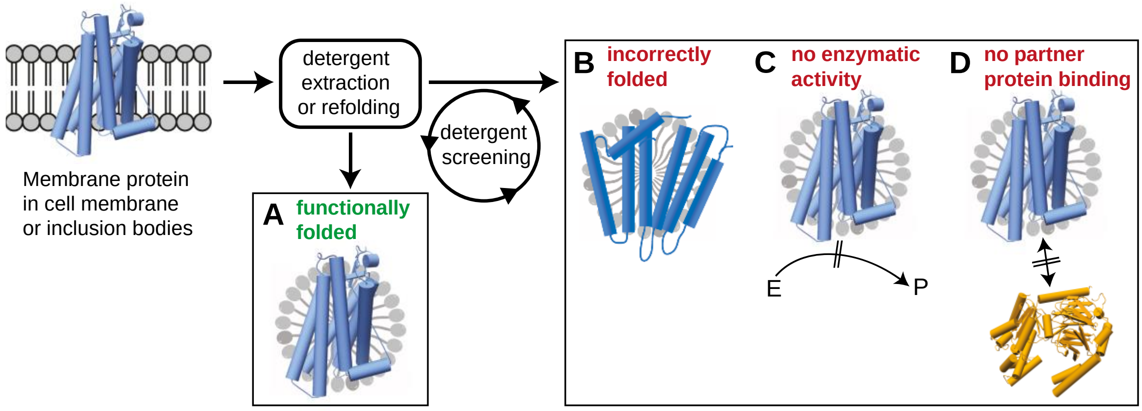


FIGURE 2: Common problems with the use of detergents as membrane mimetic

Extraction of a membrane protein from cells or refolding of inclusion bodies using detergent micelles can lead to different scenarios that may require a lengthy optimization procedure. In the best case, the membrane protein tolerates the initial detergent used for extraction (A). Other, more likely possibilities are that the membrane protein is misfolded in the chosen detergent (B), a putative enzymatic activity is inhibited by detergent molecules (C), or the detergent prevents interaction with or even leads to unfolding of a soluble partner protein (D). Scenarios (C) - (D) would require iterative screening to identify detergents that are compatible with the desired application and yet suitable for NMR spectroscopy. This figure has been adapted with permission from KLÖPFER and HAGN (2019) [2].

In general, the structures of membrane proteins can be significantly modulated by the selected membrane mimetic [2, 16]. This dilemma has been recognized and addressed by the introduction of more native membrane mimetics that have the potential to solve most or all those issues [2, 17]. Moreover, crystal structures of membrane proteins solubilized in detergents often contain bound lipids, highlighting the beneficial effect of a lipid environment on their structure and stability [2, 18]. Due to these limitations of detergent micelles, various membrane mimetics such as lipid-detergent bicelles, amphipols, phospholipid nanodiscs, or liposomes are available to provide a more native-like environment [2]. In FIGURE 3, commonly used membrane mimetics are illustrated.

Bicelles are discoidal lipid aggregates composed of long-chain phospholipids and detergents [2, 19]. They exhibit a more lipid-like behavior than micelles but can still lead to denaturation and destabilization of membrane proteins as well as their partner proteins due to the presence of detergent molecules [2]. The major advantage of bicelles is their size scalability by simply varying the molar ratio between the long-chain and detergent (or short-chain lipid) component [2]. Moreover, the lipid composition in the planar region of bicelles can be tailored to the requirements of the membrane protein under study, as is the case with any phospholipid bilayer system [2].

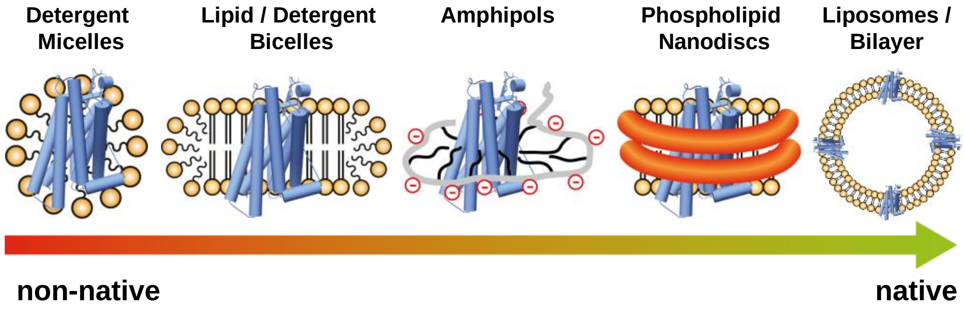


FIGURE 3: Commonly used membrane mimetics

Various membrane mimetics are available, the most commonly used of which are illustrated and ranked according to their ability to resemble a native membrane environment. This figure has been adapted with permission from KLÖPFER and HAGN (2019) [2].

Amphipols are polymer-based amphipathic membrane mimetics forming stable and homogenous particles that are more native than bicelles due to the absence of detergent and appear to be suitable for refolding certain membrane proteins [2, 20, 21]. Unfortunately, only one amphipol is commercially available, which limits the use of this system [2].

The most native system would be pure liposomes or immobilized lipid bilayers for studying membrane proteins, but these are very large and therefore not suitable for solution-state NMR or most other structural methods [2]. This issue has been resolved by the use of lipid-binding proteins such as membrane scaffolding protein (MSP) or Saposin, styrene-maleic acid copolymers, or amphiphilic peptides, which enclose a patch of the lipid bilayer and lead to the formation of nanometer-diameter lipid discs called nanodiscs [2]. These nanodiscs provide native lipid bilayer properties and ensure the associated lateral pressure properties that are often essential for proper function of membrane proteins and enable the binding of peripheral membrane proteins to their partner proteins or a plain membrane surface in a detergent-free lipid environment [2].

The formation of MSP nanodiscs relies on the ability of apolipoprotein A-I to wrap around small lipid bilayer patches, creating small membrane-like particles of defined size [2, 22, 23]. Over time, different versions of apolipoprotein A-I have been engineered for biophysical studies and have been termed membrane scaffolding proteins [2, 24]. Ideally, two copies of MSP surrounding a patch of lipid bilayer in an antiparallel manner form a nanodisc with a well-defined diameter, where the size of the nanodiscs is determined by the length of the MSP [2, 24, 25]. Any kind of bilayer-forming lipid can be used to assemble nanodiscs, depending on the requirements of the membrane protein or of the experiment [2]. In addition, circularized MSPs are available for nanodisc assembly, which exhibit an improved size homogeneity and thermal stability compared to the linear MSP nanodiscs that tend to disassemble with time due to the higher strain in the MSP belt [2, 26, 27].

Saposins are lipid-binding proteins and can be used to form nanodiscs, with Saposin A being the most common member of the family, which tolerates a wide range of lipids and lipid-to-Saposin ratios [2, 28, 29]. Saposin A can be converted from a closed apo confirmation into an open form, in which two Saposin A surrounding the bilayer are in a head-to-tail arrangement, by the addition of detergents [2, 30]. The assembly of nanodiscs with Saposin A, referred to as Salipro nanodiscs, is based on the formation of Saposin A oligomers and thus allows the size of the nanodisc to be easily adjusted by varying the Saposin A-to-lipid ratio [2]. Moreover, such an open system is useful for membrane protein oligomers that would not insert into an MSP nanodisc due to their strictly defined diameter [2].

Polymer nanodiscs are based on the principle that amphipathic polymers can solubilize lipid bilayers and form discoidal particles [31], which led to the development of nanodiscs composed of an amphipathic styrene and maleic-acid co-polymer [2, 32]. As with any chemical polymer, the polydispersity of the chain must be considered, and due to protonation of the maleic acid moiety at low pH, styrene maleic anhydride lipid particles tend to precipitate at pH values below 6.5 [2]. In addition, divalent cations such as Ca^{2+} and Mg^{2+} are chelated by the polymer, rendering the particles insoluble and limiting their use [2]. These problems have been solved by using modified polymers, such as SMA-QA that contains positively charged quaternary ammonium compounds, or a non-styrene-containing styrene maleic anhydride co-polymer called DIBMA [2, 33, 34]. The advantages of styrene maleic anhydride lipid particles are that they provide a very mild environment for membrane proteins, yet can efficiently solubilize lipids, and their ability to be used for the direct extraction of membrane proteins from cellular membranes [2, 35].

Peptide nanodiscs are based on short amphipathic peptides that can bind to the hydrophobic rim of a phospholipid bilayer membrane [2]. The most used amphipathic peptide, named 18A, is an eighteen amino acid peptide derived from apolipoprotein A-I [2, 36-38]. The size of the nanodisc can be easily adjusted by varying the peptide-to-lipid ratio [2]. However, disadvantages include the relative high cost of custom peptide synthesis and the lower stability due to their non-covalent assembly [2]. The latter problem has been solved by chemical methods involving covalent ligation of individual peptides in assembled nanodiscs [2, 39].

Advantages and disadvantages of this commonly used membrane mimetics in general and for their use in solution-state NMR are summarized in TABLE 1.

TABLE 1: Advantages / disadvantages of membrane mimetics

Some advantages and disadvantages of commonly used membrane mimetics in general and for their use in solution-state NMR are summarized as bullet points below, according to KLÖPFER and HAGN (2019) [2].

Membrane mimetic	Advantages	Disadvantages
Detergent micelles	• Large number of detergents available • Small size	• Denaturation of membrane protein or partner protein • Critical micelle concentration needs to be considered
Bicelles	• Tunable size • Easy to prepare	• Denaturing effect of detergent component
Amphipols	• Solubilization of membrane proteins • Suitable for refolding of GPCRs	• Sometimes too harsh for maintaining complexes • Only one amphipol type is commercially available
MSP nanodiscs	• Lipid-MSP-only system • Any lipid can be incorporated • Tunable size • Homogenous size • Stable after assembly	• Not suitable for large membrane protein oligomers • Dynamics of lipids affected by protein 'belt' • Optimization of assembly conditions can be time consuming
Polymer nanodiscs	• Stable • Extraction of membrane proteins from membrane is possible without use of detergents • Easy handling • Commercially available	• Size not tunable in an easy manner • Stability issues at low pH and in presence of divalent cations • Less homogenous as compared to MSP nanodiscs
Saposin nanodiscs	• Open system that can adapt to larger membrane proteins and their oligomers • Size tunable by saposin:lipid ratio • Tight packing of saposin with membrane protein (small size)	• Less stable than other nanodisc systems • Production of saposins requires some experience • Low amount of lipids leads to reduced NMR spectral quality
Peptide nanodiscs	• Suitable for membrane protein oligomers • Size tunable by peptide-to-lipid ratio	• Lower stability • Size homogeneity lower

1.1.4. Classification of membrane proteins

Since many different types of membrane proteins have been identified so far and one type of membrane gradually blends into another, they are sometimes difficult to distinguish [40]. However, they can be divided into the two main classes according to the association of the membrane protein with the membrane, namely integral and peripheral membrane proteins [41], as illustrated in FIGURE 4.

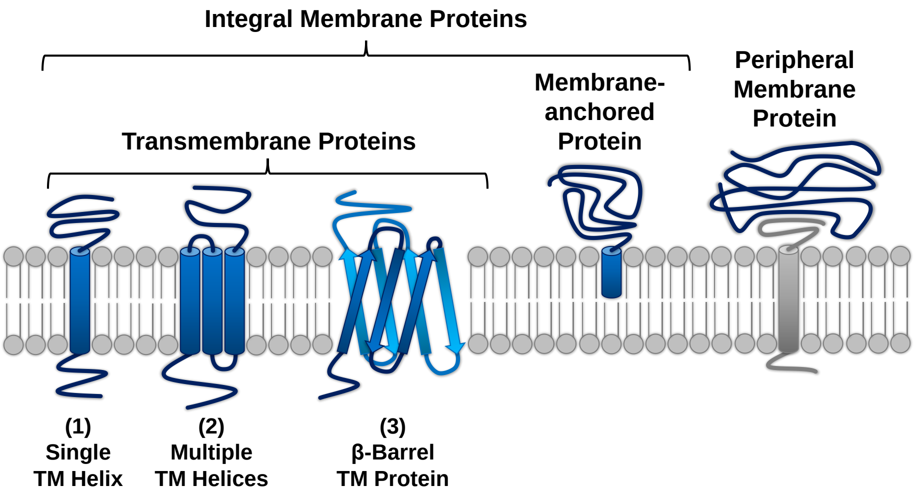


Figure 4: Classification of membrane proteins

Membrane proteins can be divided into peripheral and integral membrane proteins. For integral membrane proteins, a further distinction can be made between membrane-anchored or transmembrane proteins. Transmembrane (TM) proteins can generally consist of either a single α-helical transmembrane domain, or multiple α-helical transmembrane domains that traverse the membrane multiple times, or β-sheet structures that span the membrane in multiple passes and forming a so-called β-barrel.

Peripheral membrane proteins are associated with cell membranes, but they are generally not integrated into the hydrophobic region of the lipid bilayer interior and their entire mass is located outside the hydrophobic interior of the membrane [41]. Therefore, peripheral membrane proteins can be easily removed, e.g., by washing the membrane, changing the pH, or changing the ionic strength [41]. In contrast, integral membrane proteins can generally only be removed from the membrane by detergents or chaotropic agents that disrupt the membrane structure or the water structure, respectively, because a part of the integral membrane protein is integrated into the lipid bilayer structure [41]. Integral membrane proteins can be further divided into two subclasses, namely transmembrane proteins and anchored membrane proteins [41]. Transmembrane proteins span the lipid bilayer and a part of the protein appears on both sides of the membrane [41]. In contrast, anchored membrane proteins expose a part of their mass only on one side of the membrane and bury a part of their mass in the hydrophobic interior of the lipid bilayer, which serves as an anchor to hold the protein in the membrane [41]. As illustrated in Figure 4, transmembrane proteins are divided here into three major subclasses, whereas various definitions of the category of membrane proteins and their types exist in the literature [42]. This first class consists of generally α-helical single-pass transmembrane proteins, sometimes distinguished by whether their C-terminus or N-terminus is located in the cytosol and on which side of the membrane their soluble domain is located [42]. The second class consists of mostly α-helical transmembrane proteins that cross the lipid bilayer multiple times and are referred to as multi-pass transmembrane proteins [42].

These α-helical transmembrane segments consist of twenty to thirty amino acids spanning a lipid bilayer and exhibit a high degree of hydrophobicity, often making them easy to be identified in hydropathy plots [4]. In contrast, the third class consists of mostly β-sheet structures that span the membrane in multiple passes and form a so-called β-barrel [4, 41]. Because some β-barrel proteins are pore-forming and have a hydrophilic interior, hydropathy plots cannot be used to identify the membrane-spanning segments of a β-barrel transmembrane protein, in which ten amino acids or fewer are sufficient to traverse a lipid bilayer as an extended β-strand and only every other amino acid side chain is hydrophobic [4].

The focus of this work is on membrane proteins, and four different membrane proteins, namely outer envelope protein 21 (OEP21), translocon at the inner envelope of chloroplasts, subunit40 (Tic40), a disintegrin and metalloprotease 17 (ADAM17), and inactive rhomboid protein 2 (iRhom2), were studied. OEP21 is an example for a β-barrel porin transmembrane protein located in the chloroplast outer envelope membrane [43]. In contrast, Tic40 is a membrane-anchored protein located at the inner envelope membrane [44]. Other examples of membrane proteins include ADAM17, which contains an α-helical single-pass transmembrane domain [45] and seems to exist in most vertebrates and invertebrates [45, 46], or iRhom2, which consists of seven α-helical transmembrane segments [45] and is found in addition to iRhom1 in vertebrates [45, 47]. This work will provide detailed structural and functional insights on OEP21 and structural insights on Tic40, where for ADAM17 and iRhom2 protein production and purification is shown. In the following chapters, information on the biological background and the state of research on these proteins is provided.

1.2. The chloroplast proteins OEP21 and Tic40

1.2.1. Chloroplast proteins in general

Mitochondria and chloroplasts originated from two independent single endosymbiotic events, respectively [48]. About two billion years ago, mitochondria originated from an α-proteobacterium that was incorporated into an unknown host cell [48]. Later, about 1.6–1.5 billion years ago, a heterotrophic cell that already contained mitochondria incorporated a cyanobacteria-related organism, which was the ancestor of chloroplasts, resulting in a eukaryotic cell capable of an autotrophic life style [48]. This endosymbiotic process resulted in a massive transfer of genetic information from the endosymbiont to the emerging host nucleus [49]. However, a prerequisite for this process was the development and establishment of a protein-import machinery for the import of chloroplast-localized proteins, which were now synthesized cytosolically [49].

In general, chloroplasts are highly structured and consist of three membrane types that are important for chloroplast biogenesis, namely the outer and inner envelope membrane surrounding the organelle and the thylakoid membrane network containing the photosynthetically active protein complexes [49, 50]. Almost all chloroplast proteins are synthesized in the cytosol as precursor proteins, also known as preproteins, and imported post-translationally into the organelle [49]. Proteins destined for the thylakoid membrane, stroma, and inner envelope are usually synthesized with an amino-terminal extension, called presequence, which is proteolytically removed after import [49]. Preproteins containing a presequence necessary and sufficient for organelle recognition and translocation initiation are recognized in a guanosine triphosphate (GTP)-regulated manner by receptors of the outer envelope translocon, the translocon at the outer envelope of chloroplasts (Toc) complex [49, 51, 52]. As a result, the preproteins cross the outer envelope through an aqueous pore and are subsequently transferred to the translocon in the inner envelope, the translocon at the inner envelope of chloroplasts (Tic) complex [49]. The Toc and Tic translocons cooperate during the translocation process, while the completion of import requires energy that likely comes from the ATP-dependent function of molecular chaperones in the stroma [49, 53]. To obtain the mature form of the protein, the presequence is subsequently cleaved by the stromal processing peptidase [49]. According to the unified nomenclature, the individual subunits are named according to their calculated molecular size and their affiliation with the outer or inner envelope membrane of the chloroplast [54, 55]. So far, several Toc and Tic subunits have been identified, such as Toc34, Toc74, Toc159, Tic110 [56] or Tic20, Tic21, Tic22, Tic32, Tic55, Tic62, Tic110, and Tic40 [48]. The role Tic40 has not been fully elucidated, as Tic40 is thought to mediate the import of preproteins through the inner envelope, but protein import into chloroplast appears to occur through either Tic40-independent or Tic40-dependent pathways, depending on the types of preproteins [57].

In addition to the class of Toc proteins, several proteins have been identified that belong to another class of proteins located in the outer envelope and referred to as outer envelope protein (OEP) [58]. OEPs are unique to chloroplasts and have different substrate specificities, with OEPs named according to their molecular size in kilodaltons [58, 59]. Due to the endosymbiotic origin of the plastid, many OEPs have similar structural features to ion channels, so-called porins located at the outer membrane of Gram-negative bacteria [58]. A typical porin in the outer membrane consists of eight to twenty-four transmembrane amphipathic β-strands that form a β-barrel and thus the actual pore [58]. The general porin is water-filled and porins are often organized as trimers, while the porin monomer has an hourglass shape [58]. Since OEPs do not contain a presequence, their targeting towards the chloroplast must depend on some other intrinsic factors of the proteins, which are not yet known [58, 60]. Prominent β-barrel proteins in the outer envelope are Toc75, OEP24, OEP37, OEP40, and OEP21 [58].

1.2.2. OEP21 – outer envelope protein 21

Chloroplasts, chlorophyll-containing plastids, are the primary sites of photosynthesis in plant cells and convert atmospheric CO_2 into carbohydrates, starch, and others that support plant cell energy household and metabolism [43]. During photosynthesis, ATP and nicotinamide adenine dinucleotide phosphate (NADPH) are generated at the thylakoid membrane and in the stroma, which are used to fix CO_2 into the triosephosphates such as glyceraldehyde 3-phosphate (GAP) or 3-phophoglyceric acid (3-PGA) via the Calvin-Benson cycle [61]. These triosephosphates must be transported across the inner and outer envelope membranes into the cytosol for further metabolic processing and energy supply, supplying carbon-backbones for the synthesis of sugars or amino acids, and also providing ATP and nicotinamide adenine dinucleotide (NADH) via the enzymatic conversion of triosephosphates back to 3-PGA [43, 62].

The inner and outer membrane are equipped with a series of ion channels and transporters that facilitate the transport of nutrients, solutes, and metabolites into and out of the chloroplast [43, 63]. Transport across the inner envelope membrane is well understood and mediated by many proteins, e.g., triosephosphate / phosphate translocator (TPT), phosphoenolpyruvate / phosphate translocator (PPT), or nucleotide transport proteins (NTT), which have different substrate selectivity and specificity [43, 64-67]. In contrast, the outer envelope membrane was initially considered to be a permeable sieve that cannot form a barrier to small molecules [68]. This view was questioned by the finding that the outer envelope contains substrate-selective and charge-selective channels [69], suggesting a more specific transport mechanism [70]. Nevertheless, a detailed picture of the transport proteins in the outer envelope remains elusive [59, 71].

However, a growing number of channels for the transport of metabolites and solutes across the outer envelope membrane have been discovered [59, 71], some of which are shown in FIGURE 5. These include, in addition to the Toc75 channel, which is required for pre-protein import into chloroplasts, the amine and amino acid transporter OEP16 [72], OEP21 [43], the cation selective channels OEP23 [73] and OEP24 [74] as well as OEP37 [75], a highly conducting solute channel that mediates the passage of ATP, phosphates, dicarboxylic acids, and charged amino acids. The translocation of triosephosphates across the outer envelope membrane is mainly conducted by OEP24 and OEP21, which forms a voltage-dependent, anion-selective channel that selectively facilitates the translocation of 3-PGA and phosphate besides triosephosphates [43, 59, 71, 76].

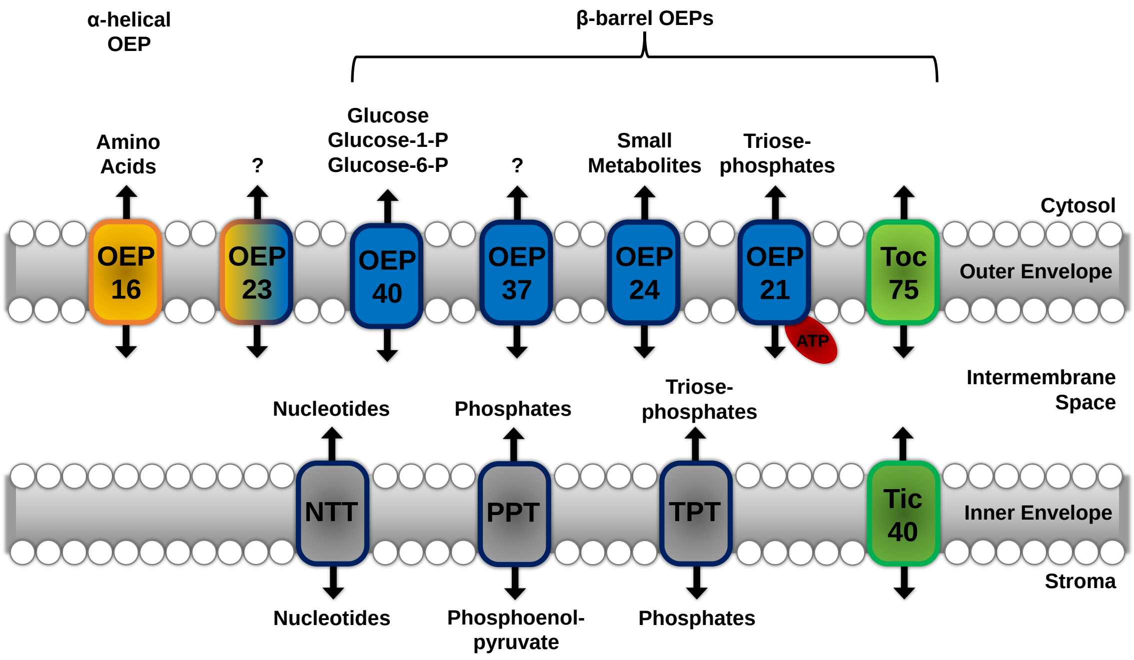


FIGURE 5: Metabolite transport across chloroplast membranes

The outer envelope membrane of chloroplasts contains outer envelope proteins (OEPs), with OEPs named according to their size in kDa. They are responsible for the transport of a variety of metabolites, e.g., amino acids, sugars, or triosephosphates, across the outer envelope membrane. Toc75, OEP40, OEP37, OEP24, and OEP21 form β-barrels, while OEP16 has an α-helical structure and OEP23 consists of a mixture of α-helices and β-sheets. Before metabolites can pass the outer envelope membrane to be released into the cytosol, they must be transported from the stroma through the inner envelope membrane into the intermembrane space. For example, triosephosphates are transported into the intermembrane space by the triosephosphate / phosphate translocator (TPT) in exchange for phosphates. Phosphates itself are transported into the intermembrane space by the phosphoenolpyruvate / phosphate translocator (PPT) in exchange for phosphoenolpyruvate. In addition, nucleotides are translocated by nucleotide transport proteins (NTT). Tic40 and Toc75 are part of the Tic and Toc complexes that facilitate protein transport through the inner and outer envelope membrane, respectively. This figure is based on FACCHINELLI and WEBER (2011) [77] and BARTH *et al.* (2022) [58].

Interestingly, the abundance of individual OEPs can vary widely in different plant species to account for the different levels of metabolite flux. For instance, the larger and less selective OEP24 pore is abundant in C4 plants, where a higher rate of carbon fixation and metabolite flux occurs, while playing a minor role in C3 plants [76]. This functional redundancy may be the reason why OEP37 knock-outs did not led to a phenotype in *Arabidopsis thaliana* [75]. In contrast, the outward-rectifying OEP21 channel is more abundant in C3 plants and has been shown to interact with ATP and triosephosphates or phosphate in a competitive manner [43, 78]. However, the molecular and structural features of metabolite transport and selectivity remain elusive for OEP21, as well as the entire OEP transporter family.

1.2.3. Tic40 – translocon at the inner envelope of chloroplasts, subunit 40

Tic40 consists of an N-terminal transmembrane domain that anchors the protein into the inner envelope membrane and a large hydrophilic C-terminal domain facing the stromal side [48, 54], as shown in FIGURE 6.

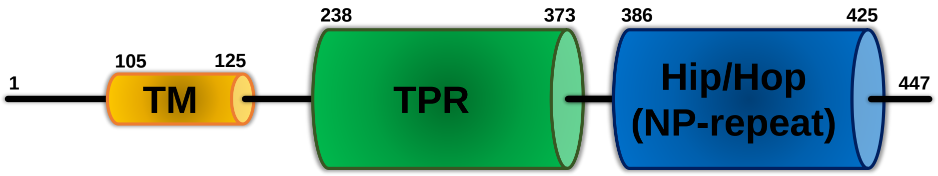


FIGURE 6: Overview of the structural features of Tic40

Tic40 consists of an N-terminal transmembrane domain (TM), a TPR-like domain, and a Hip/Hop-like domain, sometimes referred to as an NP-repeat domain. The indicated amino acids positions of the beginning and end of the structural features correspond to those of *Arabidopsis thaliana*.

The C-terminal part has a single tetratricopeptide repeat (TPR)-like domain and a domain with similarities to Hsp70 interacting protein (Hip) / Hsp70-Hsp90 organizing protein (Hop) proteins, sometimes referred to as an NP-repeat domain [44, 48]. TPR domains typically have multiple copies of two antiparallel-helices with a total length of thirty-four amino acids and are known to mediate protein-protein interactions [48, 79]. The TPR domain of Tic40 appears to be involved in the binding to Tic110, which is favored in the presence of precursor proteins, as revealed by cross-linking experiments, yeast two hybrid and bimolecular fluorescence complementation assays, suggesting proximity and interaction between Tic110 and Tic40 [44, 48, 54, 80, 81]. Hip binds to the adenosine triphosphatase (ATPase) domain of heat shock cognate protein 70 (Hsc70) and stabilizes Hsc70 in its adenosine diphosphate (ADP) state, while Hop binds to the C-termini of Hsp70 and Hsp90 and coordinates the functions of these two chaperones [44, 82]. The Hip/Hop domain of Tic40 mediates the interaction with Hsp93 and *in vitro* experiments indicated that Tic40 stimulates the ATPase activity of Hsp93 [44, 48, 81].

Compared to wild type plants, null mutants of *Arabidopsis thaliana tic40* showed a pale green phenotype, slower growth, and fewer grana stacks in the thylakoids [44, 48]. The decreased yield of nuclear-encoded proteins in the chloroplasts was caused by a reduced import rate into the chloroplasts from mutant plants, although the binding of precursor proteins according to the functional Toc complex was not affected [48]. From *in vitro* import experiments with *Arabidopsis thaliana tic40* null mutant chloroplasts, it can be concluded that Tic40 is involved in the re-insertion of Tic21, Tic40, Tic110, and itself into the envelope, as observed by increased amounts of soluble intermediates of Tic21, Tic40, and Tic110 in the stroma [48, 83].

As shown by different Tic40 constructs lacking either the transmembrane domain, TPR-like domain, or Hip/Hop-like domain that failed to complement the *tic40* phenotype, the full-length protein is required for proper activity [48, 80]. The role of Tic40 as a co-chaperone was supported by the use of a construct in which the Hip/Hop-like domain was replaced with the corresponding domain of human Hip, complementing the phenotype of *tic40* knockouts [48]. Moreover, the pattern of interaction with precursor and mature proteins during import experiments is very similar for Tic40, Tic110, and Hsp93 [44, 48]. All this data strongly suggests a model of the import channel Tic110 associated with the motor complex consisting of the co-chaperone Tic40 and the chaperone Hsp93, which facilitates protein import into chloroplasts [48].

Recently, the structure of the NP-repeat domain of Tic40 was determined by NMR spectroscopy, which revealed that NP-repeat domain consists mainly of five α-helices stabilized by two hydrophobic cores [50], as shown in FIGURE 7.

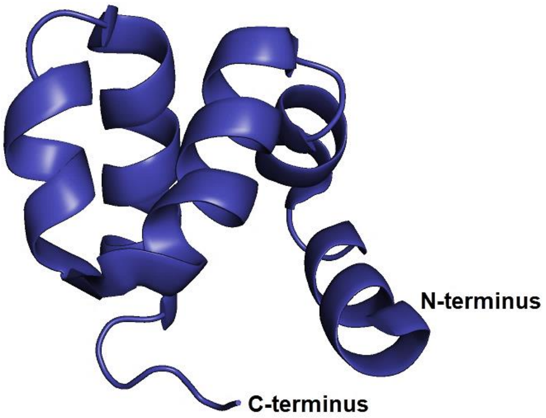


FIGURE 7: Structure of the NP-repeat domain of Tic40

The structure of the NP-repeat domain of Tic40 from *Arabidopsis thaliana*, comprising amino acids P396 to G457 (see TABLE 31 in the appendix for the sequence), was determined by NMR spectroscopy by KAO *et al.* (2012) [50]. The NP-repeat domain of Tic40 consists mainly of five α-helices stabilized by two hydrophobic cores. The NMR chemical shift information and structural coordinates of the NP-repeat domain of Tic40 can be found at the BMRB [84] under the accession code 18171 and the RCSB PDB [85] under the accession code 2LNM, respectively. This figure was created with PyMOL.

1.3. ADAM17 and iRhom2

1.3.1. ADAM17 – a disintegrin and metalloprotease 17

Proteolytic release of transmembrane proteins, known as (ectodomain) shedding, is a key mechanism in several biological processes, including cell-to-cell communication and immunity [86, 87]. Metalloproteinases such as ADAM17 are involved in this irreversible posttranslational modification [86]. Originally, ADAM17 was discovered in search for the factor that releases the proinflammatory cytokine tumor necrosis factor α (TNFα) from the cell membrane, which is why ADAM17 is also known as TNFα converting enzyme (TACE) [45, 88]. Subsequent, ADAM17 was shown to be involved in many shedding events that control the release of several members of the epidermal growth factor (EGF) family, critical adhesion molecules, cytokine receptors, and proinflammatory mediators [45]. Although the structure and function of ADAM17 are largely known except for the structure of the transmembrane domain, it remains difficult to draw a complete picture of its regulation [45]. Since ADAM17 is directly participating in the release of signaling molecules, it may be positively and negatively involved in various physiological processes as well as inflammatory, fibrotic, and malignant pathologies [45]. This participation in different processes requires tight regulation at multiple levels, including phosphorylation, various conformational changes, and endogenous inhibitors [45].

ADAM17 belongs to the ADAM family of single-pass transmembrane multi-domain metalloproteinases [45], which are mainly involved in ectodomain shedding of cell membrane proteins [86]. In mammals, around thirty ADAMs have been identified, only half of which possess the characteristic metalloproteinase domain and proteolytic potential [86, 89]. The characteristic catalytic domain of these proteolytically active ADAMs is highly conserved in all members of the metzincin superfamily [86], with metzincins, including ADAM17, use a Zn^{2+} ion for catalysis that is coordinated to three histidines of the conserved binding motif HExxHxxGxxH [86, 90]. ADAM17 is synthesized as an inactive precursor and contains an N-terminal pro-domain that restricts enzyme activity through a cysteine switch mechanism common to most metzincins [86, 90]. Therefore, a pivotal cysteine contained in the conserved PRCGxPD motif coordinates the catalytic zinc and prevents it from coordinating water molecules and carrying out catalysis [86]. Removal of the pro-domain and activation of ADAM17 is carried out by furin and occurs intracellularly within the secretory pathway [86, 91]. In addition, ADAM17 has three auxiliary domains whose physiological functions are still largely unknown [86]. These are a short stalk domain called conserved ADAM17 dynamic interaction sequence (CANDIS), a membrane proximal domain (MPD) that regulates conformational changes of the enzyme, and a disintegrin domain involved in molecular interactions with other transmembrane proteins including integrins [86].

Near the stalk region are a transmembrane domain mainly involved in the interaction of ADAM17 with its essential regulators iRhom1 and iRhom2, and an intracellular cytoplasmic domain whose physiological function is still unclear [86, 92]. A schematic representation of ADMA17 is illustrated in FIGURE 8.

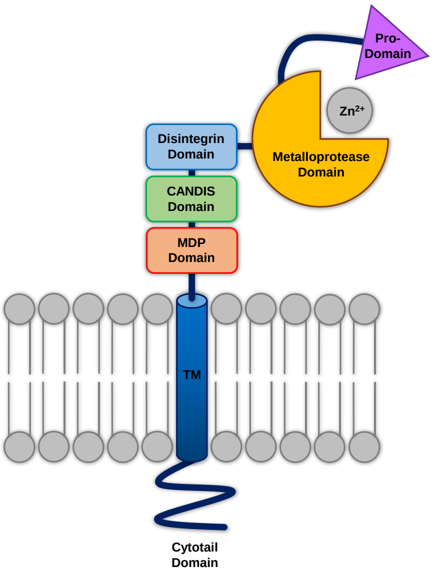


FIGURE 8: Schematic representation of ADAM17

The six domains of ADAM17, i.e., a transmembrane (TM) domain with a cytotail domain, a MPD domain, a CANDIS domain, a disintegrin domain as well as pro-domain that restricts proteolytic activity elicited by the metalloprotease domain, are illustrated with different shapes and colors. This figure is based on CALLIGARIS *et al.* (2021) [86].

To date, more than ninety different substrates of ADAM17 have been described, but it is still not clear whether each substrate identified *in vitro* has a physiological relevant role *in vivo* [45]. Besides TNFα, transforming growth factor alpha [93], amphiregulin [94, 95], interleukin 6 receptor [96], and L-selectin [93] have also been shown to play important physiological roles [45]. ADAM17 is a crucial hub for many signaling pathways and a central component of the so-called three-membrane passing signal-transactivation, in which a signal from one pathway is converted into a signal from another signaling pathway [45]. For example, parts of the large variety of different G-protein-coupled receptor signals can lead to ADAM17 activation, which then leads to the shedding of ligands that bind in an autocrine or paracrine fashion to their receptors, thereby initiating, e.g., EGF-receptor, TNFα-receptor, or interleukin 6 receptor signaling [45]. Since iRhoms critically affect the physiological and pathophysiological signaling processes regulated by ADAM17, there is interest not only in ADAM17 but also in iRhoms as therapeutic targets [45]. Therefore, the regulation of ADAM17 activity, and in particular its interaction with iRhoms, needs to be well understood in order to exploit its therapeutic potential [45].

1.3.2. iRhom2 – inactive rhomboid protein 2

iRhoms have been linked to the development of several human diseases, including cancers [97]. Moreover, the interaction of iRhoms with ADAM17 offers new therapeutic opportunities for selective and simultaneous inactivation of the key signaling pathways closely associated with disease development [97, 98].

Rhomboids are an evolutionarily conserved family of multi-span transmembrane proteins [99, 100], some of which are catalytically active serine proteases that are capable of intramembrane cleavage of their substrates [99]. They were originally discovered in *Drosophila melanogaster* as intramembrane proteases of EGF-receptor ligands [99, 101]. Although other members of the rhomboid family are structurally similar, they are catalytically inactive and known as pseudoproteases [99]. Interestingly, these rhomboid pseudoproteases are evolutionarily conserved, suggesting the presence of selection pressure despite the lack of proteolytic activity, and implying functions important enough to be conserved [47, 99]. iRhom1 and iRhom2, encoded by the genes *Rhbdf1* and *Rhbdf2*, respectively, belong to the family of pseudoproteases and have been named iRhom to indicate their proteolytic inactivity and their affiliation with the rhomboid family [99, 102].

iRhoms have been identified as important interactors for ADAM17 and promote its trafficking, maturation, and activity [45, 103, 104]. Moreover, the absence of iRhoms in cells leads to impaired exit of the ADAM17 proform from the endoplasmic reticulum (ER), preventing maturation of ADAM17 in the Golgi and subsequently preventing mature ADAM17 from reaching the cell surface, where the shedding process would occur [45, 103, 104]. Detailed insight into the molecular mechanism of the iRhom-mediated forward trafficking of ADAM17 is still incomplete, although the role of iRhoms is not limited to ADAM17 forward trafficking, but they are also critically involved in the direct regulation of its shedding activity at the cell surface [45].

iRhoms possess a characteristic core of six transmembrane helices, referred to as rhomboid fold, which is extended by an additional transmembrane helix at the C-terminus, resulting in a so-called six-plus-one topology [45, 47], as illustrated for iRhom2 in FIGURE 9. The N-terminal cytosolic region of iRhoms is approximately four hundred amino acid residues long and is involved in regulating the ADAM17-mediated shedding process [45, 105, 106]. Unique to iRhoms is the iRhom homology domain (IRHD) between the transmembrane helix one and two with a length of about two hundred and fifty amino acids, while in other rhomboids only a short loop (L1) of about thirty amino acids is located at this position [45]. Interestingly, neither the structure nor the function of the IRHD are known, but recently it was shown that the absence of this domain negatively impacts the binding and function of iRhoms to ADAM17 [45, 106].

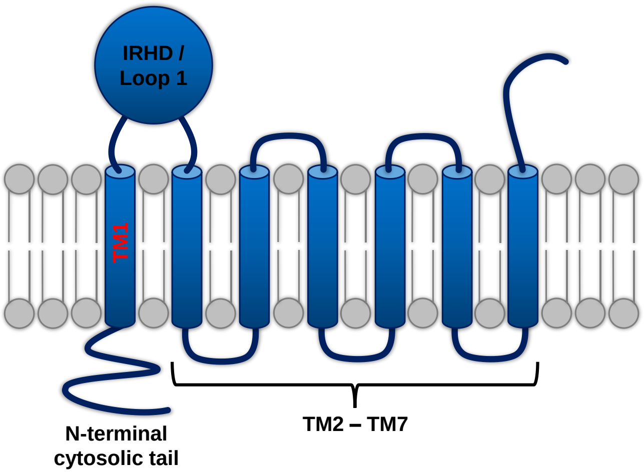


FIGURE 9: Schematic representation of iRhom2

iRhom2 has the characteristic rhomboid fold consisting of seven transmembrane (TM) helices in a so-called six-plus-one topology. Unique to iRhoms is the iRhom homology domain (IRHD) between the first and second transmembrane helix, while in other rhomboids only a short loop is located at this position. Another prominent feature is the long cytosolic N-terminal tail, which accounts for almost half of the protein. This figure is based on DÜSTERHÖFT *et al.* (2019) [45].

However, it is not yet clear whether deletion of the IRHD destabilizes the overall folding of iRhoms or whether the IRHD is directly involved in the regulation of ADAM17 [45]. The IRHD accounts for approximately one quarter of the total protein, suggesting that this domain is of vital interest and requires further investigation [45]. Therefore, it can be speculated whether the IRHD is involved in the interaction with ADAM17 or even regulates the substrate recognition of ADAM17 [45]. In comparison, loop L1 in active rhomboids appears to play an important role in protease activity, whereas deletion of the loop abrogates protease activity, probably by affecting the structural integrity of the rhomboids [45, 107, 108]. Additionally, recent studies suggest that loop L1 is involved in rhomboid dimerization as well as substrate recognition [45, 109-111].

1.3.3. Get3 – guided entry of tail-anchored proteins factor 3

In general, tail-anchored proteins represent a large and heterogeneous group of transmembrane polypeptides [112] and feature an N-terminal cytosolic domain anchored to the intracellular membrane by a single C-terminal transmembrane domain [113]. In addition, the transmembrane domain is followed by a short translocated polar region [112]. Tail-anchored proteins mediate many important biochemical activities in almost every cell membrane, with nearly 5 % of the membrane proteins being tail-anchored [114]. Since the transmembrane domain is in close proximity to the C-terminus at no more than thirty residues, a special type of biogenesis of tail-anchored proteins is enabled [112].

Therefore, the hydrophobic tail exits the ribosome only after translation is terminated, whereas translocation of the C-terminal polar domain through all target membranes, including the ER, must occur post-translationally [112]. This post-translational insertion is specific to tail-anchored proteins and distinguishes them from other ER-targeted transmembrane proteins that are inserted by a co-translational process [112]. A highly conserved 40-kDa ATPase in mammals, designated transmembrane domain recognition complex 40 (TRC40) or arsenical pump-driving ATPase protein 1 (Asna1), has been identified as a central component of the pathway of tail-anchored protein to the ER in eukaryotes, which recognizes and selectively binds the transmembrane domain of tail-anchored proteins in the cytosol [114-116]. Guided entry of tail-anchored proteins factor 3 (Get3) is the homologous system of TRC40 in *Saccharomyces cerevisiae* [114]. After recognition and binding of the transmembrane domain by Get3, the soluble complex targets to the ER by membrane-bound receptors, termed Get1 and Get2 in yeast, where the tail-anchored protein is released for insertion into the membrane [114]. Recent studies revealed a more complex mechanism involving Get4 and Get5 and a co-chaperone called small, glutamine-rich, tetratricopeptide repeat protein 2 (Sgt2), while the Get3 homodimer interacts with two copies of the Get4-Get5 complex to form an extended conformation in solution [117]. As revealed by structural studies, the fold of Get3 includes a helical subdomain that is structurally and functionally coupled to a core ATPase domain [114, 118]. In addition, Get3 occurs as a homodimer in which the two subunits are linked by a tightly coordinated zinc ion [114, 118]. A crystal structure of dimeric Get3 bound to the transmembrane domain of Pep12 is shown in FIGURE 10.

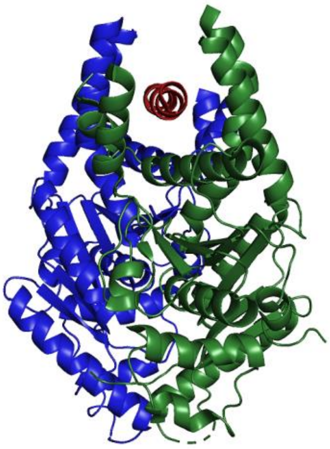


FIGURE 10: Structure of Get3 bound to the transmembrane domain of Pep12

Crystal structure of dimeric Get3 from *Saccharomyces cerevisiae* bound to a truncated Pep12 tail-anchored protein obtained by X-ray diffraction measurements by MATEJA *et al.* (2015) [119]. The helical transmembrane domain of Pep12 substrate (red) binds deeply within the composite hydrophobic groove of dimeric Get3 (blue and green). The structural coordinates of this figure, created with PyMOL, can be found at the RCSB PDB [85] under the accession code 4XTR.

As crystallographic, computational, and biophysical studies have shown, Get3 can exhibit different conformations ranging from closed to a loose to an open [118], as shown in Figure 11.

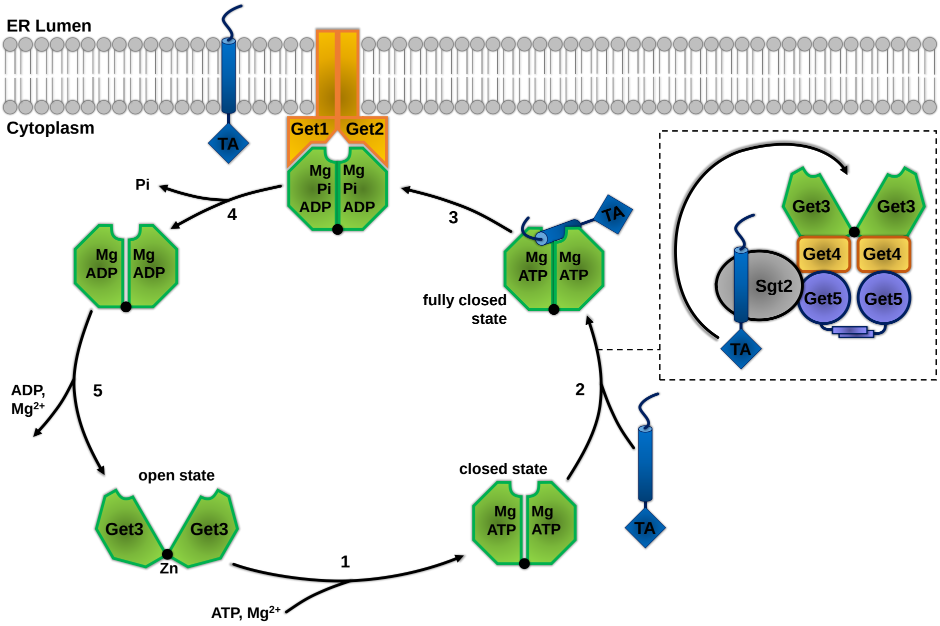


Figure 11: Model of tail-anchored protein binding and insertion by Get3

In the cytosol, the nucleotide-free Get3 dimer (green) is in an open state, and ATP binding in the presence of magnesium to Get3 causes the dimer to close, where the Get3 dimers are linked by a tightly coordinated zinc ion (1). This allows tail-anchored proteins (blue) to bind to Get3 and induce the fully closed hydrolysis-competent state (2). Docking of the Get1-Get2 receptor (orange) occurs before or after ATP hydrolysis, leaving ADP and inorganic phosphate trapped in the closed active site (3). After release of the tail-anchored protein, inorganic phosphate is released and Get3 dissociates from the Get1-Get2 receptor (4), before the Get3 dimer adopts an open structure upon release of magnesium (5). Recent studies revealed a more complex mechanism involving Get4 and Get5 and the co-chaperone Sgt2, while the Get3 homodimer interacts with two copies of the Get4-Get5 complex to form an extended conformation in solution, as shown in the dotted box. This figure is based on Bozkurt *et al.* (2009) [120] and Chang *et al.* (2012) [117].

The transition from open to close in Get3 occurs around the hinge point centered zinc ion, which is coordinated by the side chains of two cysteine residues of each monomer [114]. The closed conformation of Get3 shows an extended dimer interface that spans the entire length of each monomer and harboring the transmembrane domain binding site [118]. In the intermediate loose conformation, the dimer interface is progressively disrupted, starting with the helical subdomains, while the two Get3 subunits are spread apart in the open conformation [118]. These conformations are regulated by interactions with various ligands such as nucleotides, the Get4-Get5 scaffolding complex, tail-anchored protein clients, and the ER-localized insertion machinery Get1-Get2 to coordinate cycles of binding and release of tail-anchored proteins [118].

Protein production of the transmembrane domains of ADAM17 and iRhom2 should be enabled by co-expression of Get3. Interestingly, Get3 contains an enormous hydrophobic groove with dimensions well suited for binding to an α-helical transmembrane domain of about twenty residues [114].

1.4. Nuclear Magnetic Resonance Spectroscopy

1.4.1. Basic principles of NMR spectroscopy

NMR spectroscopy is a powerful, non-destructive technique that can be used to study the structure, dynamics, and chemical kinetics of a variety of biochemical systems [121]. Like other forms of spectroscopy, NMR spectroscopy relies on the transition from a ground state to an excited state, in particular, the absorbed photon promotes a nuclear spin from its ground state to its excited state [121]. In NMR spectroscopy, the energies of these states arise from the interaction of a nuclear magnetic dipole moment with a strong external magnetic field, and the transition excitation between these states is stimulated by radio-frequency electromagnetic radiation [121]. The nuclear magnetic dipole moment results from the spin angular momentum of the nucleus [121], but not all nuclei provide NMR signals [122]. Nuclei with an even number of charge and mass have a spin quantum number of zero, such as ^{12}C, and since these types of nuclei have no nuclear angular momentum, they will not give raise to NMR signal [122]. All nuclei with an odd mass number, such as ^{1}H, ^{13}C, or ^{15}N, have spin angular momentum due to the presence of an unpaired proton, and all nuclei with an even mass number as well as an odd charge, such as ^{2}H or ^{14}N, also have spin angular momentum [121]. Nuclei with a nonzero spin quantum number rotate about the magnetic field B_0 because the torque is generated by the interaction of the nuclear angular momentum with the magnetic field [122]. The magnetic momentum μ is either parallel or antiparallel to their nuclear angular momentum P and depends on the nuclear spin quantum number I and the nuclear gyromagnetic ratio γ [122], as shown in EQUATION 1.

EQUATION 1:

$$\mu = \gamma \cdot P = \gamma \cdot \hbar \sqrt{I \cdot (I + 1)}$$

in the case of the z-component: $\mu_z = \gamma \cdot P_z = \gamma \cdot \hbar \cdot m$

μ	magnetic moment, $J \cdot T^{-1}$
γ	nuclear gyromagnetic ratio, $rad \cdot s^{-1} \cdot T^{-1}$
P	nuclear angular momentum, $J \cdot s$
$\hbar$	reduced Plank constant (divided by $2 \cdot \pi$), $1.054 \cdot 10^{-34}$ $J \cdot s$
I	nuclear spin quantum number, -
m	magnetic quantum number, -

The nuclear gyromagnetic ratio γ is a specific value for a particular isotope and thus constant for a particular nucleus [122], as shown in TABLE 2 for nuclei commonly used in protein NMR spectroscopy.

TABLE 2: Characteristics of nuclei commonly used in protein NMR spectroscopy

The nuclear spin quantum number I, nuclear gyromagnetic ratio γ, and natural abundance of nuclei commonly used in protein NMR spectroscopy are listed below, with values taken from RULE and HITCHENS (2006) [121].

Nucleus	I	γ	Natural Abundance
^{1}H	½	$267.53 \cdot 10^{6}$ rad·s^{-1}·T^{-1}	99.98 %
^{2}H	1	$41.06 \cdot 10^{6}$ rad·s^{-1}·T^{-1}	0.016 %
^{13}C	½	$67.28 \cdot 10^{6}$ rad·s^{-1}·T^{-1}	1.108 %
^{15}N	½	$-27.12 \cdot 10^{6}$ rad·s^{-1}·T^{-1}	0.37 %

A positive value of γ leads to a parallel and a negative value of γ to an antiparallel orientation of μ to P [122]. The nuclear spin quantum number I can be either an integer or half-integer number, and since we are interested in the z-component of the nuclear angular moment P, P is restricted to integer steps of ħ in the range from -I to +I [121, 122]. Therefore, the projection of nuclear angular momentum onto the z-axis is quantized in space and has a total of 2·I + 1 possible values [122]. This means that a magnetic moment μ of a spin with I = ½ has two orientations, one up along the z-axis (+½) and the other down along the z-axis (-½) with an angle of 54.7° relative to the magnetic field [122], as illustrated in FIGURE 12.

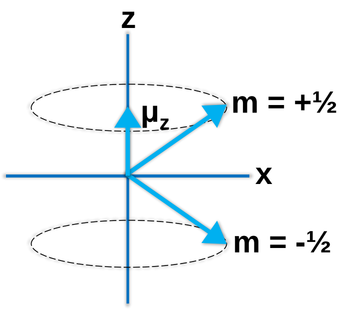


FIGURE 12: Orientation of magnetic moment μ with spin ½

The magnetic momentum μ of a spin with I = ½ has two orientations, one up along the z-axis (+½) and the other down along the z-axis (-½) with an angle of 54.7° relative to the magnetic field. The vectors represent the magnetic moment μ rotating about the magnetic field whose direction is along the z-axis of the laboratory frame. This figure was inspired by TENG (2013) [122].

Each orientation state, also called Zeeman state or spin state, is associated with an energy characterized by the frequency of precession, where the intrinsic frequency of precession is the Larmor frequency ω_0 [122]. The energy of the Zeeman state with magnetic quantum number m can be described in terms of the Larmor frequency [122], as shown in Equation 2.

Equation 2: $E = -\mu_z \cdot B_0 = -m \cdot \hbar \cdot \gamma \cdot B_0 = m \cdot \hbar \cdot \omega_0$ with $\omega_0 = -\gamma \cdot B_0$

E	energy of the Zeeman state, J
μ_z	magnetic moment of the z-component, $J \cdot T^{-1}$
B_0	magnetic field strength, T
m	magnetic quantum number, -
$\hbar$	reduced Plank constant (divided by $2 \cdot \pi$), $1.054 \cdot 10^{-34}$ J·s
γ	nuclear gyromagnetic ratio, $rad \cdot s^{-1} \cdot T^{-1}$
ω_0	Larmor frequency, $rad \cdot s^{-1}$

The energy difference in the allowed transition between, for example, the m = -½ and m = +½ Zeeman states is given by $\Delta E = \hbar \cdot \gamma \cdot B_0$, since the selection rule states that only a single-quantum transition, i.e., $\Delta m = \pm 1$, is allowed [122]. Because the Larmor frequency depends linearly on the magnetic field strength, the magnetic field strength is generally described by the proton Larmor frequency at the specific field strength [122]. For example, a proton resonance frequency of 100 MHz corresponds to the field strength of 2.35 T [122]. The frequency can also be represented in hertz with the relation $v = \omega \cdot (2 \cdot \pi)^{-1}$, where the angular frequency ω is given in radian per second [122]. As the magnetic field strength increases, the energy difference between two states becomes larger as well as the frequency associated with the Zeeman transition [122]. The intensity of the NMR signal is given by the population difference of two Zeeman states of the transition, while the ratio of the populations in the energy states is quantitatively described by the Boltzmann equation [122], as shown in Equation 3.

Equation 3: $$\frac{N_\beta}{N_\alpha} = e^{-\frac{\Delta E}{k \cdot T}} = e^{-\frac{\hbar \cdot \gamma \cdot B_0}{k \cdot T}}$$

at room temperature: $$\frac{N_\beta}{N_\alpha} \approx 1 - \frac{\hbar \cdot \gamma \cdot B_0}{k \cdot T}$$

N_α	population of the α-state, -
N_β	population of the β-state, -
ΔE	energy difference of the Zeeman states, J
k	Boltzmann constant, $1.380 \cdot 10^{-23}$ $J \cdot K^{-1}$
T	temperature, K
$\hbar$	reduced Plank constant (divided by $2 \cdot \pi$), $1.054 \cdot 10^{-34}$ J·s
γ	nuclear gyromagnetic ratio, $rad \cdot s^{-1} \cdot T^{-1}$
B_0	magnetic field strength, T

For a spin with $I = ½$ and a positive γ, such as ^{1}H or ^{13}C, the lower energy state (ground state) is defined as the α-state for $m = ½$, while the higher energy state (excited state) is defined as the β-state for $m = -½$ [122]. In contrast, $m = -½$ for ^{15}N is the lower energy α-state since it has a negative γ [122]. The Boltzmann equation states that both the energy difference of the transition states and the population difference of the states increase with the strength of the magnetic field, and that the population difference is also temperature dependent [122]. At absolute zero, there is no population in the β-state and all spins are in the α-state, while both states are equally populated when the temperature is infinitely high [122]. At room temperature, the term $\hbar\cdot\gamma\cdot B_0$ is much smaller than the term $k\cdot T$, so a first-order Taylor expansion can be used to describe the population difference [122], as shown in Equation 3. Therefore, the population of the β-state is slightly lower than that of the α-state, where, for example, the population ratio for protons at 800 MHz is 0.99987 [122]. This means that due to the low energy difference, only a small fraction of the spins contribute to the signal intensity and that NMR spectroscopy is a very insensitive spectroscopic technique requiring a high magnetic field to achieve sufficient sensitivity [122].

A general feature of NMR spectroscopy is that the observed resonance frequencies are dependent on the local environments of the individual nuclei and differ slightly from the frequencies predicted by $\omega_0 = -\gamma\cdot B_0$ [123]. This differences in resonance frequencies are referred to as chemical shifts and provide the ability to distinguish between otherwise identical nuclei in different chemical environments [123]. The chemical shift phenomenon occurs because the electron motions induced by the external magnetic field generate secondary magnetic fields [123]. The net magnetic field at the location of a given nucleus depends on the static magnetic field and the local secondary fields, where the effect of the secondary fields is called nuclear shielding and can enhance or weaken the effect of the main field [123]. The net effect can be described with the help of a property, the so-called shielding constant σ, by Equation 4 [122].

EQUATION 4: $\nu = 0.5\cdot\pi^{-1}\cdot(1 - \sigma)\cdot\gamma\cdot B_0$

ν	resonance frequency of the nucleus, $Hz = s^{-1}$
σ	shielding constant, -
γ	nuclear gyromagnetic ratio, $rad\cdot s^{-1}\cdot T^{-1}$
B_0	magnetic field strength, T

The shielding constant is always less than one, since the induced local magnetic field cannot be greater than the applied magnetic field [122]. The absolute zero value of a chemical shift is that obtained from a bare nucleus without electrons [122].

Although the absolute zero value of a chemical shift can be obtained for bare nuclei such as protons, it is convenient to use as a reference a particular compound whose resonance frequency is set to the chemical shift value of zero [122]. The chemical shifts of the other resonances are expressed as the difference of the electron shielding to the reference nucleus [122], as shown in EQUATION 5.

EQUATION 5:

$$\delta = \frac{\nu - \nu_{\mathrm{ref}}}{\nu_{\mathrm{ref}}} \cdot 10^{6}$$

δ	chemical shift of the nucleus with frequency v, ppm
ν	resonance frequency of the nucleus, MHz
ν_{ref}	resonance frequency of the reference nucleus, MHz

The chemical shift δ is independent of the magnetic field strength, i.e., the resonances in ppm in a spectrum remain the same when measured at different magnets with different field strengths [122]. The reference compound must meet certain conditions, i.e., it must be stable in a variety of solvents, have an unchanged chemical shift value over a wide range of temperature and pH values, and be easy to handle [122]. Two compounds are commonly used for ^{1}H NMR reference, namely tetramethylsilane (TMS) and 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS), and one of the two reference compounds can be added into an NMR sample as an internal reference or used alone as an external reference [122].

1.4.2. Basics of protein structure determination

To study protein structures and dynamics, spectral resonances must be identified prior structural information and relaxation parameters can be used [122]. Homonuclear nuclear Overhauser effect (NOE) is a widely used technique in various types of calculation methods such as distance geometry, restrained molecular dynamics, or simulated annealing, and variable target function methods in the structural characterization of proteins based on information about approximate distances between protons obtained in nuclear Overhauser effect spectroscopy (NOESY) experiments [122]. When the resonance of a spin in an NMR spectrum is perturbed by saturation or inversion of magnetization, this can lead to a change in the spectral intensities of other resonances in the spectrum, and this phenomenon is known as NOE [122]. This intensity change caused by NOE is due to population changes of the Zeeman states of coupled spins as a result of the perturbation by the dipolar interaction [122]. This distance information is strictly local, i.e., no information is obtained if the spins are farther than about 5 Å apart [121]. The use of a fully deuterated protein allows the measurement of long-range NOEs between amide protons corresponding to distances up to 8 Å [124].

NOE restraints can be derived from a set of 3D-NOESY spectra consisting of, e.g., HNH-NOESY [125], NNH-NOESY [126], CNH-NOESY [127], HCH-NOESY [128] , and CCH-NOESY [127] experiments that allow the assignment not only of proton-proton chemical shift pairs but also of the attached ^{13}C- and ^{15}N-heteronuclei. HNH-NOESY, NNH-NOESY, and CNH-NOESY are based on a [^{15}N,^{1}H]-correlated spectrum with an additional hydrogen, nitrogen, or carbon dimension, respectively. Similarly, HCH-NOESY and CCH-NOESY are based on a [^{13}C,^{1}H]-correlated spectrum with an additional hydrogen or carbon dimension, respectively. This sophisticated approach makes the structure determination process more reliable and accurate, since the probability of incorrect NOE assignments is lower. [122]

Prior using NOE information, the origin of each resonance in the NMR spectrum must be linked to a nucleus in the molecular sequence, a process known as sequence-specific assignment, where the assignment must be completed for a sufficient number of atoms in the sequence to obtain a high-resolution structure [122].

1.4.3. Assignment of backbone and side chain resonances

Numerous triple resonance experiments have been developed and optimized for protein structure determination using heteronuclear multidimensional NMR spectroscopy, where these experiments use one-bond and two-bond heteronuclear scalar couplings to correlate backbone and side chain ^{1}H, ^{15}N, and ^{13}C spins of isotopically labeled protein [122]. The nomenclature that is used for triple resonance experiments relies on the coherence transfer pathway in the experiment [122]. The name of an experiment is formed by the spins involved in the coherence transfer in the order of the transfer pathway, with the spins in parentheses if their chemical shifts do not evolve [122]. In addition, the name is formed only by the first half of the coherence transfer when the magnetization of the proton spin is transferred to neighboring spins and then back to the proton along the same pathway, which is why this type of experiments is referred to as out-and-back [122].

Although a variety of triple resonance experiments are available, often only a certain number of experiments are used for backbone and side chain assignments [122, 129], for example, HNCO [130-132], HN(CA)CO [133-135], HNCA [130, 131, 136-138], HN(CO)CA [131, 136, 137, 139, 140], HNCACB [141], and HN(CO)CACB [142] are used for backbone assignment. Deviating from the out-and-back notation, the HN(CO)CACB experiment is also referred to as CBCA(CO)NH and the HNCACB experiment is referred to as CBCANH. These six NMR experiments are illustrated in FIGURE 13, where observed atoms and atoms through which magnetization flows are highlighted.

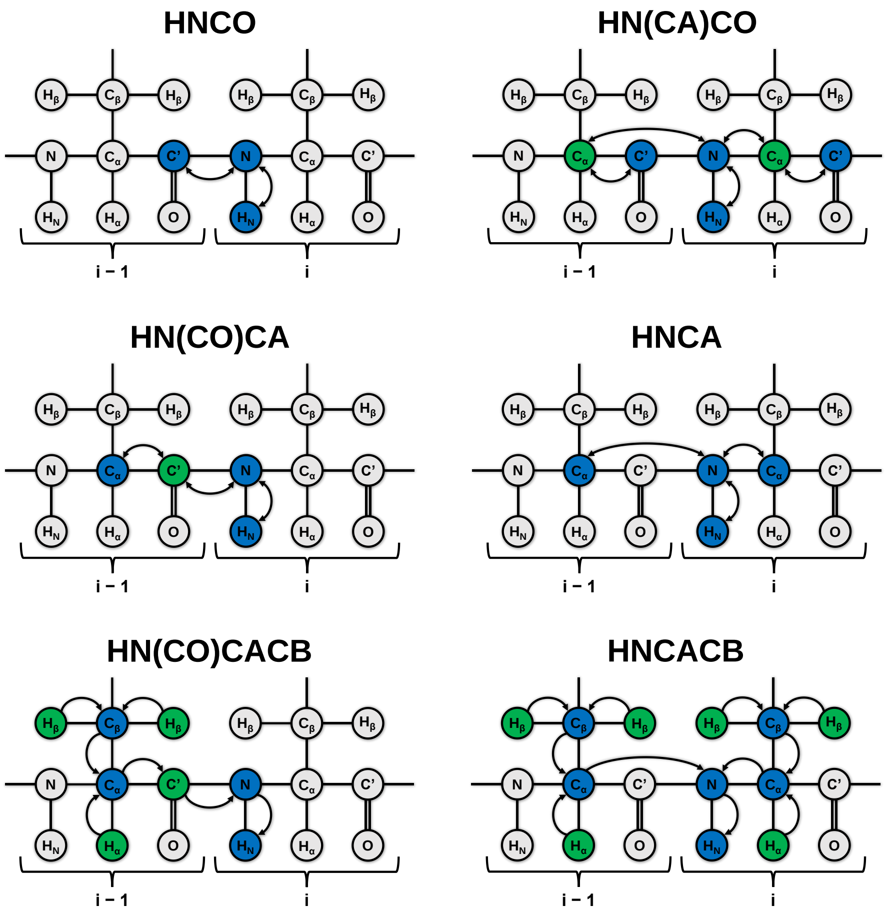


FIGURE 13: NMR experiments used for backbone assignment

All six NMR experiments are based on a [^{15}N,^{1}H]-correlated spectrum with an additional carbon dimension. In the HN(CA)CO spectrum, the carbonyl carbons of the observed residue and the preceding residue are visible, while the HNCO spectrum shows only the carbonyl carbon of the preceding residue. Similarly, the HNCA spectrum shows the alpha carbons of the observed residue and the preceding residue, while in the HN(CO)CA spectrum only the alpha carbon of the preceding residue is visible. In addition, the HNCACB spectrum shows the alpha carbons and beta carbons of the observed residue and the preceding residue, while the HN(CO)CACB spectrum shows only those of the preceding residue. In each experiment, the observed atoms and atoms through which magnetization flows are highlighted in blue and green, respectively.

Since all six NMR experiments are based on an [^{15}N,^{1}H]-correlated spectra such as a heteronuclear single quantum coherence (HSQC) spectrum or transverse relaxation optimized spectroscopy (TROSY) spectrum with an additional carbon dimension, it is possible to assign the apparent peaks in the HSQC spectrum with the respective chemical shifts of the six spectra. From the HN(CA)CO spectrum, information about the carbonyl carbons of the observed residue and the preceding residue can be obtained, while the HNCO spectrum contains information only about the preceding residue. Similarly, the HNCA spectrum contains information about the alpha carbons of the observed residue and the preceding residue, while the HN(CO)CA spectrum contains information only about the preceding residue. [122]

In addition, the HNCACB spectrum shows the alpha carbons and beta carbons of the observed residue and the preceding residue, while the HN(CO)CACB spectrum shows only those of the preceding residue. With this sophisticated approach of matching chemical shifts of carbonyl, alpha, and beta carbons of observed residues and preceding residues, a sequential assignment of the peaks in the [^{15}N,^{1}H]-HSQC spectrum can be performed. [122]

In practice, these six experiments will provide the backbone assignment, where the obtained C_α and C_β assignments can also be used to determine side chain connectivity that are necessary for high-resolution structure determination using NOE distance constraints [122]. The side chain amide resonances of asparagine and glutamine can be assigned using the HN(CO)CACB experiment, in which the signals in the carbon dimension correspond to the C_β and C_α or C_γ and C_β chemical shifts of asparagine and glutamine, respectively [143]. Further side chain resonances can be assigned using ^{13}C-edited and ^{15}N-edited CC(CO)NH [144] and H(CCO)NH [144, 145] experiments, which have an additional heteronuclei dimension and can be based on an [^{15}N,^{1}H]-HSQC spectra. The CC(CO)NH spectrum provides information about all carbon nuclei in the side chain of the preceding residue, while in the H(CCO)NH spectrum hydrogen chemical shifts in the side chain of the preceding residue are visible [144, 145]. In addition, HNHA [146] and HNHB [147] spectra can be recorded, which are based on an [^{15}N,^{1}H]-HSQC spectrum and provide information about the H_α and H_β chemical shift of the preceding residue, respectively. A [^{13}C,^{1}H]-HSQC spectrum can be used to visualize all carbon side chain resonances and allows further assignment in case of ambiguity.

1.4.4. Protein structure determination with experimental NMR data

After assigning backbone and side chain resonances, the next step is to transform the information into distances and torsion angles for structure calculation [148]. For structure calculations, it is useful to determine the secondary structure using a combination of chemical shifts, J coupling constants, and NOE data [122]. The secondary structure of a protein can be estimated using the chemical shifts obtained from the backbone resonance assignment, since secondary structure and chemical shifts are closely related [149, 150]. For this purpose, the observed chemical shifts of amino acid residue (e.g., C_α and C_β) are compared with the chemical shifts of the corresponding amino acid residues in a random coil conformation, as shown in EQUATION 6, where a positive value corresponds to an α-helical and a negative value to a β-sheet secondary structure [149, 150].

EQUATION 6: $$\Delta\delta(C_\alpha) - \Delta\delta(C_\beta) = (\delta C_{\alpha,exp} - \delta C_{\alpha,rc}) - (\delta C_{\beta,exp} - \delta C_{\beta,rc})$$

$\delta C_{\alpha,exp}$ / $\delta C_{\beta,exp}$	experimental C_α / C_β chemical shift, ppm
$\delta C_{\alpha,rc}$ / $\delta C_{\beta,rc}$	random coil C_α / C_β chemical shift, ppm

The chemical shift difference between its observed value and its random coil value is called secondary shift [149, 150]. Since homologous proteins have similar patterns in the chemical shifts of secondary structure, this relation was used to develop TALOS, which searches a database for triplets of adjacent residues with secondary chemical shifts and sequence similarity that show the best match to the triplet being searched for [148, 151]. The result is the predicted dihedral angles Φ and Ψ, which can be used directly for structure calculation [148].

Another approach uses homonuclear $^3J(H_N–H_\alpha)$ coupling constants to define the secondary structure of a protein, since the magnitude of the coupling constant depends on the dihedral angle Φ [146]. The $^3J(H_N–H_\alpha)$ coupling constant can be described by the Karplus equation [152], which was empirically parameterized with the relationship between $^3J(H_N–H_\alpha)$ and the dihedral angle Φ derived from the structure of ubiquitin [153], as shown in EQUATION 7.

EQUATION 7: $$^3J(H_N–H_\alpha) = 6.98\cdot\cos^2(\Phi - 60°) - 1.38\cdot\cos(\Phi - 60°) + 1.72$$

$^3J(H_N–H_\alpha)$ coupling constants are often small for helical residues (<5 Hz), but large for β-sheet residues (>8 Hz) [122], as shown in TABLE 3.

TABLE 3: Typical $^3J(H_N–H_\alpha)$ coupling constants in secondary structure elements

$^3J(H_N–H_\alpha)$ coupling constant are typically <5 Hz for helical and >8 Hz for β-sheet protein secondary structure elements. The values below were obtained from WÜTHRICH (1986) [154].

Secondary structure	**Φ**	**$^3J(H_N–H_\alpha)$**
α-helix	-57°	3.9 Hz
3_{10}-helix	-60°	4.2 Hz
antiparallel β-sheet	-139°	8.9 Hz
parallel β-sheet	-119°	9.7 Hz
random coil	–	6 – 8 Hz

$^3J(H_N–H_\alpha)$ coupling constants can be measured by quantitative J experiments, in which the intensity ratio between the cross peak and the diagonal signal in an HNHA experiment is determined, where the $^3J(H_N–H_\alpha)$ coupling constant can be determined according to EQUATION 8 [146].

EQUATION 8: $$^3J(H_N–H_\alpha) = (2\cdot\pi\cdot\xi\cdot e)^{-1}\cdot \arctan \sqrt{-(S_{cross}\cdot S_{dia}^{-1})}$$

with $$S_{cross}\cdot S_{dia}^{-1} = -\tan^2(2\cdot\pi\cdot J_{HH}\cdot\xi)$$

S_{cross}	intensity of the cross peak
S_{dia}	intensity of the diagonal peak
J_{HH}	H_N-H_N J coupling
ξ	time span of $^3J(H_N–H_\alpha)$ coupling evolvement
e	correction factor for relaxing properties of $H_N–H_\alpha$ / $H_N–H_N$

In addition, secondary structure elements can be identified by sequential NOEs, since each type of secondary structure element is characterized by a specific pattern of short-range NOEs ($|r_i – r_j| < 5$ Å) [122], as shown in TABLE 4.

TABLE 4: Useful distances in secondary structures elements

Sequential and medium range ^1H-^1H distances for identification of secondary structure elements in polypeptide chains. The distances below were obtained from WÜTHRICH *et al.* (1995) [155].

$^1H–^1H$ Distance	**α-helix**	**3_{10}-helix**	**antiparallel β-sheet**	**parallel β-sheet**
α–N	3.5 Å	3.4 Å	2.2 Å	2.2 Å
α–N (i, i+2)	4.4 Å	3.8 Å	–	–
α–N (i, i+3)	3.4 Å	3.3 Å	–	–
α–N (i, i+4)	4.2 Å	–	–	–
N–N	2.8 Å	2.6 Å	3.3 Å	4.0 Å
N–N (i, i+2)	4.2 Å	4.1 Å	–	–
α–β (i, i+3)	2.5 – 4.4 Å	3.1 – 5.1 Å	–	–

For example, α-helices are characterized by a stretch of strong to medium $NH_i–NH_{i+1}$ NOEs and medium to weak $C_\alpha H_i–NH_{i+3}$ and $C_\alpha H_i–NH_{i+1}$ NOEs, sometimes supplemented by $NH_i–NH_{i+2}$ and $C_\alpha H_i–NH_{i+4}$ NOEs [122]. In contrast, β-strands show very strong $C_\alpha H_i–NH_{i+1}$ NOEs, while other short-range NOEs involving the NH and $C_\alpha H$ protons are absent [122].

NOEs between the β-strands involving the NH, $C_{\alpha}H$, and $C_{\beta}H$ protons can be easily used to identify and align β-sheets structures [122]. NOEs are the most important NMR parameters for structure determination, as they provide information on short-range distances that are used for defining secondary structure elements and long-range distances that provide crucial tertiary structural information [122, 154]. Another preliminary step in structural determination is the stereospecific assignment of diasterotopic protons, where stereospecific assignments provide more accurate distance information and also dihedral angle information including χ_1 and χ_2 [122]. The assignment of β-methylene protons allows the determination of the side chain torsion angle χ_1, whereas only three different amino acid side chain χ_1 rotamers are possible for steric reasons [156], as shown in FIGURE 14.

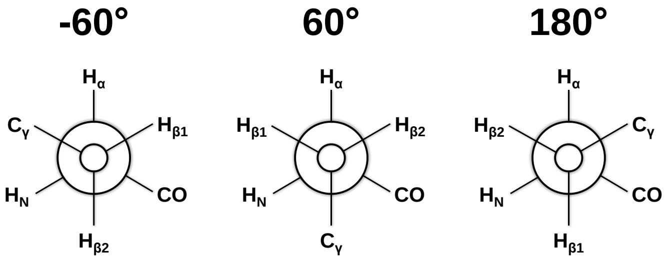


FIGURE 14: Newman projection of χ_1 rotamers

Due to steric reasons, three different confirmations are possible for the amino acid side chain χ_1 rotamer (-60°, 60°, 180°). The hydrogen and carbon atoms are shown according to Newman projection and labeled according to XPLOR nomenclature.

The diastereotopic H_{β} protons can be assigned by a combined approach where the ^3J coupling constant between the backbone ^{15}N and the two β-methylene protons taken from an HNHB experiment [147] and the NOE pattern observed in 3D-^{15}N and ^{13}C edited NOESY experiments are used [156], as shown in TABLE 5.

TABLE 5: Determination of the χ_1 rotamer

The amino acid side chain χ_1 rotamer (-60°, 60°, 180°) can be determined from the signal intensity of $H_{\beta 1}$ and $H_{\beta 2}$ in different NMR experiments, according to CLORE and GRONENBORN (1989) [156].

χ_1 rotamer	$^3J(H_N–H_{\beta 1})$	$^3J(H_N–H_{\beta 2})$	NOE ($H_N–H_{\beta 1}$)	NOE ($H_N–H_{\beta 2}$)	NOE ($H_{\alpha}–H_{\beta 1}$)	NOE ($H_{\alpha}–H_{\beta 2}$)
-60°	strong	weak	weak	strong	strong	weak
60°	weak	strong	strong	weak	strong	strong
180°	weak	weak	strong	strong	weak	strong

The amino acid side chain torsion angle χ_2 can only be determined from NOE data, although the determination of χ_1 and χ_2 can be supported by empirical information on the relative occurrence in known protein structures [157]. In cases where diastereotopic assignment is not possible, pseudoatoms can be introduced to replace the two protons, but the use of pseudoatoms reduces the accuracy of experimentally determined conformational constraints [158]. As a rule, the diastereotopic side chain amide protons of asparagine and glutamine can be assigned with the proton chemical shifts of $H_{\delta1}$ / $H_{\varepsilon1}$ (E-configuration) located downfield and $H_{\delta2}$ / $H_{\varepsilon2}$ (Z-configuration) located upfield [159], and in addition, strong NOEs should be present between the H_β / H_γ protons of asparagine and the $H_{\delta1}$ / $H_{\varepsilon1}$ protons of glutamine [160]. The conformation of proline residues can be determined using chemical shift and NOE information. The dihedral angle Φ of prolines is always in the range of -60° due to the ring conformation [161]. The dihedral angle Ψ can take values of -55° or 145° [161], where a strong NOE contact between the amide proton of the preceding residue and the proline H_δ corresponds to a Ψ-angle of -55° while a weak or no NOE contact corresponds to a Ψ-angle of 145°. The ω-angle indicating a cis or trans conformation of proline can be determined using chemical shift and NOE data. In the cis conformation, the C_γ chemical shift is about 25 ppm and the C_β chemical shift is about 34 ppm, resulting in a large difference between C_β and C_γ [162]. In the trans conformation, the C_γ chemical shift is about 27 ppm and the C_β chemical shift is about 32 ppm, leading to a small difference between C_β and C_γ [162]. In addition, NOE data can be used to determine the conformation of the proline dihedral angle ω [154], as shown in TABLE 6.

TABLE 6: Determination of the proline dihedral angle ω

The conformation of the proline dihedral angle ω (cis or trans) can be determined by NOE contacts of the indicated atoms, where (i) corresponds to the proline residue and (i-1) to the preceding residue, according to WÜTHRICH (1986) [154].

Dihedral angle ω	$H_N^{(i-1)}$–$H_\alpha^{(i)}$	$H_\alpha^{(i-1)}$–$H_\alpha^{(i)}$	$H_N^{(i-1)}$–$H_\delta^{(i)}$	$H_\alpha^{(i-1)}$–$H_\delta^{(i)}$
cis (0°)	strong	strong	weak	weak
trans (180°)	weak	weak	strong	strong

The final structure determination is performed in an iterative process of NOE assignment and structure calculation, where additional NOEs can be obtained using an initial structure. The general principle of structure calculation is that the data for distance and torsion angle, which are based on empirical input data, are converted into a visible structure using calculation software [148].

Starting from a randomly folded structure calculated based on the empirical data and the amino acid sequence of the protein, the structure calculation software attempts to fold the initial structure in a way that satisfies all experimentally derived constraints [148]. Importantly, NMR data do not uniquely define the three-dimensional structure of a protein because constraints are included as ranges of allowed values, the data contain experimental uncertainties, and only a small subset of all possible constraints are observable [123]. Therefore, structure calculations are repeated several times to find an ensemble of (low energy) structures that match the input NMR data, whereas a good ensemble of structures minimizes the violations of the input constraints and minimizes the root-mean-square deviation between members of the ensemble [123]. Various methods have been developed to calculate protein structures using constraints derived from experimental NMR data [123], as shown in [163-165].

1.4.5. Binding site mapping and binding affinity determination

NMR spectroscopy is not only used to determine structures and dynamics of molecules, but has also become a powerful approach for studying the interactions between proteins and ligands, where the interactions can be studied by observing a binding-induced change in NMR signals [122]. For this purpose, different pulse sequences are used to observe changes in, for example, chemical shifts, mobility, relaxation properties, and NOEs [122]. Chemical shifts perturbations resulting from complex formation is a highly sensitive tool for identification of protein binding sites [123]. This procedure is known as chemical shift mapping and has been used in a variety of ways, including defining contacts between macromolecules and high-throughput screening of small ligands to determine structure-activity relationships [123, 166]. For chemical shift mapping, HSQC, TROSY, and other ^1H-^{15}N correlated spectra are widely used since these spectra are very sensitive and often well resolved [123]. However, [^1H-^{13}C]-spectra can also be used for chemical shift mapping. To identify the binding site on the protein, unlabeled ligand is titrated into a solution containing ^{15}N-labeled protein, and the associated spectral changes in the [^1H-^{15}N]-spectrum are observed [123]. The chemical shift changes upon ligand titration are often quantified using the weighted chemical shift change $\Delta\delta_{NH}$ [123, 167], as shown in EQUATION 9.

EQUATION 9:

$$\Delta\delta_{NH} = \sqrt{0.5 \cdot \Delta\delta_H^{2} + 0.02 \cdot \Delta\delta_N^{2}}$$

$\Delta\delta_{NH}$	weighted chemical shift change, ppm
$\Delta\delta_H$	change in chemical shift of ^1H, ppm
$\Delta\delta_N$	change in chemical shift of ^{15}N, ppm

The simplest kinetic scheme for binding of a ligand to a protein with equimolar stoichiometry is described as a second-order exchange process [123, 168], as shown in EQUATION 10.

EQUATION 10: $$\mathrm{P} + \mathrm{L} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{PL} \qquad \text{with} \qquad K_A = k_{on}\cdot k_{off}^{-1} \text{ and } K_D = k_{off}\cdot k_{on}^{-1}$$

P	protein
L	ligand
PL	protein-ligand complex
k_{on}	rate of association
k_{off}	rate of dissociation
K_A	association constant
K_D	dissociation constant

The observed chemical shift δ_{obs} for labeled protein can be determined as shown in EQUATION 11 [123, 168].

EQUATION 11: $$\delta_{obs} = X_P\cdot\delta_P + X_{PL}\cdot\delta_{PL} \qquad \text{or:}$$

$$\delta_{obs} = \delta_P + 0.5\cdot(\delta_{PL} - \delta_P)\cdot P_0^{-1}\cdot\left(L_0 + P_0 + K_D - \sqrt{(L_0 + P_0 + K_D)^2 - 4\cdot L_0\cdot P_0}\right)$$

$$\text{with} \quad X_P + X_{PL} = 1\ ,\ P_0 = P + PL\ ,\ L_0 = L + PL$$

$$\text{and} \quad K_D = P\cdot L\cdot PL^{-1} = (P_0 - X_{PL}\cdot P_0)\cdot(L_0 - X_{PL}\cdot P_0)\cdot(X_{PL}\cdot P_0)^{-1}$$

δ_{obs}	observed chemical shift, ppm
δ_P	chemical shift of the protein, ppm
δ_{PL}	chemical shift of the protein-ligand complex, ppm
X_P	fraction of the free protein, -
X_{PL}	fraction of the bound protein, -
P_0	total protein concentration, M
P	protein concentration, M
PL	concentration of the protein-ligand complex, M
L_0	total ligand concentration, M
L	ligand concentration, M
K_D	dissociation constant, M

It should be noted that the relationship between δ_{obs} and dissociation constant K_D is non-linear and that δ_{PL} is not directly determinable, which means that the K_D cannot usually be determined from a single NMR experiment [168]. The unknown parameters K_D and δ_{PL} can be determined by measurements with a series of different concentrations of protein / ligand and followed by data analysis using a linearization method or a non-linear curve fitting procedure [168].

2. MATERIALS AND METHODS

2.1. Materials

2.1.1. Chemicals and consumables

All chemicals and consumables, unless otherwise noted, were purchased from the following suppliers: AppliChem GmbH (Darmstadt, GER), Carl Roth GmbH + Co. KG (Karlsruhe, GER), Merck KGaA (Darmstadt, GER), SERVA Electrophoresis GmbH (Heidelberg, GER), Thermo Fisher Scientific (Waltham, USA), and VWR International, LLC. (Radnor, USA).

DPC and LDAO were purchased from Cube Biotech GmbH (Monheim, GER). NADH was purchased from AppliChem GmbH (Darmstadt, GER). ATP and NADPH were purchased from Carl Roth GmbH + Co. KG (Karlsruhe, GER). 3-PGA, ADP, adenosine monophosphate (AMP), cytidine triphosphate (CTP), GAP, guanosine diphosphate (GDP), GTP, and uridine triphosphate (UTP) were purchased from Merck KGaA (Darmstadt, GER). Phosphoenolpyruvate (PEP) was purchased from Roche Holding AG (Basel, CHE).

2.1.2. Kits, enzymes, bacteria strains, DNA and protein standards

The kits, enzymes, bacteria stains, deoxyribonucleic acid (DNA) standard, and protein standards used are listed in TABLE 7, TABLE 8, TABLE 9, and TABLE 10, respectively.

TABLE 7: List of kits used

Kit	**Company**
QuikChange Lightning Site-Directed Mutagenesis Kit	Agilent Technologies, Inc. (Santa Clara, USA)
Wizard® SV Gel and PCR Clean-Up System	Promega Corporation (Madison, USA)
Wizard® *Plus* SV Minipreps DNA Purification System	Promega Corporation (Madison, USA)

TABLE 8: List of enzymes used

Enzyme	Company
Deoxyribonuclease (DNase) I grade II, from bovine pancreas, 100 U·µL^{-1}	Roche Holding AG (Basel, CHE)
DpnI, 10 U·µL^{-1}	Promega Corporation (Madison, USA)
Lysozyme from chicken egg white, powder (crystalline), ~ 70,000 U·mg^{-1}	Merck KGaA (Darmstadt, GER)
Pfu DNA Polymerase	Promega Corporation (Madison, USA)
Thrombin, Bovine, ≥ 1,800 U·mg^{-1}	Merck KGaA (Darmstadt, GER)

TABLE 9: List of bacteria strains used

Escherichia coli (E. coli) BL21 (DE3) cells were used for recombinant protein expression, while *E. coli* NEB 5-Alpha and *E. coli* XL10 gold cells were used for plasmid amplification.

Strain	Genotype	Company
E. coli BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	New England Biolabs, Inc. (Massachusetts, USA)
E. coli NEB 5-Alpha	*fhuA2* Δ(*argF-lacZ*)*U169 phoA glnV44* Φ*80* Δ(l*acZ*)*M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	New England Biolabs, Inc. (Massachusetts, USA)
E. coli XL10 gold	Tet^{r} Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F´ *proAB lacI*^{q}*Z*Δ*M15* Tn*10* (Tet^{r}) Amy Cam^{r}]*	Agilent Technologies, Inc. (Santa Clara, USA)

TABLE 10: List of DNA and protein standards used

Enzyme	Company
Dual Color Protein Standard III	SERVA Electrophoresis GmbH (Heidelberg, GER)
FastLoad 100 bp DNA Ladder	SERVA Electrophoresis GmbH (Heidelberg, GER)
ROTI®Mark TRICOLOR	Carl Roth GmbH + Co. KG (Karlsruhe, GER)

2.1.3. Bacterial growth media

The bacterial growth media used, and their compositions can be found in TABLE 11.

TABLE 11: List of bacterial growth media used

All bacterial growth media were sterilized before use, either by autoclaving at 121 °C for 20 min (LB) or by sterile filtration with a 0.2 µm filter (M9 medium). LB medium was purchased as a premixed powder from Carl Roth GmbH + Co. KG (Karlsruhe, GER). Pre-mixed LB medium powder with 15 g·L^{-1} agar from Carl Roth GmbH + Co. KG (Karlsruhe, GER) was used to prepare agar plates. M9 medium was prepared using D_2O instead of H_2O. D_2O and $^{15}NH_4Cl$ were purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, USA). SOC medium was purchased pre-mixed from New England Biolabs, Inc. (Massachusetts, USA). Ampicillin purchased from Merck KGaA (Darmstadt, GER) was added to LB and M9 at a final concentration of 100 µg·mL^{-1} just before use. (1) Ethylenediaminetetraacetic acid (EDTA).

Medium	**Component**		**Medium**	**Component**	
LB	Tryptone	10 g·L^{-1}	**SOC**	Tryptone	20 g·L^{-1}
	Yeast extract	5 g·L^{-1}		Yeast extract	5 g·L^{-1}
	NaCl	10 g·L^{-1}		Glucose	20 mM
M9	Na_2HPO_4	7.5 g·L^{-1}		$MgCl_2$	10 mM
	KH_2PO_4	3.5 g·L^{-1}		$MgSO_4$	10 mM
	NaCl	0.5 g·L^{-1}		NaCl	10 mM
	$^{15}NH_4CL$	1 g·L^{-1}		KCl	2.5 mM
	Glucose	5 g·L^{-1}	**Trace elements**	EDTA (1)	134 mM
	Biotin	1 mg·L^{-1}		$FeCl_3$	31 mM
	Thiamine	1 mg·L^{-1}		$ZnCl_2$	6.2 mM
	$MgSO_4$	5 mM		$CuCl_2$	760 µM
	$CaCl_2$	0.2 mM		$CoCl_2$	420 µM
	Trace elements	1 mL·L^{-1}		H_3BO_3	1.62 mM
				$MnCl_2$	81 µM

2.1.4. Equipment

The equipment used can be found in TABLE 12.

TABLE 12: List of equipment used

Instrument	Company
ÄKTA™ pure protein purification system equipped with HiLoad Superdex 200 pg and Superdex 200 Increase 10/300 GL	GE Healthcare (Chicago, USA)
Autoclave Systec VX-150	Systec GmbH (Linden, GER)
CD spectrometer J-1500 equipped with	JASCO Corporation (Tokyo, JPN) /
air-to-water recirculating cooler AWC100	Julabo GmbH (Seelbach, GER)
Cell density meter WPA CO8000	Biochrome Ltd. (Cambridge, GBR)
Centrifuge Sorvall LYNX 6000 Superspeed Centrifuge equipped with fixed angle rotors Fiberlite™ F9-6 x 1000 LEX and T29-8 x 50	Thermo Fisher Scientific (Waltham, USA)
Centrifuge 5810 R - Benchtop Centrifuge	Eppendorf SE (Hamburg, GER)
Diaphragm pump MD 1	VACUUBRAND GmbH & Co. KG (Wertheim, GER)
Balance Acculab® ATL-84-I Atilon	Sartorius AG (Göttingen, GER)
Balance Entris 2202-1S	Sartorius AG (Göttingen, GER)
Balance Pioneer PA114C Analytical Balance	OHAUS Corporation (Parsippany, USA)
Freezer CD615-E	Polar Refrigeration (Bristol, GBR)
Freezer FN6192CW	Gorenje gospodinjski aparati, d. o. o (Velenje, SVN)
Gel electrophoresis PowerPac™ Basic Power Supply equipped with SE 250 Mini-Vertical Gel Electrophoresis Unit or equipped with	Bio-Rad Laboratories, Inc. (Hercules, USA)
Sub-Cell® GT Agarose Gel Electrophoresis Systems	Hoefer Inc. (Holliston, USA)
Heating block Analog Dry Block Heater	VWR International, LLC. (Radnor, USA)
Homogenizer SilentCrusher M	Heidolph Instruments GmbH & Co.KG (Schwabach, GER)
Incubator Heratherm IGS60	Thermo Fisher Scientific (Waltham, USA)
Incubator Shakers New Brunswick™ Innova® 40 and 44	Eppendorf SE (Hamburg, GER)
ITC MicroCal PEAQ-ITC	Malvern Panalytical Ltd. (Malvern, GBR)
Laboratory glassware washer PG 8583 CD	Miele & Cie. KG (Gütersloh, GER)
Magnetic hotplate stirrer VMS-C7 Advanced	VWR International, LLC. (Radnor, USA)

Instrument	Company
Magnetic stirrer IKA® small squid and IKA® big squid	IKA®-Werke GmbH & CO. KG (Staufen, GER)
Microcentrifuges 5424 and 5424 R	Eppendorf SE (Hamburg, GER)
Micro centrifuge Micro Star 17R	VWR International, LLC. (Radnor, USA)
Microwave HF12M240	Siemens AG (München, GER)
Mini centrifuge ROTILABO®	Carl Roth GmbH + Co. KG (Karlsruhe, GER)
Nutating Mixer	VWR International, LLC. (Radnor, USA)
PCR cycler Mastercycler® nexus X2	Eppendorf SE (Hamburg, GER)
Peristaltic pump P-1	GE Healthcare (Chicago, USA)
pH bench meter FiveEasy™ FE20-ATC kit	Mettler-Toledo International Inc. (Columbus, USA)
Photometer NanoPhotometer® N60	Implen GmbH (München, GER)
Pipette controller accu-jet™ pro	BRAND GmbH & Co. KG (Wertheim, GER)
Pipettes Research® plus (0.1 - 2.5 µL / 2 - 20 µL / 20 - 200 µL / 100 - 1,000 µL)	Eppendorf SE (Hamburg, GER)
Refrigerator FK 3640	Liebherr-International S.A. (Bulle, CHE)
Refrigerator FC156.3AA	Amica S.A. (Wronki, POL)
Refrigerator-Freezer KG33SV10	Siemens AG (München, GER)
Refrigerator-Freezer RK6192EX	Gorenje gospodinjski aparati, d. o. o (Velenje, SVN)
Shaker ROCKER 3D digital	IKA®-Werke GmbH & CO. KG (Staufen, GER)
Spectrofluorometer FP-8300 equipped with mini water circulation bath MCB-100	JASCO Corporation (Tokyo, JPN)
Thermal mixer ThermoMixer® C	Eppendorf SE (Hamburg, GER)
Ultra-Low Temperature Freezer HERAfreeze™ HFU T	Thermo Fisher Scientific (Waltham, USA)
Ultrasonic bath Sonorex RK 103 H	BANDELIN electronic GmbH & Co. KG (Berlin, GER)
Ultrasonic cell disruptor Sonifier™ S-250D	Branson Ultrasonics (Brookfield, USA)
UV Transilluminator Mini Benchtop M-10E	UVP, LLC. (Upland, USA)
Vortex Mixer	VWR International, LLC. (Radnor, USA)

2.1.5. Software

The software used can be found in TABLE 13.

TABLE 13: List of software used

Software	**Company**
Adobe Illustrator CS5 Version 15.0.2	Adobe, Inc. (San José, USA)
EndNote Version X9.3.3	Clarivate Plc (London, GBR)
Expasy ProtParam Tool	SIB Swiss Institute of Bioinformatics (Lausanne, CHE) https://web.expasy.org/protparam
Expasy Translate Tool	SIB Swiss Institute of Bioinformatics (Lausanne, CHE) https://web.expasy.org/translate
MicroCal PEAQ-ITC Analysis Software Version 1.21	Malvern Panalytical Ltd. (Malvern, GBR)
Microsoft Office 2013	Microsoft Corporation (Redmond, USA)
NMRFAM-SPARKY Version 1.470 powered by Sparky Version 3.190	National Magnetic Resonance Facility at Madison (Madison, USA) / [169]
OriginPro® Version 9.0 and 9.9.5	OriginLab Corporation (Northampton, USA)
PyMOL Version 1.8.4.2	Schrödinger, Inc. (New York, USA)
TopSpin® Version 3.2	Bruker Corporation (Billerica, USA)
Unicorn 7.1	GE Healthcare (Chicago, USA)
Spectra Manager^{TM} Version 2.13.00 and 2.15.01	JASCO Corporation (Tokyo, JPN)

2.2. Molecular biological methods

2.2.1. Generation of the OEP21 variants by site-directed mutagenesis

The gene encoding for OEP21 from *Pisum sativum* (GeneBank: AJ009987.1) was inserted into a modified pET-21a vector with ampicillin resistance (Merck KGaA / Darmstadt, GER) by polymerase chain reaction (PCR)-based restriction-free cloning, resulting in OEP21 harboring a non-cleavable C-terminal His$_{10}$-tag, as shown in the appendix in TABLE 30. This procedure, the generation of OEP21 variants K19A, R51A, R66A, R84A, and C109A, and the generation of OEP21-ΔHis (without His$_{10}$-tag, L178, and E179) was performed by Dr. Elisabeth Häusler [170]. To generate further OEP21 variants, site-directed mutagenesis was performed using the QuikChange Lightning Site-Directed Mutagenesis Kit. The sample preparation and thermocycler protocol are shown in TABLE 14, while the oligonucleotide primers used are listed in TABLE 15. After site-directed mutagenesis, the generated products were digested with DpnI for 1 h at 37 °C. To amplify the generated plasmids, transformation of *E. coli* and plasmid extraction were conducted as described in chapter 2.2.2. To verify the mutations, the plasmids were sent for sequencing as described in chapter 2.2.3.

TABLE 14: QuikChange site-directed mutagenesis

Shown are the sample preparation and thermocycler protocol used to generate the OEP21 variants using the QuikChange Lightning Site-Directed Mutagenesis Kit from Agilent Technologies, Inc. (Santa Clara, USA). The enzyme was added just before the start of the reaction. (1) Deoxyribonucleoside triphosphate (dNTP).

Preparation		**PCR-Protocol**		
10x QuikChange Lightning Buffer	2.5 µL			
dNTP (1) mix	0.5 µL	95 °C	2 min	
QuikSolution reagent	0.75 µL	95 °C	30 s	20 cycles
Primer 1 and 2 [10 µM]	each 0.5 µL	60 °C	30 s	
Plasmid DNA template	50 ng	68 °C	7 min	
QuikChange Lightning Enzyme	0.5 µL	68 °C	10 min	
H_2O	add to 25 µL			

Table 15: Primers used for generating the OEP21 variants

OEP21 variants were generated by site-directed mutagenesis with the following oligonucleotides, all shown in the 5’-3’ direction.

Mutation	Primer 1	Primer 2
ΔLoop5	GGGTTGGTGT GCGGTATGAT AAACGTGATA AGTTACGGTA TACTGTACTT GCTAAG	CTTAGCAAGT ACAGTATACC GTAACTTATC ACGTTTATCA TACCGCACAC CAACCC
R33A	CACCAACACC TTCTTCCAGG TTGCTGGAGG GCTTGACACA AAAAC	GTTTTTGTGT CAAGCCCTCC AGCAACCTGG AAGAAGGTGT TGGTG
R66W	GCAACGCTTG GGGTTGGTGT GTGGTATGAT AAACAAGATT CTGTAGGAG	CTCCTACAGA ATCTTGTTTA TCATACCACA CACCAACCCC AAGCGTTGC
N164R (used along with R66A)	CAGATTAGGG AGAATAATTG GACATTTAGA GCAGATTACA AAGGTAGATG G	CCATCTACCT TTGTAATCTG CTCTAAATGT CCAATTATTC TCCCTAATCT G
K90A	GATAAGTTAC GGTATACTGT ACTTGCTGCG AAGACGTTTC CTGTGACAAA TG	CATTTGTCAC AGGAAACGTC TTCGCAGCAA GTACAGTATA CCGTAACTTA TC
K104A	GACAAATGAT GGTCTTGTTA ACTTCGCAAT TAAAGGAGGA TGTGATGTTG AC	GTCAACATCA CATCCTCCTT TAATTGCGAA GTTAACAAGA CCATCATTTG TC
R140A	CAATTTTCAG AAGGATCAAG ATGTTGCACT CAGAATTGGT TATGAAGCTT TTG	CAAAAGCTTC ATAACCAATT CTGAGTGCAA CATCTTGATC CTTCTGAAAA TTG
R142A	CAGAAGGATC AAGATGTTAG ACTCGCAATT GGTTATGAAG CTTTTGAACA G	CTGTTCAAAA GCTTCATAAC CAATTGCGAG TCTAACATCT TGATCCTTCT G
R157A	GAACAGGTTC CTTATCTGCA GATTGCGGAG AATAATTGGA CATTTAATGC AG	CTGCATTAAA TGTCCAATTA TTCTCCGCAA TCTGCAGATA AGGAACCTGT TC
R174A	GATTACAAAG GTAGATGGAA TGTGGCATAT GACTTGCTCG AGCATCATC	GATGATGCTC GAGCAAGTCA TATGCCACAT TCCATCTACC TTTGTAATC

2.2.2. Transformation of *E. coli* and plasmid amplification

For plasmid amplification, the respective plasmid was transformed into *E. coli* NEB 5-Alpha or XL10 gold competent cells. For this purpose, about 50 ng plasmid were incubated with the competent cells for 30 min on ice. After incubation for 30 s at 42 °C, cells were incubated for 2 min on ice. Then, 200 µL SOC medium (see Table 11 for composition) was added to the cells before gentle shaking at 37 °C for 1 h. Subsequently, cells were plated out on LB agar plates containing 100 µg·mL^{-1} ampicillin and incubated overnight at 37 °C. The next day, a single colony was picked and transferred into 10 mL LB medium containing additionally 2 wt% glucose and 100 µg·mL^{-1} ampicillin. Cells were grown overnight at 37 °C with gentle shaking.

Subsequently, plasmid extraction and purification were performed using the Wizard® Plus SV Minipreps DNA Purification System according to the manufacturer's protocol. Finally, the concentration of the purified plasmid was determined by measuring the absorbance at 260 nm using a photometer. The plasmids were stored at -20 °C.

2.2.3. Sequencing of plasmid DNA

Sequencing of plasmid DNA was conducted by GENEWIZ Germany GmbH (Leipzig, GER). For each reaction, 7.5 µL plasmid DNA with a concentration of 50 ng·µL^{-1} to 100 ng·µL^{-1} were mixed with 2.5 µL of the respective sequencing primer, which had a concentration of 10 µM. TABLE 16 lists the primer used for sequencing.

TABLE 16: Primer used for DNA sequencing

The oligonucleotides used for DNA sequencing are shown in the 5'-3' direction.

Primer	Primer 1	Primer 2
T7	`TAA TAC GAC TCA CTA TAG GG`	`GCT AGT TAT TGC TCA GCG G`
pQE	`CCC GAA AAG TGC CAC CTG`	`GTT CTG AGG TCA TTA CTG G`

2.2.4. Generation of the ADAM17 and iRhom2 constructs

To produce the transmembrane domains of ADAM17 and iRhom2, co-expression of Get3 should be performed. Therefore, the corresponding DNA sequence of the transmembrane domain of ADAM17 / iRhom2 was cloned into a pT5*L*/T7 vector (derived from pQE80) consisting of Get3-maltose-binding protein (MBP) under the T5/*Lac* promotor and the target protein fused to a tobacco etch virus protease (TEV)-cleavable Z-tag under the T7 promotor. The modified pT5*L*/T7 vector with the corresponding features was generated by Dr. Kolio Raltchev [171]. The genes encoding human ADAM17-TM (GeneBank: NP_001369706.1) and human iRhom2-TM1 / iRhom2-7TM (GeneBank: AAH35829.1) were obtained from constructs that were generated by Dr. Elisabeth Häusler. TABLE 17 lists all primers used to generate the ADAM17 and iRhom2 constructs (see TABLE 33 and TABLE 34 in the appendix for amino acid sequences).

TABLE 17: Primer used for generating the ADAM17 and iRhom2 constructs

Primers marked with an asterisk were identical.

Construct	Primer 1	Primer 2
ADAM17-TM-TEV-Z-tag	CATGTCAGGA GAAAACCTTT ATTTTCAGGG TAATACCTTC GGCAAATTTC TGGCGG	CGAGTGCGGC CGCAAGCTTT CATCCTTCGT ACTGTTTGTC CAGTTTCTTA TC
ADAM17-TM-Z-tag	GCCGAAAGTA GCCATGTCAG GATTAGTTCC GCGTGGATCC AATACCTTCG GCAAATTTCT GGCGG	CCGCCAGAAA TTTGCCGAAG GTATTGGATC CACGCGGAAC TAATCCTGAC ATGGCTACTT TCGGC
ADAM17-TM-GB1	GGCACCACCA TCACCATCAC CACCATCACC ACGATTACGA TATCCCAACG (*)	GTTATCCGCC AGAAATTTGC CGAAGGTATT AGAAGAACCA CCAGAACCAC GCG
iRhom2-TM1-Z-tag	CATGTCAGGA TTAGTTCCGC GTGGATCCGG TTCTCAACGC CAGTTAGAGT CGTTTGATTC CC	CGAGTGCGGC CGCAAGCTTT CATGATTCAT ACACCCCTTT GTTGCGCAGA AC
iRhom2-TM1-GB1	GGCACCACCA TCACCATCAC CACCATCACC ACGATTACGA TATCCCAACG (*)	GGAATCAAAC GACTCTAACT GGCGTTGAGA AGAACCACCA GAACCACGCG
iRhom2-7TM-Z-tag	CATGTCAGGA TTAGTTCCGC GTGGATCCGG TTCTTCGTTC GATTCCCATC GCCCG	CGAGTGCGGC CGCAAGCTTT CACTCGAGGG TCAGATGCTC AATCCACGGC CAG
iRhom2-7TM-Integrin	GGTTCTGGTT TAGTTCCGCG TGGATCCTCG TTCGATTCCC ATCGCCCG	GGTGGTGGTG ATGATGATGA TGCTCGAGGG TCAGATGCTC AATCCACGGC CAG

ADAM17-TM-TEV-Z-tag

To generate the ADAM17-TM-TEV-Z-tag construct, PCR-based restriction-free cloning was conducted as described in chapter 2.2.5. In a first PCR, a megaprimer consisting of the ADAM17-TM DNA sequence elongated with the non-overlapping primer ends was generated, which was inserted into the modified pT5*L*/T7 vector in a second PCR.

ADAM17-TM-Z-tag

The sequence for TEV cleavage was modified into a sequence for thrombin cleavage using the QuikChange Lightning Site-Directed Mutagenesis Kit from Agilent Technologies, Inc. (Santa Clara, USA). The sample preparation and thermocycler protocol are shown in TABLE 14, where the ADAM17-TM-TEV-Z-tag construct was used as originating plasmid.

ADAM17-TM-GB1

Because the protein G B1 domain (GB1) is more suitable for refolding after chemical denaturation, the Z-tag was changed to GB1. Therefore, PCR-based restriction-free cloning was conducted as described in chapter 2.2.5. The generated megaprimer consisted of a sequence encoding GB1, which was derived from a construct generated by Dr. Elisabeth Häusler, and the priming sequences for insertion into the ADAM17-TM-Z-tag construct.

iRhom2-TM1-Z-tag

The sequence encoding the first transmembrane domain of iRhom2 was inserted into the ADAM17-TM-Z-tag construct by PCR-based restriction-free cloning according to chapter 2.2.5.

iRhom2-TM1-GB1

Analogous to ADAM17-TM-Thrombin-GB1, PCR-based restriction-free cloning was conducted as described in chapter 2.2.5. The generated megaprimer consisted of a sequence encoding GB1, which was derived from a construct generated by Dr. Elisabeth Häusler, and the priming sequences for insertion into the iRhom2-TM1-Z-tag construct.

iRhom2-7TM-Z-tag

The sequence encoding all seven transmembrane domains of iRhom2 was inserted into the ADAM17-TM-Z-tag construct by PCR-based restriction-free cloning according to chapter 2.2.5.

iRhom2-7TM-Integrin

The sequence encoding all seven transmembrane domains of iRhom2 was inserted into a modified pET-21a vector (Merck KGaA / Darmstadt, GER) by PCR-based restriction-free cloning according to chapter 2.2.5. The vector was generated by Dr. David Goričanec according to ARCEMISBÉHÈRE *et al.* (2010) [172], with a thrombin-cleavable α_5 integrin fragment (amino acid residues 231 to 514) and a His_{10}-tag fused to a target sequence. Since only a fragment of α_5 integrin was used, the resulting construct was insoluble and was therefore expressed as inclusion bodies.

2.2.5. PCR-based restriction-free cloning

To produce a megaprimer for further use, a first PCR was conducted according to TABLE 18. Subsequently, the amplified DNA product was separated via agarose gel electrophoresis and purified with a kit. The composition of the buffers used for agarose gel electrophoresis is listed in TABLE 19, and electrophoresis was carried out at 180 V for 15 min while the camber was cooled with ice. To visualize the separation of DNA fragments, the agarose gel was irradiated with ultraviolet (UV) light. The FastLoad 100 bp DNA Ladder was used as a length reference for the DNA fragments. Fragments that matched the theoretical length were cut out of the agarose gel and purified using the Wizard® SV Gel and PCR Clean-Up System according to the manufacturer’s protocol. DNA concentrations of the purified samples were determined using a photometer at 260 nm.

Table 18: Preparation and protocol for the first PCR

The sample preparation and thermocycler protocol used for the first PCR of restriction-free cloning are shown. The Buffer, dNTP mix, and polymerase were purchased from Promega Corporation (Madison, USA). The polymerase was added just before the start of the reaction.

Preparation		**PCR-Protocol**		
10x Buffer with $MgSO_4$	5 µL			
dNTP mix [10 mM each]	1 µL	95 °C	2 min	
DNA-template	100 ng	95 °C	30 s	30 cycles
Primer 1 and 2 [10 µM]	each 2.5 µL	60 °C	30 s	
DMSO	2.5 µL	68 °C	1 min	
Pfu DNA Polymerase	1 µL	68 °C	10 min	
H_2O	add to 50 µL			

Table 19: Buffers used for agarose gel electrophoresis

TBE buffer was used for agarose gel preparation and as running buffer during gel electrophoresis. DNA samples were mixed with sixfold concentrated DNA loading buffer prior to gel electrophoresis. DNA Stain G was purchased from SERVA Electrophoresis GmbH (Heidelberg, GER). (1) Tris(hydroxymethyl)aminomethane (Tris).

TBE Buffer		**Agarose Gel**		**DNA Loading Buffer, 6x**	
Tris (1), pH 8.3	89 mM	Agarose	1 wt%	Tris, pH 7.5	60 mM
Boric acid	89 mM	DNA Stain Clear G	0.006 vol%	Sodium acetate	30 mM
EDTA	2 mM	all dissolved in TBE buffer		ETDA	12 mM
				Glycerol	60 wt%
				Bromophenol blue	0.18 wt%
				Xylene cyanole FF	0.12 wt%

The purified DNA product served as a megaprimer for a second PCR, which was conducted according to Table 20. Subsequently, the generated products were digested with DpnI for 1 h at 37 °C. To amplify the generated plasmids, transformation of *E. coli* and plasmid extraction were conducted as described in chapter 2.2.2. To verify cloning, plasmids were sent for sequencing as described in chapter 2.2.3.

TABLE 20: Preparation and protocol for the second PCR

The sample preparation and thermocycler protocol used for the second PCR of restriction-free cloning are shown. The QuikChange Lightning Site-Directed Mutagenesis Kit from Agilent Technologies, Inc. (Santa Clara, USA) was used, and the enzyme was added just before the start of the reaction. The megaprimer was used at a twenty-fold molar excess to the target plasmid.

Preparation		**PCR-Protocol**		
Reaction buffer	2.5 µL			
dNTP mix	0.5 µL	95 °C	2 min	
QuikSolution reagent	0.75 µL	95 °C	30 s	20 cycles
Destination plasmid	50 ng	68 °C	7 min	
Insert (Megaprimer)	20x molar excess	68 °C	10 min	
QuikChange Lightning Enzyme	0.5 µL			
H_2O	add to 25 µL			

2.3. Protein production and purification methods

2.3.1. Transformation of *E. coli* and pre-culture preparation

For protein expression, the respective plasmid was transformed into *E. coli* BL21(DE3) competent cells. For this purpose, about 50 ng plasmid were incubated with 25 µL competent cells for 30 min on ice. After incubation for 30 s at 42 °C, cells were incubated for 2 min on ice. Then, 200 µL SOC medium (see TABLE 11 for composition) were added to the cells before gentle shaking for 1 h at 37 °C. Cells were then transferred into LB medium containing additionally 2 wt% glucose and 100 $\mu g \cdot mL^{-1}$ ampicillin. This pre-culture was incubated overnight at 37 °C while shaking at 140 revolutions per minute (rpm). The next day, a certain amount of the pre-culture was transferred into the main culture to start with an initial optical density measured at a wavelength of 600 nm (OD_{600}) of around 0.2.

2.3.2. OEP21

E. coli BL21(DE3) cells were transformed with the respective plasmid according to chapter 2.3.1, with the constructs used shown in the appendix in TABLE 30. Cells were grown in LB medium containing 100 µg·mL^{-1} ampicillin to an OD_{600} of around 0.8 at 37 °C while shaking at 140 rpm. Protein production was then induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and further incubating the cells for 4 h to 5 h at 37 °C. Cells were finally harvested by centrifugation at 6,000*g* and 4 °C for 15 min before cell pellets were stored at -80 °C.

All buffers used for OEP21 purification are listed in TABLE 21. For lysis, cells were resuspended with a homogenizer in buffer A with an additional 1 vol% Triton X-100, 0.2 mg·ml^{-1} lysozyme, and one tablet of the cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche Holding AG / Basel, CHE). Cells were sonicated for 10 min (30 % amplitude, 1 s pulse, 2 s pause) and DNA was digested simultaneously by addition of 1 U·mL^{-1} DNaseI and 5 mM $MgCl_2$. After centrifugation at 38,759*g* and 4 °C for 20 min, inclusion bodies were washed once by resuspending in buffer A with an additional 1 vol% Triton X-100 and centrifuged again. Finally, inclusion bodies were dissolved in buffer B using a homogenizer. Before further purification by nickel^{2+} ions coupled to nitrilotriacetic acid (Ni-NTA) affinity chromatography, the solution was centrifuged at 38,759*g* and 4 °C for 20 min. The supernatant was filtered with a 0.45-µm-filter before incubation with the Ni-NTA resin (GE Healthcare / Chicago, USA) for 1 h on a shaker. The applied protein sample was then allowed to flow through the column before washing the resin with 10 column volume (CV) of buffer B, followed by elution of OEP21 with 10 CV of buffer B containing a total of 500 mM imidazole. The elution fraction was dialyzed overnight against dialysis buffer 1 in a dialysis tube with 3.5 kDa molecular weight cut-off (MWCO), and precipitated OEP21 was dissolved in buffer C at a concentration of 5 mg·mL^{-1}. Refolding of the protein was performed by rapid dropwise dilution into a tenfold excess volume of refolding buffer with moderate stirring at 4 °C for approximately 3 h. To remove residual guanidinium chloride (GdmCl), dialysis was performed overnight in dialysis buffer 2 in dialysis tube with 3.5 kDa MWCO. Final purification was done by size-exclusion chromatography (SEC) using an ÄKTA™ pure protein purification system equipped with a HiLoad Superdex 200 pg or a Superdex 200 Increase 10/300 GL column in SEC buffer at a flow rate of 1.5 mL·min^{-1} or 0.5 mL·min^{-1}, respectively. 500 mM NaCl was added to the samples before they were applied to SEC. Amicon® Ultra centrifugal filter units with 10 kDa or 30 kDa MWCO (Merck KGaA / Darmstadt, GER) were used to concentrate protein samples, e. g. for SEC. Protein samples were stored at -80°C and thawed just before use. To adapt the buffer for different methods, protein samples were dialyzed in the respective buffer.

TABLE 21: Buffers used for OEP21 purification

(1) β-Mercaptoethanol (BME), (2) sodium phosphate (NaPi), (3) dithiothreitol (DTT), (4) 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES).

Buffer A		Buffer B		Buffer C		Refolding Buffer	
50 mM	Tris, pH 8.0	6 M	GdmCl	6 M	GdmCl	20 mM	NaPi, pH 6.0
100 mM	NaCl	50 mM	Tris, pH 8.0	50 mM	NaPi (2), pH 6.0	50 mM	NaCl
1 mM	EDTA	100 mM	NaCl	100 mM	NaCl	1 mM	EDTA
10 mM	BME (1)	10 mM	Imidazole	5 mM	EDTA	5 mM	DTT
		5 mM	BME	10 mM	DTT (3)	10 vol%	Glycerol
						0.5 wt%	LDAO

Dialysis Buffer 1		Dialysis Buffer 2		SEC Buffer	
20 mM	Tris, pH 8.0	20 mM	NaPi, pH 6.0	20 mM	HEPES (4), pH 7.0
50 mM	NaCl	50 mM	NaCl	50 mM	NaCl
1 mM	EDTA	1 mM	EDTA	0.5 mM	EDTA
2.5 mM	BME	2.5 mM	BME	5 mM	DTT
				0.1 wt%	LDAO

2.3.3. OEP21-ΔHis

Expression and purification of OEP21-ΔHis (without His_{10}-tag, L178, and E179) was analogous to the procedure described in chapter 2.3.2, except that cation exchange chromatography was used instead of Ni-NTA affinity chromatography. For this purpose, inclusion bodies were finally dissolved in 8 M urea and 50 mM NaPi 6.0 instead of buffer B using a homogenizer and then centrifuged at 38,759*g* and 4 °C for 20 min. The supernatant was filtered with a 0.45-µm-filter before incubation with the SP Sepharose Fast Flow cation exchange chromatography resin (GE Healthcare / Chicago, USA) for 1 h on a shaker. The applied protein sample was then allowed to flow through the column before washing the resin with 10 CV of 8 M urea and 50 mM NaPi 6.0, followed by elution of the protein with 10 CV of 8 M urea, 250 mM NaC, and 50 mM NaPi 6.0. The elution fraction was dialyzed overnight and further processed according to chapter 2.3.2.

2.3.4. Isotope labelling

For NMR experiments, isotope-labeled protein had to be produced first. Transformation of *E. coli* and pre-culture preparation was conducted according to chapter 2.2.2. The next day, the pre-culture was centrifuged at 2,000*g* for 6 min before resuspending the cell pellet in LB medium with an additional 2 wt% glucose and 100 µg·mL^{-1} ampicillin, all prepared in D_2O purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, USA). The initial OD_{600} was adjusted to approximately 0.2, which was achieved by resuspending an appropriate amount of cells in the medium. Subsequently, cells were grown at 37 °C while shaking at 140 rpm in LB/D_2O medium until an OD_{600} of 1–1.5 was reached, which usually took around 6 h. Cells were then centrifuged at 2,000*g* for 6 min before resuspending the cell pellet in M9 medium containing a total of 1 wt% glucose and 100 µg·mL^{-1} ampicillin, all prepared in D_2O. Starting from an adjusted OD_{600} of approximately 0.2, cells were grown overnight at 37 °C while shaking at 140 rpm. The next day, this cell culture was diluted 1:10 with fresh M9 medium in D_2O containing 100 µg·mL^{-1} ampicillin to start the growth of the main culture. In this way, the main culture typically started with an OD_{600} of around 0.2. Cells were grown at 37 °C while shaking at 140 rpm to an OD_{600} of around 0.8. Protein production was then induced by adding 1 mM IPTG and further incubating the cells for 4 h to 5 h at 37 °C. Finally, cells were harvested by centrifugation at 6,000*g* and 4 °C for 15 min before cell pellets were stored at -80 °C.

2.3.5. Bna-Tic40

The ^{15}N-labelled *Brassica napus* (Bna)-Tic40 variants A10 and C9, and the ^{13}C,^{15}N-labelled variant A10 produced recombinantly in *E. coli* were purchased from Crelux GmbH (Martinsried, GER). The amino acid sequences for both proteins are shown in the appendix in TABLE 31. Proteins were stored in 20 mM NaPi pH 6.0, 50 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), and 0.5 mM EDTA at -80 °C.

2.3.6. ADAM17 and iRhom2

E. coli BL21(DE3) cells were transformed with the respective plasmid for ADAM17-TM, iRhom2-TM1, or iRhom2-7TM according to chapter 2.3.1, with the constructs used shown in the appendix in TABLE 33, TABLE 34, and TABLE 35, respectively. Cells were grown in LB medium containing 100 µg·mL^{-1} ampicillin to an OD_{600} of around 0.8 at 37 °C while shaking at 140 rpm. Protein production was then induced by adding 1 mM IPTG and further incubating the cells for 4 h to 5 h at 37 °C.

Finally, cells were harvested by centrifugation at 6,000*g* and 4 °C for 15 min before cell pellets were stored at -80 °C.

All buffers used for purification are listed in TABLE 22. For lysis, cells were resuspended with a homogenizer in buffer A with an additional 0.2 mg·ml^{-1} lysozyme and one tablet of the cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche Holding AG / Basel, CHE). Cells were sonicated for 10 min (30 % amplitude, 1 s pulse, 2 s pause) and DNA was simultaneously digested by addition of 1 U·mL^{-1} DNaseI and 5 mM $MgCl_2$. After centrifugation at 38,759*g* and 4 °C for 20 min, the supernatant was filtered with a 0.45-µm-filter before further purification by Ni-NTA affinity chromatography. The protein sample was incubated with the Ni-NTA resin (GE Healthcare / Chicago, USA) on a shaker for 30 min before the sample was allowed to flow through the column. The resin was washed with 10 CV of buffer B, followed by elution of OEP21 with 10 CV of buffer B containing a total of 500 mM imidazole. The elution fraction was dialyzed overnight against SEC buffer without DPC in a dialysis tube with 3.5 kDa MWCO. Final purification was done by SEC using an ÄKTA™ pure protein purification system equipped with a HiLoad Superdex 200 pg or a Superdex 200 Increase 10/300 GL column in SEC buffer at a flow rate of 1.5 mL·min^{-1} or 0.5 mL·min^{-1}, respectively.

Amicon® Ultra centrifugal filter units with 3 kDa or 10 kDa MWCO (Merck KGaA / Darmstadt, GER) were used to concentrate protein samples for SEC. Protein samples were stored at -80°C and thawed just before use.

TABLE 22: Buffers used for ADAM17 and iRhom2 purification

Buffer A		**Buffer B**		**Cleavage Buffer**		**SEC Buffer**	
50 mM	Tris, pH 8.0	50 mM	Tris, pH 8.0	20 mM	Tris, pH 8.0	20 mM	NaPi, pH 6.0
500 mM	NaCl	500 mM	NaCl	50 mM	NaCl	50 mM	NaCl
1 mM	EDTA	5 mM	BME	5 mM	BME	1 mM	EDTA
10 mM	BME	10 mM	Imidazole	5 mM	$CaCl_2$	5 mM	DTT
1 vol%	Triton X-100	0.2 wt%	DPC			0.1 wt%	DPC

If thrombin cleavage was performed, the elution fraction was dialyzed overnight against cleavage buffer. The next day, thrombin was added to the protein sample, and cleavage was conducted at 37 °C until complete cleavage was achieved, which usually took 1–2 h. The cleaved protein sample was then purified by a second Ni-NTA affinity chromatography under the same conditions as above, with the target protein was found in the flow-through and wash fractions. Both fractions were pooled and concentrated before final purification by SEC.

2.4. PROTEIN CHARACTERIZING methods

2.4.1. Protein concentration determination

To determine protein concentrations, absorbance was measured at 280 nm using NanoPhotometer® N60. Molar protein concentrations were calculated using the protein-specific extinction coefficients under reducing conditions found in the appendix in TABLE 30 (OEP21), TABLE 31 (Bna-Tic40), TABLE 33 (ADAM17), TABLE 34 and TABLE 35 (iRhom2).

2.4.2. Size estimation from SEC

To roughly estimate the size of protein samples based on the corresponding elution time in SEC, the columns were calibrated using different protein standards. Since the size of the protein standards is known, linear regression was applied to obtain column-specific elution times. This allowed size estimation with EQUATION 12 and EQUATION 13 and the column-specific parameters from TABLE 23. Blue Dextran 2000 was used to determine the void volume. Ferritin (444 kDa), Conalbumin (75 kDa), Carbonic anhydrase (29 kDa), Ribonuclease A (13.7 kDa), and Aprotinin (6.5 kDa) were used as protein standards, with all standards purchased from GE Healthcare (Chicago, USA).

EQUATION 12: $K_{av} = (V_E - V_0) \cdot (V_C - V_0)^{-1}$

K_{av}	partition coefficient, -
V_E	elution volume, mL
V_0	column void column, mL
V_C	column volume, mL

EQUATION 13: $\log(MW) = (K_{av} - C) \cdot M^{-1}$

MW	molecular weight, Da
C	constant from linear regression, -
M	slope from liner regression, $(\log(\text{Da}))^{-1}$

TABLE 23: Parameters used for size estimation from SEC

Column	V_C	V_0	C	M
HiLoad Superdex 200 pg	120.6 mL	46.1 mL	2.08	-0.336
Superdex 200 Increase 10/300 GL	23.56 mL	8.70 mL	2.13	-0.355

2.4.3. SDS-PAGE

Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (-PAGE) was performed to visualize the proteins and to check their purity and molecular weight. Protein samples were mixed with Laemmli buffer in a ratio of 1:2 and boiled at 95 °C for 5 min. Electrophoretic separation was then carried out at 50 mA for 30 min. The composition of the stacking gel, separation gel, running buffer, and Laemmli buffer is shown in TABLE 24. After electrophoretic separation, gels were stained with PageBlue™ Protein Staining Solution (Thermo Fisher Scientific / Waltham, USA) according to the manufacturer's protocol. Either Dual Color Protein Standard III or ROTI®Mark TRICOLOR was used as the standard marker for protein molecular weight.

TABLE 24: List of used buffers for SDS gel electrophoresis

The composition of the stacking gel, separation gel, running buffer, and Laemmli buffer used for SDS gel electrophoresis is shown below. *BME was used only for reducing gels. (1) Ammonium persulfate (APS), (2) tetramethylethylenediamine (TEMED).

Buffer	**Component**		**Buffer**	**Component**	
Stacking Gel	Acrylamide	6 wt%	**Separation Gel**	Acrylamide	15 wt%
	Tris, pH 6.8	125 mM		Tris, pH 8.8	375 mM
	SDS	0.1 wt%		SDS	0.1 wt%
	APS (1)	0.1 wt%		APS	0.1 wt%
	TEMED (2)	0.1 wt%		TEMED	0.1 wt%
	Bromophenol blue for staining				
Running Buffer	Glycine	192 mM	**Laemmli Buffer, 2x**	Tris, pH 6.8	120 mM
	Tris	25 mM		SDS	4 wt%
	SDS	0.1 wt%		Glycerol	20 vol%
				Bromophenol blue	0.02 wt%
				*BME	10 vol%

2.4.4. Circular dichroism spectroscopy

Circular dichroism (CD) spectra and thermal melting studies were recorded using a CD spectrometer J-1500. Far-UV CD spectra were recorded according to the instrument settings in TABLE 25 using a cuvette with a path length of 1 mm. Thermal melting studies were conducted by monitoring the CD signal at 215 nm during heating from 20 °C to 100 °C with a heating rate of 1 $°C\cdot min^{-1}$, a bandwidth of 5 nm, a digital integration time of 2 s, and a data interval of 0.5 °C. Spectra Manager^{TM} Version 2.15.01 was used to derive thermal melting points. Unless otherwise specified, the used buffer contained 10 mM NaPi pH 7.0, 0.5 mM DTT, and 0.1 wt% LDAO, with a protein concentration of 10 µM. Raw ellipticity data were converted to mean residue ellipticity [173], as shown in EQUATION 14.

TABLE 25: Instrument settings for far-UV CD spectroscopy measurements

Scanning Range	190–260 nm	Temperature:	20 °C
Bandwidth:	2 nm	Digital Integration Time:	1 s
Scanning Speed:	50 $nm\cdot min^{-1}$	Data Pitch:	0.1 nm
Scanning Mode:	Continuous	Number of Accumulations:	5

EQUATION 14:

$$\Theta_{mrw} = \Theta\cdot 10^{6}\cdot[d\cdot c\cdot(N - 1)]^{-1}$$

Θ_{mrw}	mean residue weight ellipticity, $deg\cdot cm^{2}\cdot dmol^{-1}$
Θ	ellipticity, mdeg
d	path length of the cuvette, mm
c	protein concentration, µM
N	number of amino acids in the protein, -

2.4.5. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were performed using a MicroCal PEAQ-ITC, with instrument settings listed in TABLE 26. Data analysis, curve fitting, and determination of binding affinities were done using MicroCal PEAQ-ITC Analysis Software Version 1.21. Unless otherwise specified, the used buffer containing 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 50 mM LDAO.

For the titrations with the metabolites, the OEP21-WT concentration in the cell was 75 µM. For ATP, CTP, GTP, and UTP, a concentration of 1.5 mM was used in the syringe, while the concentration was 2.5 mM for ADP, 2 mM for GDP, and 5 mM for AMP.

To determine K_D values for the OEP21 variants, 75 µM of the respective OEP21 variant was in the cell, while 1.2 mM ATP was in the syringe.

TABLE 26: Instrument settings for ITC measurements

Three different setups were used for injection, differing in the number of injections, injection volume, and injection duration.

Temperature:	25 °C	Reference Power:	10 µcal·s^{-1}	Stir Speed:	750 rpm
Initial Delay:	60 s	Injection Spacing:	150 s	Feedback:	High
13 Injections: 1·0.4 µL, 12·3 µL		19 Injections: 1·0.4 µL, 18·2 µL		25 Injections: 1·0.4 µL, 24·1.5 µL	
Injection Duration:	6 s	Injection Duration:	4 s	Injection Duration:	3 s

2.4.6. Fluorescence polarization spectroscopy

Fluorescence polarization (FP) assays were performed using a spectrofluorometer FP-8300 equipped with an MCB-100mini water circulation bath. Samples were transferred into a quartz cuvette with a path-length of 1 cm, and the FP of fluorescently labeled 2'-(or-3')-*O*-(*N*-Methylanthraniloyl) (MANT)-ATP was measured. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 0.1 % LDAO, while a buffer without EDTA was used in the presence of $MgCl_2$. The instrument settings for the FP measurements are listed in TABLE 27.

TABLE 27: Instrument settings for FP measurements

Excitation:	356 nm	Excitation Bandwidth:	5 nm	Duration:	30 s
Emission:	448 nm	Emission Bandwidth:	5 nm	Data Interval:	1 s
Temperature:	25 °C	Sensitivity:	High	Response:	1 s

To determine the binding affinities of MANT-ATP for the OEP21 variants, stepwise increasing protein concentrations were titrated into 200 nM MANT-ATP dissolved in buffer. Depending on the binding affinity of the OEP21 variants, the final protein concentration ranged from 500 nM to 800 nM. K_D values were derived by fitting the data to a one-site binding model using OriginPro®.

Using the same workflow and adjusted final protein concentrations of OEP21-WT, binding affinities for MANT-ATP were determined in the presence of different concentrations of NaCl (0, 0.5, 1, 2 M) and $MgCl_2$ (0, 1, 10, 50 mM).

To determine IC_{50} values for competitive binding between MANT-ATP and NaPi, GAP, or GAP plus $MgCl_2$ (2:1 concentration ratio), FP of a preformed complex of 1 µM MANT-ATP and 1 µM OEP21 (WT or ΔLoop5) was measured in the presence of increasing concentrations of the respective competitor. All FP experiments were performed as triplicates, except for the competitive binding experiment with GAP, which was performed as duplicate. IC_{50} values were derived by fitting the data to sigmoidal dose-response model using OriginPro®.

In addition, binding affinities for fluorescently labeled MANT-GDP and MANT-5'-guanylyl imidodiphosphate (GMPPNP), were determined analogously to MANT-ATP. All fluorescently labeled MANT nucleotides were purchased from Thermo Fisher Scientific (Waltham, USA).

2.4.7. CD spectroscopy with Bna-Tic40

Far-UV CD spectra were recorded with 10 µM BnaA10 or BnaC9 in 10 mM NaPi pH 6.0, 20 mM NaCl, 0.5 mM EDTA, and 2 mM β-mercaptoethanol on a CD spectrometer J-1500 in cuvette with a pathlength of 1 mm. The instrument settings are listed in TABLE 25, except that the bandwidth was 1 nm, and the scanning speed was 100 $nm \cdot min^{-1}$. The raw ellipticity data were converted to mean residue ellipticity [173], with BnaA10 and BnaC9 has a calculated mean residue weight of 116.1 $g \cdot mol^{-1}$ and 116.8 $g \cdot mol^{-1}$, respectively. In addition, the ellipticity for both samples was recorded during heating up from 20 °C to 100 °C at a wavelength of 222 nm with a bandwidth of 5 nm, a heating rate of 1 $°C \cdot min^{-1}$, a digital integration time of 2 s, and a data interval of 0.5 °C.

For titrations with 2,2,2-Trifluoroethanol (TFE), separate samples were prepared with 20 µM BnaA10 protein containing 0, 5, 10, or 20 vol% TFE. Deviating from TABLE 25, the band width was 1 nm, the digital integration time was 0.5 s, and the number of accumulations were ten.

The fractional helicity of a protein was estimated from the ellipticity at 222 nm, with a minimum at 222 nm and other nearby wavelengths characteristic of a helical structure [174]. The fractional helicity was calculated according to EQUATION 15.

EQUATION 15: $$FH = (\Theta_\lambda - \Theta_\lambda^0)\cdot(\Theta_\lambda^{100} - \Theta_\lambda^0)^{-1}$$

FH	fractional helicity, -
Θ_λ	mean residue weight ellipticity at wavelength λ, deg·cm^2·dmol^{-1}
Θ_λ^0	ellipticity for a protein with 0 % helical content at wavelength λ: -3,000 deg·cm^2·dmol^{-1} at 222 nm
Θ_λ^{100}	ellipticity for a protein with 100 % helical at wavelength λ: -39,500 deg·cm^2·dmol^{-1} at 222 nm

2.5. NMR spectroscopy

2.5.1. General

All NMR experiments were recorded in shigemi tubes with a volume of typically 300 µL and 7 vol% D_2O on Bruker Avance III instruments equipped with cryogenic probes and operating at 600 MHz, 800 MHz, 900 MHz, or 950 MHz proton frequency (Bruker Corporation / Billerica, USA). The acquisition and processing of data was performed using TopSpin® Version 3.2. Data analysis and resonance assignment was performed using NMRFAM-SPARKY Version 1.470 powered by Sparky Version 3.190.

2.5.2. OEP21

Various 2D-[^{15}N,^{1}H]-TROSY NMR spectra of ^{2}H,^{15}N-labelled OEP21 were typically recorded in 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 7 vol% D_2O at 303 K. 2D-[^{15}N,^{1}H]-TROSY NMR spectra were recorded in the presence of different LDAO concentrations or of stepwise titrated ATP, GAP, or phosphate. More detailed information on samples preparation and measurement conditions can be found in the respective figure legend. The LDAO concentration in the samples was estimated from 1D-[^{1}H]-NMR experiments using the integral areas from the water signal and the lower intensity triplet signal from the methyl group of the dodecyl alkyl tail.

2.5.3. Bna-Tic40

Initial 2D-[^{15}N,^{1}H]-TROSY NMR spectra of BnaA10 and BnaC9 were recorded on an 800 MHz spectrometer at different temperatures (303 K, 313 K, 323 K) and at different pH values (6.0, 7.0, 8.0) with 200 µM ^{15}N-labelled protein in 20 mM NaPi, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, and 7 vol% D_2O with 40 scans per increment and 256 complex points in the indirect ^{15}N dimension.

In presence of TFE, further 2D-[^{15}N,^{1}H]-TROSY NMR spectra of BnaA10 and BnaC9 were recorded on a 600 MHz spectrometer with 400 µM ^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, 7 vol% D_2O, and 10 vol% TFE at 303 K with 32 scans per increment and 256 complex points in the indirect ^{15}N dimension.

All NMR structure determination experiments were conducted with a 1.1 mM ^{13}C,^{15}N-labelled BnaA10 sample in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, 7 vol% D_2O, and 10 vol% d_3-trifluoroethanol at 308 K. As a starting point, a 2D-[^{15}N,^{1}H]-HSQC NMR spectrum with 8 scans per increment and 256 complex points in the indirect ^{15}N dimension was recorded on a 600 MHz spectrometer. For backbone resonance assignment of BnaA10, a set of 3D triple resonance experiments [129] were recorded on 600 MHz spectrometer, which consisted of HNCO [130-132], HN(CA)CO [133-135], HNCA [130, 131, 136-138], HN(CO)CA [131, 136, 137, 139, 140], HNCACB [141], and HN(CO)CACB [142]. The side chain amide resonances of asparagine and glutamine were assigned using the HN(CO)CACB experiment. Further side chain resonances were assigned using ^{13}C-edited and ^{15}N-edited CC(CO)NH [144] and H(CCO)NH [144, 145] experiments. HNHA [146] and HNHB [147] spectra were recorded to obtain information about the H_α and H_β chemical shift of the preceding residue, respectively. A [^{13}C,^{1}H]-HSQC spectrum with 8 scans per increment and 320 complex points in the indirect ^{13}C dimension was used to visualize all carbon side chain resonances and allowed further assignment in case of ambiguity. All of the above experiments were recorded on a 600 MHz spectrometer.

The protein secondary structure was estimated using the chemical shifts obtained from the backbone resonance assignment, as shown in EQUATION 6. NOE contacts were derived from a set of 3D-NOESY spectra consisting of HNH-NOESY [125], NNH-NOESY [126], CNH-NOESY [127], HCH-NOESY [128] , and CCH-NOESY [127] experiments. All 3D-NOESY experiments were recorded on a 900 MHz spectrometer. The diastereotopic H_β protons were assigned by a combined approach where the ^{3}J coupling constant between the backbone ^{15}N and the two β-methylene protons taken from an HNHB experiment and the NOE pattern observed in 3D-^{15}N and ^{13}C edited NOESY experiments were used, as shown in TABLE 5. In cases where diastereotopic assignment was not possible, pseudoatoms were introduced to replace the two protons. The diastereotopic side chain amide protons of asparagine and glutamine were assigned with the proton chemical shifts of $H_{\delta1}$ / $H_{\varepsilon1}$ (E-configuration) located downfield and $H_{\delta2}$ / $H_{\varepsilon2}$ (Z-configuration) located upfield.

3. Results

3.1. OEP21

3.1.1. Protein production and purification

Initial experiments in the field of OEP21 were performed by Dr. Elisabeth Häusler [170]. She screened for suitable conditions for expression of OEP21 in *E. coli* as inclusion bodies, the purification of OEP21 under denaturing conditions, and the refolding of OEP21 into LDAO detergent micelles. In addition, she performed biophysical and NMR spectroscopy experiments with refolded OEP21.

To perform further experiments for characterization of OEP21, protein first had to be produced and purified. OEP21 was expressed as inclusion bodies in *E. coli* BL21(DE3), as exemplified in FIGURE 15. After induction with IPTG, a protein band appeared on the SDS-PAGE gel with molecular weight in good agreement with the theoretical one of OEP21-WT (22,058 g·mol^{-1}).

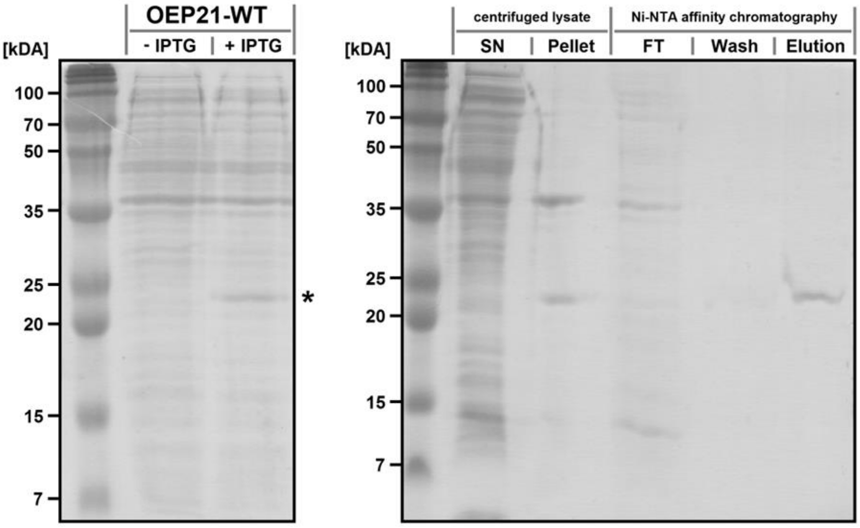


FIGURE 15: Expression and purification OEP21-WT

On the left, a Coomassie-stained SDS-PAGE gel shows protein expression of OEP21-WT at 37 °C by *E. coli* BL21(DE3) in LB medium before induction (-) and 4 h after induction with 1 mM IPTG (+). The asterisk indicates a protein band that came up during expression and represented OEP21-WT. Purification of His$_{10}$-tagged OEP21-WT via Ni-NTA affinity chromatography is shown on the right. Prior to this, the cell lysate was centrifuged to separate insoluble OEP21-WT in the pellet fraction from the supernatant (SN). Subsequently, the pellet fraction was dissolved in a buffer containing 6 M GdmCl before purifying by Ni-NTA affinity chromatography under denaturing conditions. In the first step, the applied protein sample was allowed to flow through the column (FT) after incubating with the Ni-NTA slurry for 1 h. The column was then washed with buffer before OEP21-WT was eluted from the column with a buffer containing 500 mM imidazole.

Purification of His$_{10}$-tagged OEP21-WT from inclusion bodies via Ni-NTA affinity chromatography under denaturing conditions with GdmCl yielded a large amount of highly pure OEP21-WT after elution with imidazole, as shown exemplarily in FIGURE 15. Dialysis was performed to remove GdmCl from the samples, resulting in precipitation of OEP21-WT. After separation of protein pellet and supernatant by centrifugation, a yield of around 40 mg unfolded OEP21-WT per liter LB medium was usually obtained. Subsequently, the protein pellet was dissolved highly concentrated in a buffer containing GdmCl, followed by refolding in a buffer containing LDAO as membrane mimetic and no GdmCl. During refolding, precipitation of OEP21-WT was observed, resulting in a protein loss of approximately 25 % after refolding and concentrating the protein for SEC. OEP21-WT was then further purified by SEC, with a typical SEC chromatogram of OEP21-WT shown in FIGURE 16.

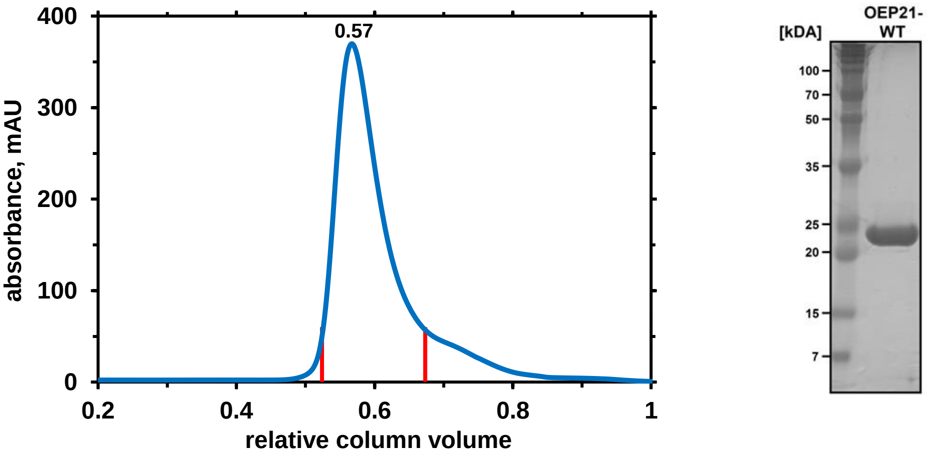


FIGURE 16: SEC chromatogram of refolded OEP21-WT

On the left, refolded OEP21-WT was applied onto a HiLoad Superdex 200 pg SEC column. The running buffer consisted of 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 0.1 wt% LDAO. The main peak eluted at 0.57 CV, as indicated by the maximum absorbance measured at 280 nm (blue curve). The calculated molecular weight of the OEP21-WT/LDAO complex at 0.57 CV was around 200 kDa. The interval between the red lines indicates all SEC fractions that were pooled for further use. On the right, the purity of these pooled OEP21-WT fractions is verified by a Coomassie-stained SDS-PAGE gel.

Since the main peak eluted at 0.57 CV, a theoretical size of about 200 kDa was calculated for the OEP21-WT/LDAO complex. It could be concluded that OEP21 formed higher oligomers, since LDAO has a detergent micelle size of 21.5 kDa ± 0.7 kDa [175] and the size of OEP21-WT is 22.1 kDa. Interestingly, the tendency of OEP21 to form oligomers could be modulated by the LDAO concentration, with a higher LDAO concentration allowing the appearance of smaller species of OEP21/LDAO besides higher oligomeric species.

This was observed as shown in FIGURE 17, where a sample with a relatively high LDAO-to-protein ratio was applied on SEC. As a result, various OEP21-WT/LDAO species were detectable, with all fractions containing pure OEP21-WT, as confirmed by SDS-PAGE (data not shown). The peak that eluted 0.67 CV was approximately 45 kDa in size and represented a smaller species of OEP21-WT/LDAO that was most likely the monomeric form. Interestingly, this monomeric species could be transformed back to a higher oligomeric species again by applying the monomer fractions onto another SEC, as shown in FIGURE 17, thus reducing the relative LDAO-to-protein ratio. Different LDAO-to-protein ratios were caused, for example, by differences in refolding efficiency or protein purity. After concentration of the main peak from SEC, a typical yield of around 15 mg OEP21-WT per liter LB medium was obtained.

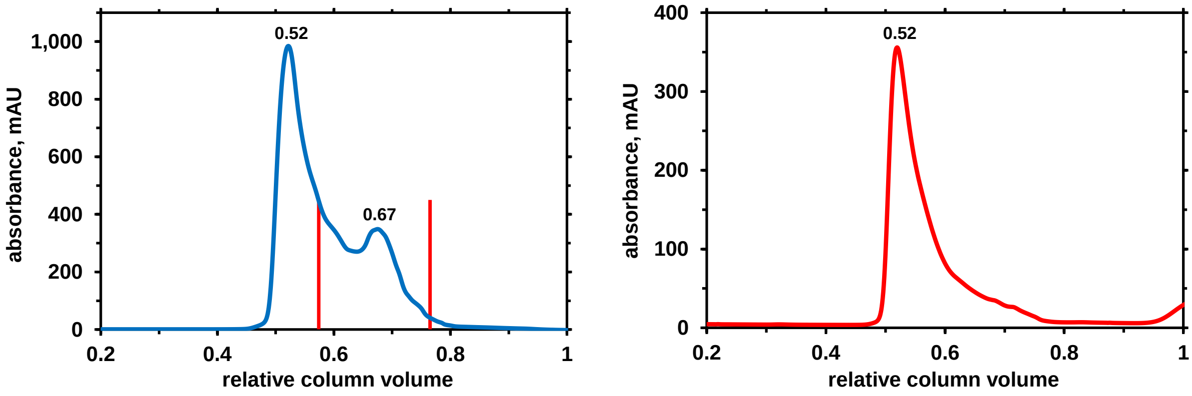


FIGURE 17: LDAO affected the oligomeric state of OEP21

On the left, refolded OEP21-WT with a relatively high LDAO-to-protein was applied onto a Superdex 200 Increase 10/300 GL SEC column. The main peak eluted at 0.52 CV, as shown by the absorbance measured at 280 nm (blue curve). The elution volume of the main peak was slightly different from that in FIGURE 16 (0.57 CV) because a different column type was used. When samples with a relatively high LDAO-to-protein ratio were applied, tailing of the main peak was observed as well as the appearance of additional peaks. Here, a second peak was eluting at 0.67 CV, corresponding to a theoretical molecular weight of around 45 kDa. The interval between the red lines indicates all SEC fractions that were applied onto another SEC. The chromatogram of this second SEC is shown on the right (red curve). Since the relative LDAO-to-protein ratio was reduced by this second SEC, the main peak eluting at 0.52 CV again indicated higher oligomeric species. The running buffer for both runs consisted of 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 0.1 wt% LDAO.

3.1.2. Characterization of OEP21 by CD spectroscopy

To gain insight about the secondary structure of OEP21, a far-UV spectrum of OEP21-WT in LDAO detergent micelles was recorded. The spectrum is shown in FIGURE 18 and indicates a β-sheet secondary structure, as a β-sheet secondary structure has a characteristic minimum at 215 nm and a maximum at 198 nm [174, 176].

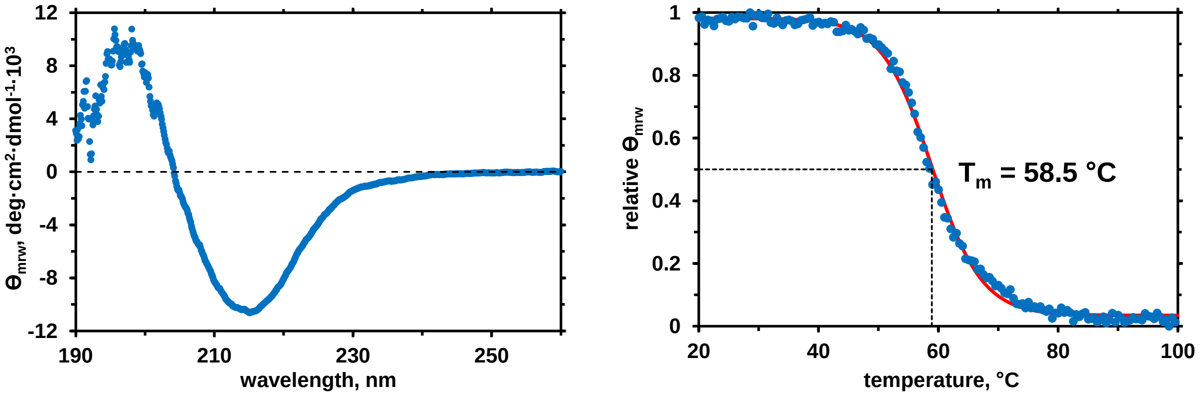


FIGURE 18: Far-UV CD spectrum and thermal melting of OEP21-WT

On the left, the mean residue weight ellipticity Θ_{mrw} of OEP21-WT in LDAO micelles at a wavelength from 190 nm to 260 nm at 20 °C is shown. On the right, the thermal melting behavior of OEP21-WT by heating up from 20 °C to 100 °C at a wavelength of 215 nm is shown, with the resulting Θ_{mrw} normalized. The thermal melting point T_m of OEP21-WT was 58.5 °C. Both spectra were recorded with 10 µM protein in 10 mM NaPi pH 7.0, 0.5 mM DTT, and 0.1 wt% LDAO.

Indeed, OEP21-WT exhibited a maximum at 198 nm and a minimum at 215 nm. In addition, a thermal unfolding curve of OEP21-WT was recorded, which revealed a clear sigmoidal transition corresponding to a thermal melting point of 58.5 °C, as shown in FIGURE 18. After heating up to 100 °C, precipitation of the protein occurred, indicating irreversible unfolding of OEP21-WT.

3.1.3. Structure determination by NMR spectroscopy

To better understand the function and relevance of OEP21, it was very helpful to obtain structural information about the protein. Therefore, OEP21 was structurally analyzed using NMR spectroscopy by Dr. Elisabeth Häusler [170] and Prof. Dr. Franz Hagn. The NMR chemical shift information and structural coordinates of OEP21 can be found at the BMRB [84] under the accession code 34589 and the RCSB PDB [85] under the accession code 7BGH, respectively.

To obtain a high-resolution structure, a uniformly ^2H,^{15}N and selectively methyl-group labeled (Ile-δ_1, Leu-δ_2, Val-γ_2, Ala-β) as well as ^1H,^{15}N-labeled phenylalanine- and tyrosine-containing sample in complex with the stabilizing ligand ATP was used. Secondary ^{13}C chemical shift information indicated twelve β-strand regions of varying lengths, as shown in FIGURE 19.

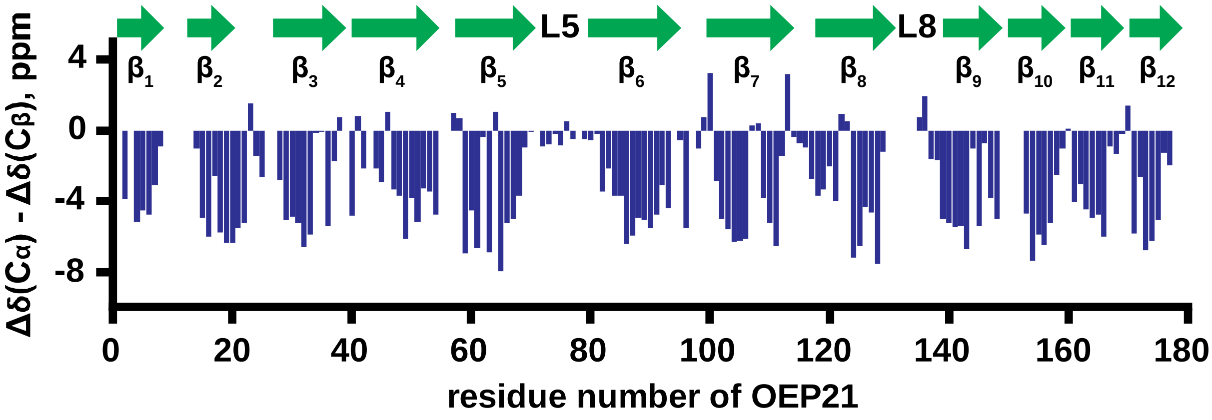


Figure 19: Secondary structure estimation of OEP21 by ^{13}C chemical shifts

The secondary structure of OEP21 was estimated using ^{13}C secondary chemical shift information obtained from the backbone resonance assignment. $\Delta\delta(C_\alpha) - \Delta\delta(C_\beta)$ was calculated according to Equation 6, where a positive value corresponds to an α-helical and a negative value to β-sheet secondary structure. Therefore, twelve β-strands regions of varying lengths were identified, visualized by the green boxes above. This figure was provided by Prof. Dr. Franz Hagn based on NMR experiments performed by Dr. Elisabeth Häusler [170].

In addition, a set of heteronuclear 3D-NOESY NMR experiments was used to assign side chain methyl resonances and extract NOE distance constrains for structure calculation. Finally, a well-defined structural bundle of OEP21 was obtained, as shown in Figure 20, which had a root mean square deviation in ordered secondary structure elements of 0.5 Å. This shows that OEP21 resembled a cone with a diameter of 22 Å on one side and 11 Å on the opposite side. Except for two longer loops, namely loop five and loop eight, located on either side of the β-barrel structure, most connecting loops were structurally well-defined. Loop five gave rise to strong NMR signals and showed no NOE contacts with other parts of the protein, indicating an intrinsic flexibility. Therefore, the solvent accessibility of loop five was probed by NMR paramagnetic relaxation enhancement experiments using a spin-labeled fatty acid, which resulted in selective signal attenuation of membrane-incorporated parts of the protein but left loop five almost unaffected (data not shown). Interestingly, the determined structure reveals that the pore has a highly positively charged inner surface to which negatively charged metabolites could bind, as shown in Figure 20.

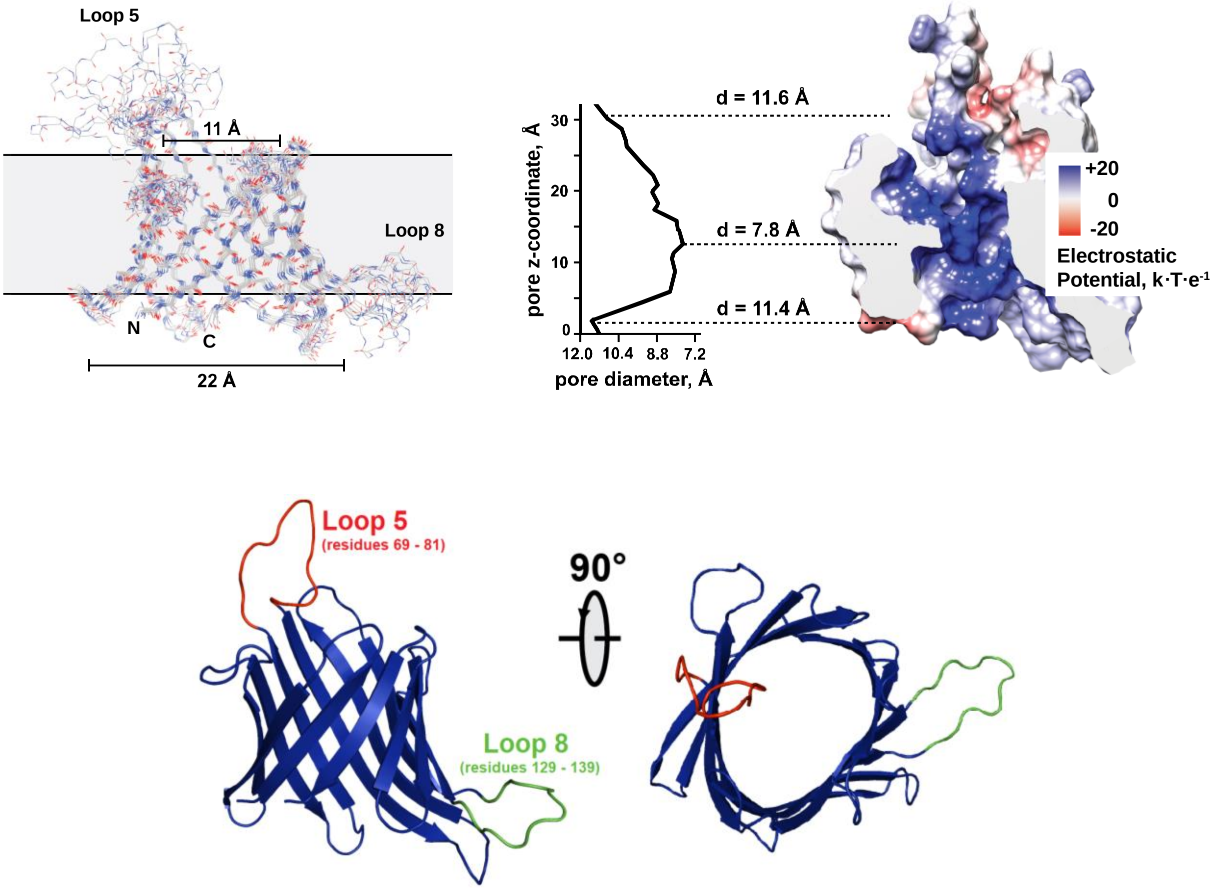


FIGURE 20: NMR structure of OEP21

The NMR structure of OEP21 was determined by Dr. Elisabeth Häusler [170] and Prof. Dr. Franz Hagn. The NMR chemical shift information and structural coordinates of OEP21 can be found at the BMRB [84] under the accession code 34589 and the RCSB PDB [85] under the accession code 7BGH, respectively. In the upper left, the NMR structural bundle of the OEP21 β-barrel pore shows well-defined secondary structure elements with a root mean square deviation in ordered secondary structure elements of 0.5 Å. Thereby, OEP21 resembled a cone with a diameter of 22 Å on one side and 11 Å on the opposite side. Except for two longer loops, namely loop five and loop eight, located on either side of the β barrel structure, most connecting loops were structurally well-defined. In the upper right, an analysis of the pore geometry indicated a 7.8 Å wide constriction site, with the pore interior having a high abundance of positive charges, suggesting binding to negatively charged metabolites. The schematic representation of the structure of OEP21 in the lower panel was created using PyMOL. The two upper figures were provided by Prof. Dr. Franz Hagn.

3.1.4. Determination of the orientation and oligomerization

As seen by the NMR structure, OEP21 exhibits a cone-shaped structure indicative of directional metabolite flow. Therefore, the orientation of OEP21 in the chloroplast outer envelope membrane was determined by limited proteolysis experiments with isolated right-side-out outer envelope vesicles [177, 178], resulting in two distinct proteolytic fragments of around 10 kDa and 14 kDa after treatment with thermolysin according to Bölter *et al.* (1999) [43]. The protease thermolysin exhibits specificity against hydrophobic amino acid residues such as L-leucine, L-isoleucine, and L-phenylalanine [179]. To identify the cleavage site in OEP21, the larger fragment of trypsin-treated OEP21 was subjected to N-terminal Edman sequencing, and the identified peptide (YAKNDK) was assigned to loop five. Trypsin itself is known to strongly restrict its action to the basic amino acids lysine and arginine [180]. This indicated that loop five is oriented toward the cytosol and that the N-termini and C-termini are both located in the intermembrane space, as illustrated in Figure 21.

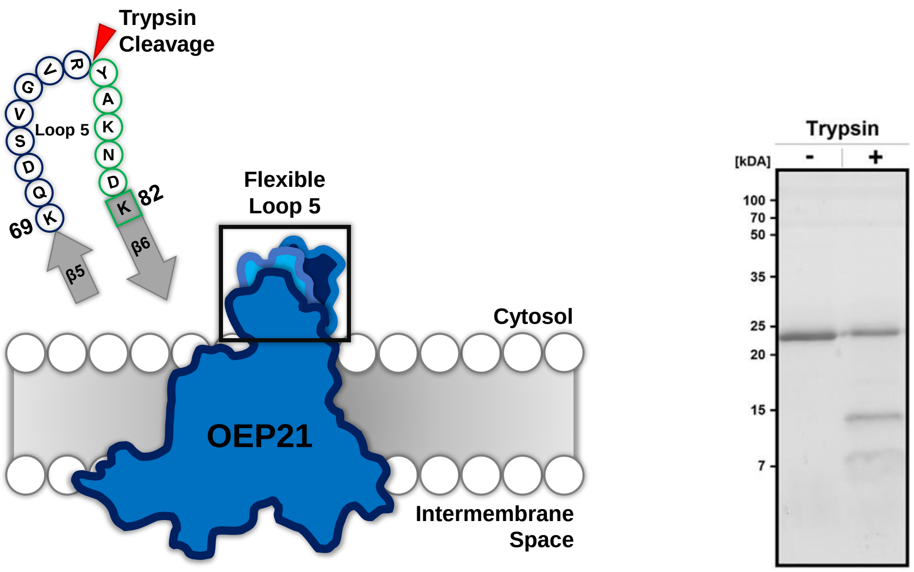


Figure 21: Orientation of OEP21 in the chloroplast outer envelope

The larger cleavage fragment of trypsin-treated outer envelope vesicles was subjected to N-terminal Edman sequencing, and the identified peptide (YAKNDK) was assigned to loop five. This indicated that loop five is oriented toward the cytosol and that the N-termini and C-termini are both located in the intermembrane space, as shown schematically on the left. On the right, recombinant OEP21-WT reconstituted in liposomes was treated (+) or not treated (-) with 0.1 µM trypsin for 15 min. The reaction was quenched by addition of 1 mM PMSF before being analyzed by Coomassie-stained SDS-PAGE. Liposomes were prepared with 10 $g\cdot L^{-1}$ soybean polar lipid extract (Croda International Plc / Snaith, GBR) solubilized in 20 mM HEPES pH 7.5, 250 mM KCl, and 1 wt% LDAO. 10 µM OEP21-WT was added to the liposomes and the detergent was removed using Bio-Beads SM-2 Resin (Bio-Rad Laboratories, Inc. / Hercules, USA). This was followed by three freeze-thaw cycles and fifteen passes through a 0.2-µm-filter using a mini-extruder. These data were provided by Dr. Umut Günsel.

Interestingly, trypsin treatment of recombinantly produced OEP21 reconstituted in liposomes showed an identical cleavage pattern, indicating a native conformation of refolded OEP21, as shown in FIGURE 21. However, the presence of ATP or GAP slightly protected loop five from proteolytic cleavage, suggesting an interaction with loop five, as shown in FIGURE 22. Since the orientation of OEP21 in the liposome membrane was not controlled, the degree of cleavage was only approximately 50 %. The addition of detergent and the consequent solubilization of the liposomes, which allowed complete exposure of OEP21, resulted in a cleavage degree of approximately 100 %, as shown in FIGURE 22.

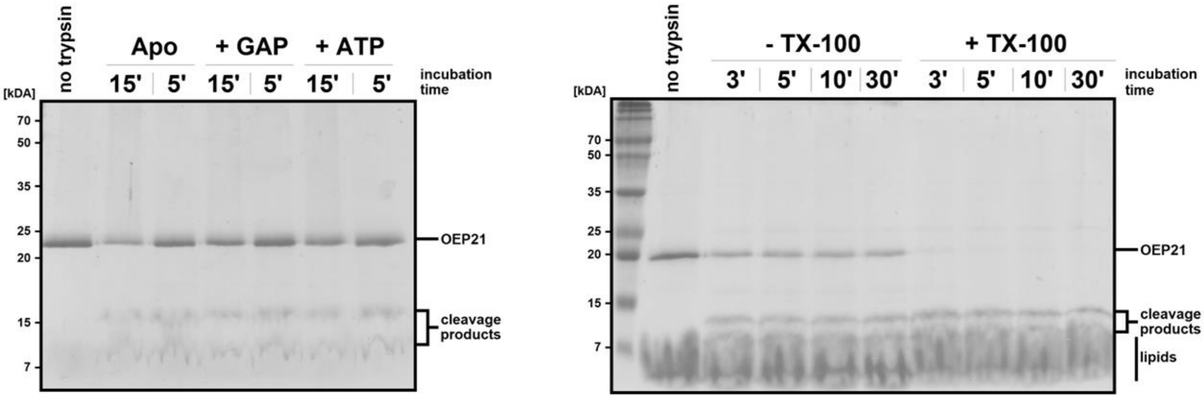


FIGURE 22: Trypsin digestion experiments with OEP21 reconstituted in liposomes

On the left, trypsin digestion experiments were performed with 10 µM OEP21-WT reconstituted in liposomes prepared according to FIGURE 21 in the apo form or in the presence of 5 mM GAP or 5 mM ATP. After 5 min and 15 min incubation with 0.1 µM trypsin, samples were taken for analysis by Coomassie-stained SDS-PAGE, where the reaction was stopped by addition of 1 mM PMSF. OEP21-WT without trypsin served as a control. Since the orientation of OEP21-WT in liposomes was not controlled, only half of OEP21-WT was assumed to expose loop five for trypsin cleavage, with trypsin added from the outside, resulting in approximately 50 % reduction in band intensity of OEP21-WT. On the right, addition of 0.2 vol% Triton X-100 dissolved the liposomes and allowed complete cleavage of OEP21-WT by trypsin, using 0.4 µM trypsin. These data were provided by Dr. Umut Günsel.

As previously shown by chemical crosslinking, OEP21 forms larger oligomers in the outer envelope [43]. Further experiments performed by Dr. Bettina Bölter using blue native PAGE of extracted outer envelope vesicles revealed that OEP21 oligomers were detectable in a molecular weight range from around 40 kDa to 200 kDa and possibly represented dimers in groups from one to five (data not shown). Recombinant OEP21 reconstituted in liposomes or in LDAO also showed the appearance of higher oligomers in chemical crosslinking experiments, as shown in FIGURE 23, while the formation was not altered by addition of ATP or GAP. Interestingly, the tendency of OEP21 to form oligomers could be influenced by the LDAO concentration, with the oligomer population decreasing with increasing LDAO.

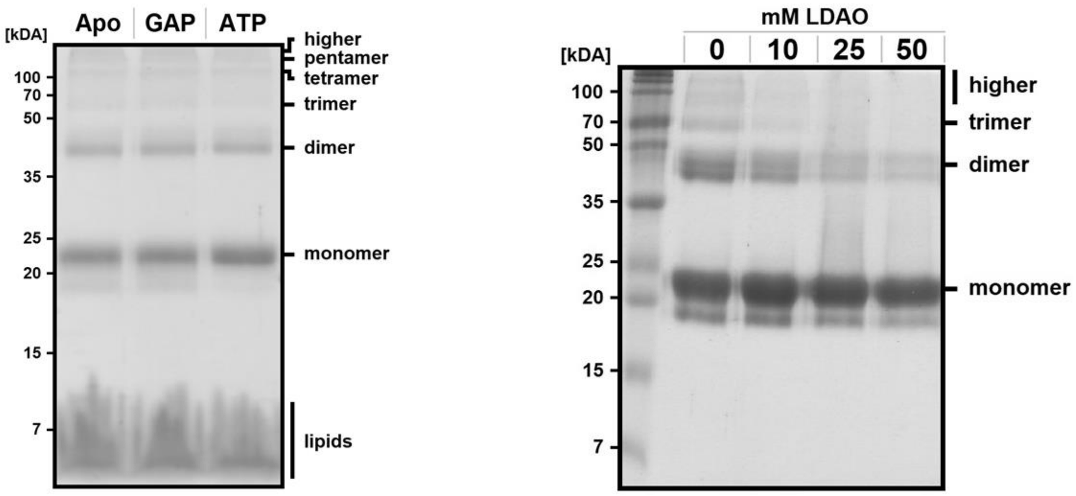


FIGURE 23: Oligomerization of OEP21

On the left, a Coomassie-stained SDS-PAGE gel shows chemical crosslinking of OEP21-WT reconstituted in liposomes either in the apo form or in the presence of GAP or ATP, with stable oligomerization indicated in each case. Liposomes were prepared according to FIGURE 21 in the absence or in the presence of 5 mM ATP or 5 mM GAP. The amino-selective crosslinker bis(sulfosuccinimidyl)suberate (Thermo Fisher Scientific / Waltham, USA) at a concentration of 0.5 mM was applied to the liposomes and incubated at room temperature for 30 min before quenching the reaction by adding 50 mM Tris pH 7.5 and incubating for 15 min. These data were provided by Dr. Umut Günsel. On the right, a Coomassie-stained SDS-PAGE gel shows chemical crosslinking of OEP21-WT in the presence of different LDAO concentrations (with additional 0 mM, 10 mM, 25 mM, or 50 mM LDAO), where the oligomer population decreased with increasing LDAO concentration. 0.5 mM bis(sulfosuccinimidyl)suberate was applied to 50 µM OEP21-WT in 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 0.1 wt% LDAO. After incubation at room temperature for 30 min, the reaction was quenched by adding 50 mM Tris pH 8.0 and incubating for 15 min.

2D-[^{15}N,^{1}H]-TROSY NMR experiments were recorded at high and low detergent concentrations to obtain further insight into the oligomerization behavior of OEP21, as shown in FIGURE 24. At a low detergent concentration, there was a strong line broadening of the backbone amide resonances in the folded β-barrel, indicating the formation of higher oligomers. In addition, loop five appeared to be affected by oligomerization, as strong NMR signals originated from it. Strong NMR signals were also present in the random coil chemical shift region, which could be caused by a slight compression of the β-barrel. Interestingly, the addition of ATP improved the quality of the NMR spectrum and caused the disappearance of peaks in the random coil chemical shift region, suggesting that ATP could stabilize the β-barrel of OEP21 by binding.

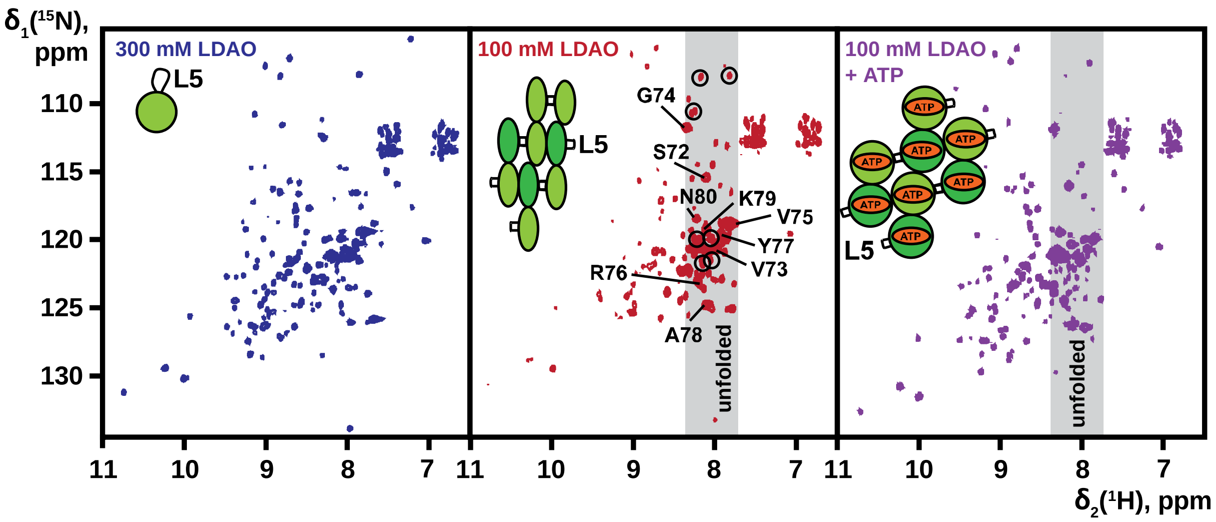


FIGURE 24: NMR experiments recorded at high and low detergent concentrations

2D-[^{15}N,^{1}H]-TROSY NMR spectra of OEP21-WT were recorded in the presence of different LDAO concentrations (approximately 100 mM or 300 mM) and in the presence of 2 mM ATP at 100 mM LDAO. All spectra were recorded on a 950 MHz spectrometer with 350 µM ^{2}H,^{15}N-labelled OEP21-WT in 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 7 vol% D_2O with 48 scans per increment and 192 complex points in the indirect ^{15}N dimension at 303 K. At 100 mM LDAO, there was a strong line broadening of the backbone amide resonances in the folded β-barrel, indicating the formation of higher oligomers. In addition, loop five appeared to be affected by oligomerization, as strong NMR signals originated from it. Strong NMR signals were also present in the random coil chemical shift region, which could be caused by a slight compression of the β-barrel, as shown schematically. Interestingly, the addition of 2 mM ATP improved the quality of the NMR spectrum and caused the disappearance of peaks in the random coil chemical shift region, suggesting that ATP could stabilize the β-barrel of OEP21 by binding. The encircled NMR signals in the random coil region are visible only in the oligomeric apo form. This figure was provided by Prof. Dr. Franz Hagn.

As indicated by blue native PAGE experiments, OEP21 appeared to occur as dimer groups forming oligomers. Therefore, the dimerization interface of OEP21 was investigated by analyzing NMR chemical shift perturbations in 2D-[^{15}N,^{1}H]-TROSY experiments of ^{2}H,^{15}N-labeled OEP21 at medium and high detergent concentrations. The analysis showed that only one side of the β-barrel exhibited distinct chemical shift perturbations, as shown in FIGURE 25.

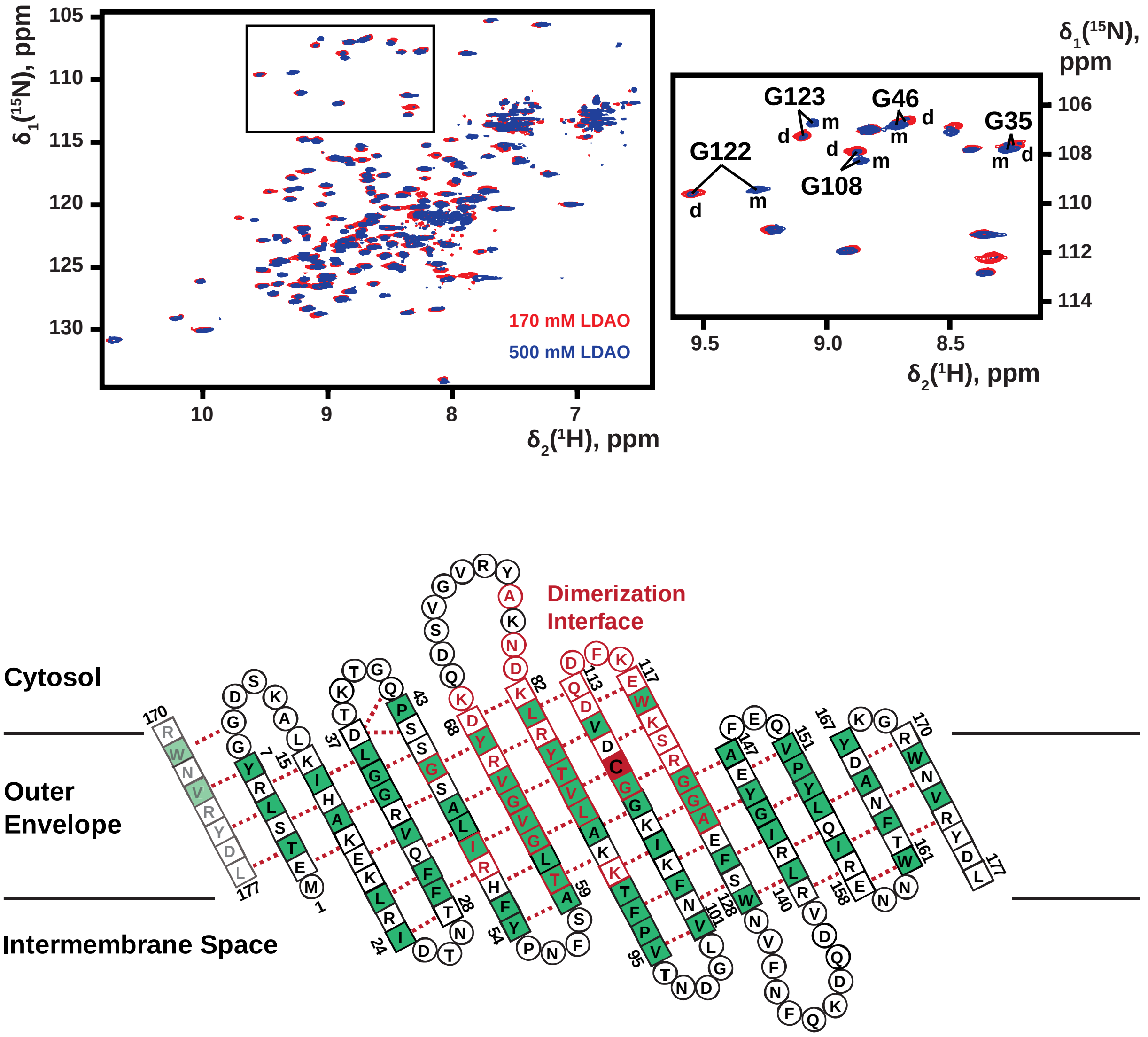


FIGURE 25: NMR chemical shift perturbations indicate dimerization of OEP21

Above, 2D-[^{15}N,^{1}H]-TROSY NMR spectra of OEP21-WT recorded in the presence of approximately 170 mM (red) and 500 mM LDAO (blue) showed large chemical shift perturbations. For the glycine region in the TROSY spectrum (see box on the right), the monomeric (m) and dimeric (d) species were labeled respectively, with a small fraction of OEP21 remaining in the dimeric state even at 500 mM LDAO. Below, the residues with pronounced chemical shift perturbations are marked in red in the topology plot, whereas hydrophobic amino acids are indicated by green boxes. The long β-strands on one side of the β-barrel located around cysteine 109 (shown as a red box) are strongly affected, suggesting that cysteine 109 may form a disulfide bridge between two monomers. Both figures were provided by Prof. Dr. Franz Hagn based on NMR experiments performed by Dr. Elisabeth Häusler [170].

Examination of the affected region reveals a cysteine residue (C109) in the center that is oriented toward the membrane and well-positioned to form a disulfide bridge under oxidizing conditions. The possibility of cysteine-mediated crosslinking was clearly demonstrated by analysis of a non-reducing SDS-PAGE with Cu^{2+}-oxidized OEP21, as shown in FIGURE 26.

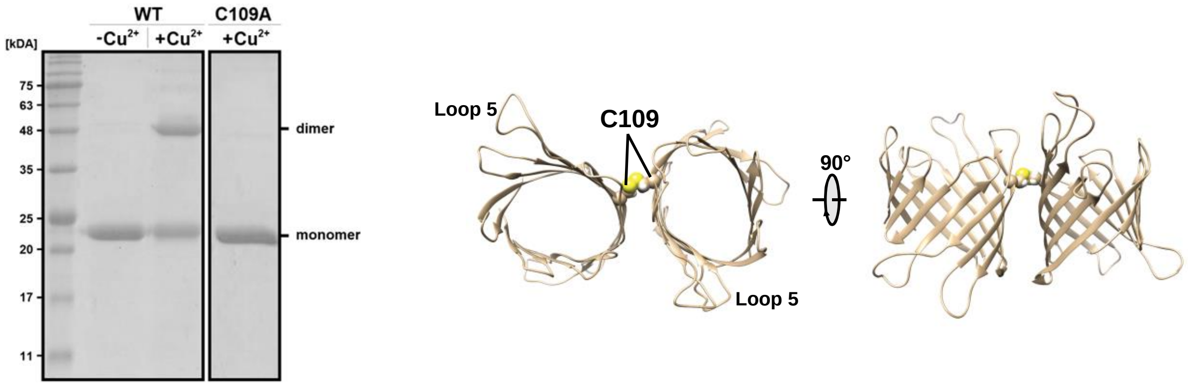


Figure 26: OEP21 dimer formation in an oxidizing environment

On the left, a Coomassie-stained non-reducing SDS-PAGE gel shows the formation of a OEP21-WT dimer via an inter-monomer disulfide bridge in an oxidizing environment. For this purpose, 20 µM protein in 20 mM HEPES pH 7.5, 50 mM NaCl, and 0.5 mM DTT was exposed reducing conditions or oxidizing conditions (with additional 1 mM $CuSO_4$) for 15 min at room temperature. No dimer formation was observed when cysteine 109 was mutated to alanine (C109A). These data were provided by Dr. Umut Günsel. On the right, a structural model of the disulfide-bridged dimer is shown, which could be the smallest unit of OEP21 present in the outer envelope membrane. This figure was provided by Prof. Dr. Franz Hagn.

Interestingly, the presence of the disulfide bridge could be verified by an analysis of the NMR $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts of cysteine 109, according to Martin *et al.* (2010) [181]. Based on these results, a dimeric, disulfide-bridged structural model of OEP21 was constructed, as shown in Figure 26, where the dimer could be the smallest unit present in the outer envelope membrane. The question now was what would happen if the cysteine at position 109 was replaced by alanine. As shown later in Figure 45, the mutation had no effect on the secondary structure of OEP21, while thermal stability was slightly reduced compared to the wildtype.

3.1.5. Effect of metabolites on OEP21 detected by CD spectroscopy

To further characterize OEP21, CD spectroscopy experiments were performed in the presence of various metabolites, as the inner surface of OEP21 has been shown to be highly positively charged, suggesting the binding of negatively charged metabolites. First, far-UV spectra of OEP21-WT were recorded in the presence of various metabolites, as shown in Figure 27. All spectra exhibited a minimum around 215 nm and a maximum around 198 nm indicating a β-sheet secondary structure. Because the aromatic groups of some metabolites impacted the absorbance, the spectra differed at the affected wavelengths. Therefore, the nucleoside-containing metabolites ATP, ADP, AMP, CTP, GTP, GDP, UTP, NADH, and NADPH were affected. The results of the CD-detected thermal melting experiments are summarized in Figure 27.

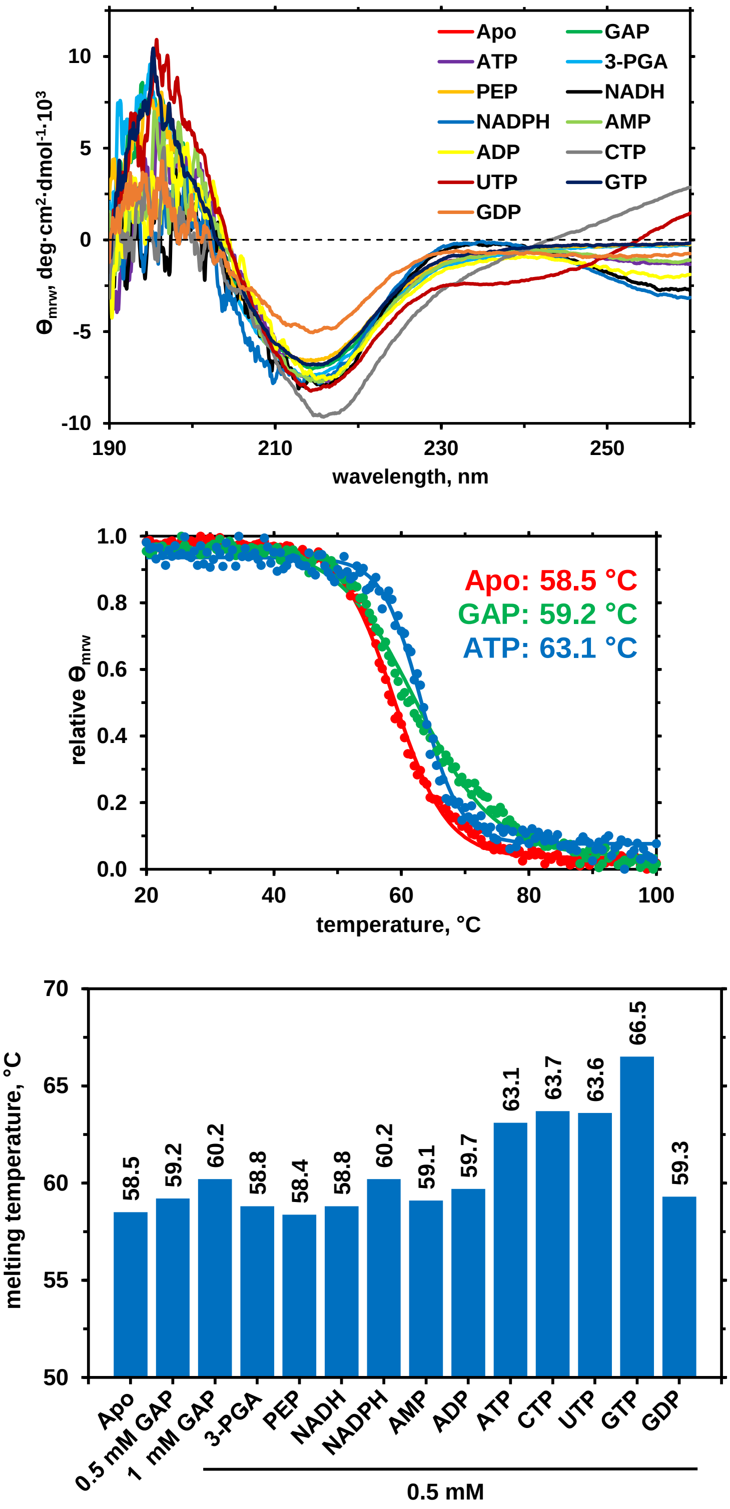


FIGURE 27: Thermal stability of OEP21-WT in the presence of various metabolites

Above, the mean residue weight ellipticity Θ_{mrw} of OEP21-WT in LDAO micelles at a wavelength of 190 nm to 260 nm at 20 °C is shown. The thermal melting behavior of OEP21-WT in the apo form (red), with 0.5 mM GAP (green), or with 0.5 mM ATP (blue), observed by heating up from 20 °C to 100 °C at a wavelength of 215 nm, is shown in the middle, with the resulting mean residue weight ellipticity Θ_{mrw} normalized. Below, the thermal melting points determined for various metabolites are summarized. All spectra were recorded with 10 µM protein in 10 mM NaPi pH 7.0, 0.5 mM DTT, 0.1 wt% LDAO, and 0.5 mM of the respective metabolite, except for GAP where additionally 1 mM was used. The reference (Apo) was recorded without any metabolite.

In summary, nucleoside triphosphates such as ATP, CTP, GTP, or UTP led to a significantly higher thermal melting point compared to the apo form, indicating increased stabilization of OEP21-WT. Interestingly, the stabilizing effect of these metabolites is much less pronounced when a nucleoside diphosphate, such as ADP or GDP, or the nucleoside monophosphate AMP was used. Similarly, stability was slightly increased when NADPH was used instead of NADH. Monophosphorylated metabolites such as GAP, 3-PGA, or PEP were also tested. In the presence of 3-PGA and PEP, the stability did not change significantly, whereas the presence of GAP increased stability. This stabilizing effect could be further increased by using 1 mM GAP instead of 0.5 mM. In conclusion, negatively charged metabolites can bind to OEP21, as evidenced by increased stability of OEP21, with the stabilizing effect depending on the number of charges in the metabolite.

3.1.6. Determination of the affinity of OEP21 for metabolites by ITC

As shown in the previous chapter, the presence of certain metabolites affected the thermal stability of OEP21, indicating binding of metabolites. Therefore, the binding affinities of various metabolites were investigated in more detail using ITC. First, the binding of ATP to OEP21 was probed, since ATP had a significant effect on thermal melting behavior and is present in plant cells at a concentration of about 1–2 mM [182]. It was also suggested that ATP acts as an inhibitor of triosephosphate transport across OEP21 [43]. FIGURE 28 shows a typical result of ATP titrated to OEP21-WT observed by ITC. As indicated by the negative spikes in the differential power, an exothermic reaction is taking place, suggesting binding of ATP to OEP21. Looking at the relative enthalpy plot, a clear sigmoidal transition can be seen, indicating saturation of OEP21 with ATP at an excess of ATP. Using the analysis software, a binding affinity with a K_D of 11.9 µM was derived for ATP, where the stoichiometry was approximately equimolar. As a control, ATP was titrated into buffer without protein as shown in FIGURE 29, which gave only minor positive spikes in the differential power, indicating an endothermic reaction.

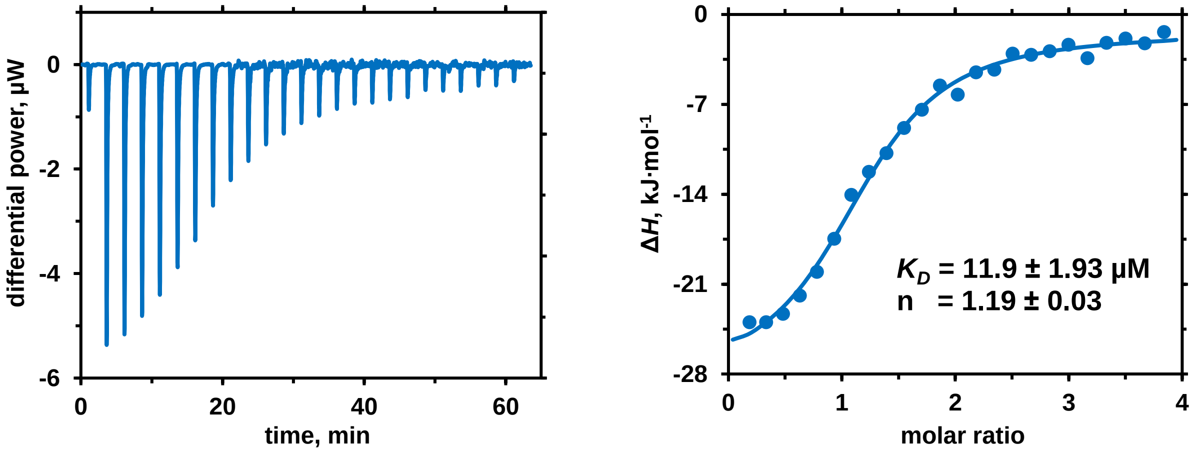


FIGURE 28: Probing the binding affinity of ATP to OEP21-WT by ITC

ATP at a concentration of 1.5 mM in the syringe was titrated into the cell containing 75 µM OEP21-WT in 25 injection steps, and the effect on differential power was measured at 25 °C. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 50 mM LDAO. On the right, the relative enthalpy change over the molar ratio (ATP:OEP21-WT) is shown, which was obtained from the raw data of the differential power on the left. For the binding affinity of OEP21-WT to ATP, a K_D of 11.9 µM ± 1.93 µM and a stoichiometry of 1.19 ± 0.03 were obtained.

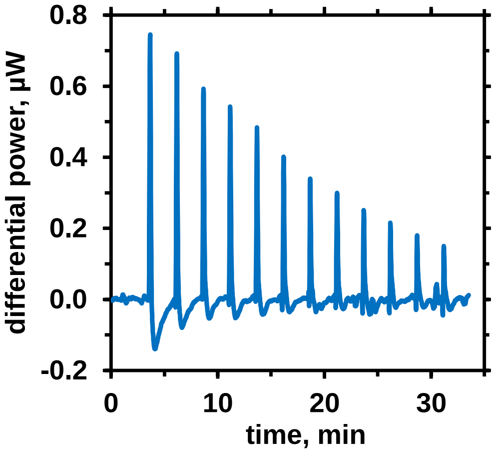


FIGURE 29: ITC control of ATP titrated into buffer

ATP at a concentration of 1.2 mM in the syringe was titrated into the cell containing buffer in 13 injection steps, and the effect on differential power was measured at 25 °C. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 50 mM LDAO.

Interestingly, a significantly different result was obtained at a low LDAO concentration of 0.1 wt% LDAO instead of 50 mM LDAO, as illustrated in FIGURE 30

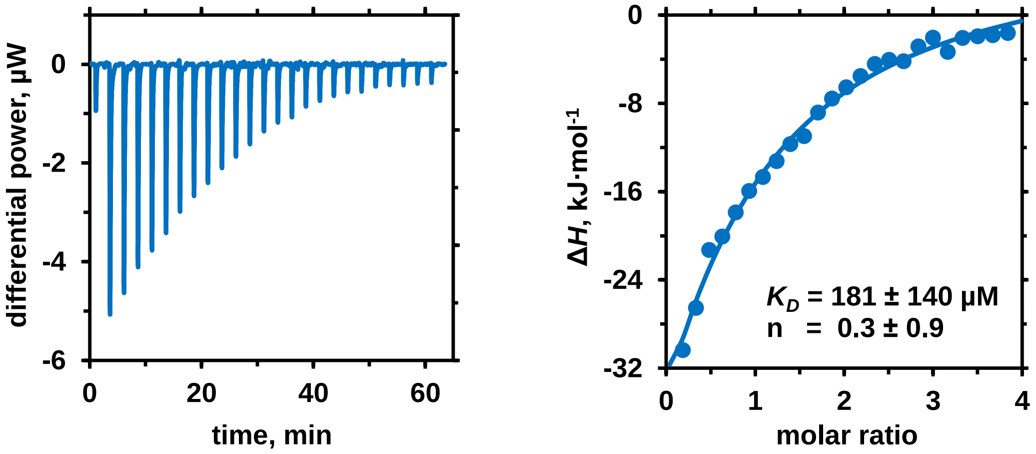


FIGURE 30: Effect of a low LDAO concentration on the binding affinity of ATP

ATP at a concentration of 1.5 mM in the syringe was titrated into the cell containing 75 µM OEP21-WT in 25 injection steps, and the effect on differential power was measured at 25 °C. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 0.1 wt% LDAO. On the right, the relative enthalpy change over the molar ratio (ATP:OEP21-WT) is shown, which was obtained from the raw data of the differential power on the left. For the binding affinity of OEP21-WT to ATP at a low LDAO concentration, a K_D of 181 µM ± 140 µM and a stoichiometry of 0.3 ± 0.9 were obtained. Obviously, the curve could only be insufficiently fitted by the ITC analysis software, as shown by the serious error deviations of K_D and stoichiometry.

In contrast to the results obtained with 50 mM LDAO, no sigmoidal transition in the relative enthalpy was observed at 0.1 wt% LDAO. Therefore, the ITC analysis software obviously had problems fitting this saturation curve without a clear inflection point, which was evident from the serious error deviations of K_D and stoichiometry. FIGURE 31 summarizes the effect of different LDAO concentrations on the K_D and stoichiometry obtained by ITC, where higher LDAO concentrations resulted in higher affinity of ATP to OEP21-WT. Since the LDAO concentration affects oligomerization of OEP21, the reduced binding affinity could be explained by an impaired accessibility of ATP to oligomeric OEP21. Note that different binding affinities are reported when comparing FIGURE 28 and FIGURE 30 with FIGURE 31, which could be due to different setups and different protein batches.

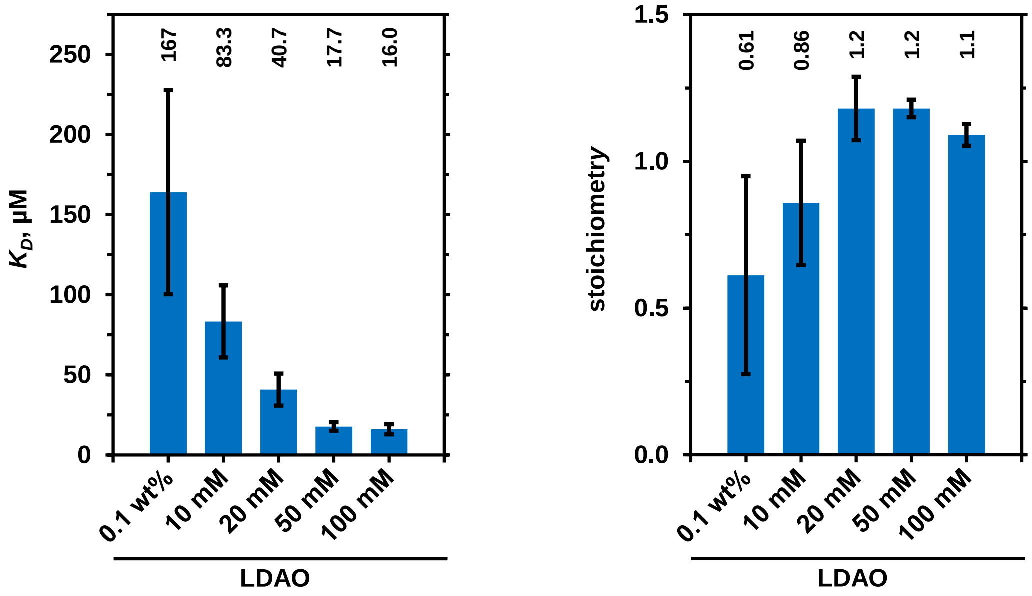


FIGURE 31: Effect of the LDAO concentration on the binding affinity of ATP

ATP in the syringe was titrated into the cell containing OEP21-WT in 25 injection steps, and the effect on differential power was measured at 25 °C. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and different concentrations of LDAO (0.1 wt%, 10 mM, 20 mM, 50 mM, 100 mM). The ATP concentration in the syringe was 2 mM, 2.35 mM, or 4 mM in the case of 50 mM / 100 mM, 0.1 wt%, or 10 mM / 20 mM LDAO, respectively. The OEP21-WT concentration in the cell was 80 µM, except in the case of 0.1 wt% LDAO, where 94 µM protein were used. On the left, the K_D values obtained by ITC at different LDAO concentrations are summarized. On the right, the respective stoichiometries are shown. The error bars represent the standard deviation of the fitted K_D and stoichiometry values, respectively.

To gain a deeper understanding of the function relevance with respect to the binding and transport mechanism of OEP21, the affinity of various negatively charged metabolites bound to OEP21 was investigated. FIGURE 32 shows the K_D values obtained for these metabolites and the effect of different conditions on the binding affinity of ATP. Interestingly, the binding affinity appeared to be dependent on the number of negative charges in the metabolite. This is evident from the lower binding affinity of ADP with 87.5 µM compared to ATP with 11.9 µM. No binding could be detected for AMP by ITC, indicating an even lower binding affinity compared to ADP. Other nucleoside triphosphates also had similar affinities, with a K_D of 10.5 µM and 20.8 µM for CTP and UTP, respectively. Since the K_D of 87.4 µM for GTP did not fit the scheme, binding affinity was verified by FP experiments using fluorescently labeled analogs for GTP and GDP, with MANT-ATP serving as the reference. Using this setup, it was shown that GTP indeed had a comparable binding affinity as ATP, with the GTP analog even binding slightly better to OEP21-WT. Consistent with the results of ITC and FP, GDP exhibited a lower binding affinity than GTP or ATP.

Since the charge of metabolites may depend on pH, the binding of ATP to OEP21-WT was verified at different pH conditions. At pH 7.5, a comparable K_D of 10.4 µM was determined, whereas for pH 6.0, the HEPES buffer system had to be replaced. A comparable K_D of 2.9 µM was also obtained for 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.0. However, when NaPi pH 6.0 was used, the K_D increased significantly to 52.7 µM. When HEPES buffer pH 7.0 was used with an additional 10 mM NaPi, the binding affinity of ATP was reduced to a K_D of 44.9 µM. Therefore, the presence of other negatively charged metabolites likely acted as competitors for ATP.

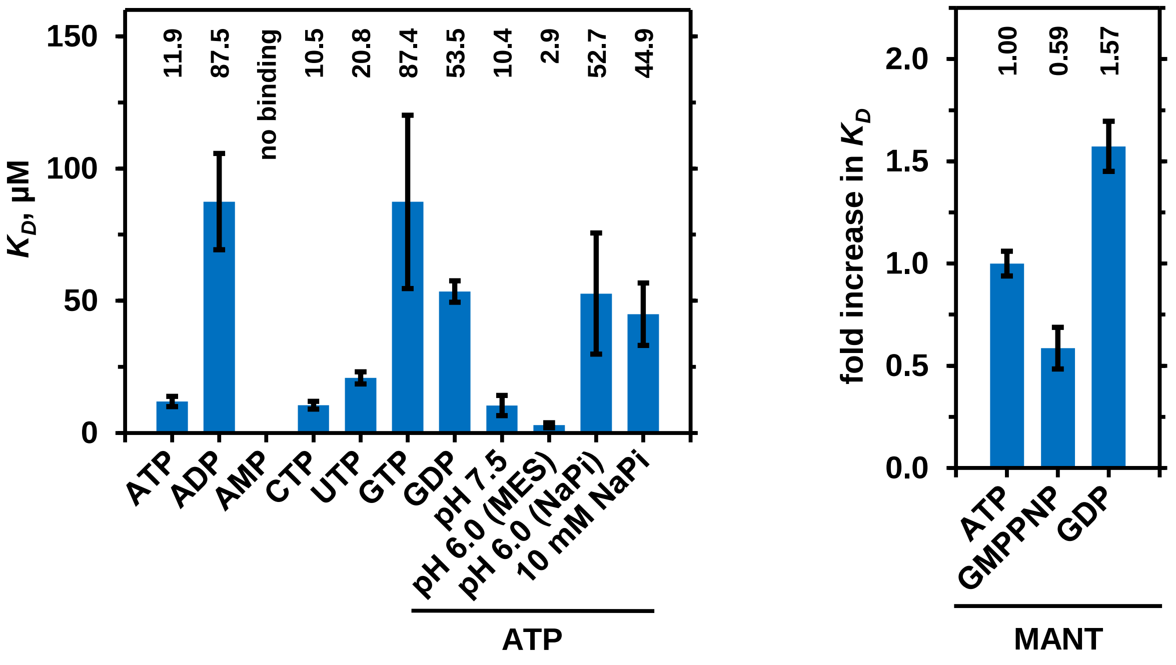


Figure 32: Affinity of various metabolites for OEP21-WT

On the left, the K_D values obtained by ITC for various metabolites to OEP21-WT are summarized. For this purpose, the respective metabolite in the syringe was titrated into the cell containing 75 µM OEP21-WT in 25 injection steps, and the effect on differential power was measured at 25 °C. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 50 mM LDAO. A concentration of 1.5 mM in the syringe was used for ATP, CTP, GTP, and UTP, while the concentration was 2.5 mM for ADP, 2 mM for GDP, and 5 mM for AMP. In addition, the binding affinity of OEP21-WT for ATP was probed under different pH conditions (6.0 and 7.5). At pH 6.0, the buffer system had to be changed by using either 10 mM MES or 10 mM NaPi instead of HEPES. To investigate the effect of phosphate on the ATP affinity, a titration was performed in HEPES buffer with an additional 10 mM NaPi. The error bars represent the standard deviation of the fitted K_D values. On the right, normalized K_D values probed by FP with fluorescently labeled MANT-ATP (reference), MANT-GMPPNP (comparable to GTP), and MANT-GDP are shown. The mean values of three independent FP experiments were used to determine the K_D values, and the error bars represent the standard deviations of the fitted K_D values.

FIGURE 33 shows the titration of GAP into buffer and to OEP21-WT, where positive spikes in differential power were observed for the buffer reference, indicating an endothermic reaction.

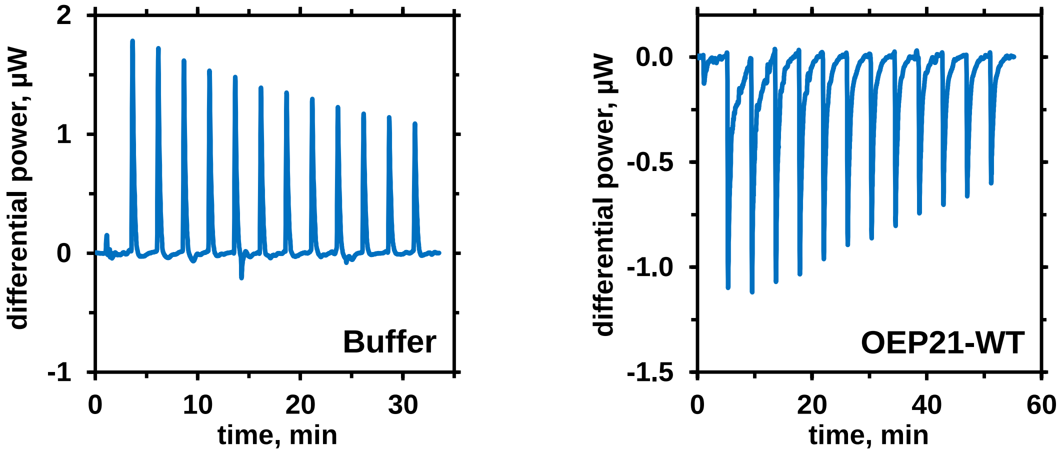


FIGURE 33: Probing the binding affinity of GAP to OEP21-WT by ITC

GAP at a concentration of 6 mM in the syringe was titrated into the cell containing either buffer without protein (left) or 150 µM OEP21-WT (right) in 13 injection steps, and the effect on differential power was measured at 25 °C. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 50 mM LDAO. Due to the poor baseline of differential power in the case of OEP21-WT, the injection spacing was elongated to 250 s, but still no equilibrium was observed for the first injections steps.

In the case of OEP21-WT, minor negative spikes in differential power were detected, indicating an exothermic reaction and thus suggesting binding of GAP to OEP21-WT. Because the changes in differential power due to titration of GAP were too minor compared with ATP, determination of K_D and stoichiometry between multiple titration experiments did not yield reliable values. In addition, equilibrium for the baseline of differential power could not be reached during GAP titration, even with an elongated injection spacing. Therefore, titrations to OEP21-WT were conducted with the more negatively charged 3-PGA (data not shown), resulting in a similar pattern as overserved for GAP. Due to the limitations of ITC, no reliable K_D values could be obtained for GAP or 3-PGA, suggesting no binding or rather binding with a K_D in a high micromolar to low millimolar range, as indicated by the development of small exothermic heats during titration. To demonstrate that GAP can bind to OEP21, and that GAP can compete with ATP, the binding of ATP to OEP21-WT was studied in the presence of increasing GAP concentrations, as shown in FIGURE 34.

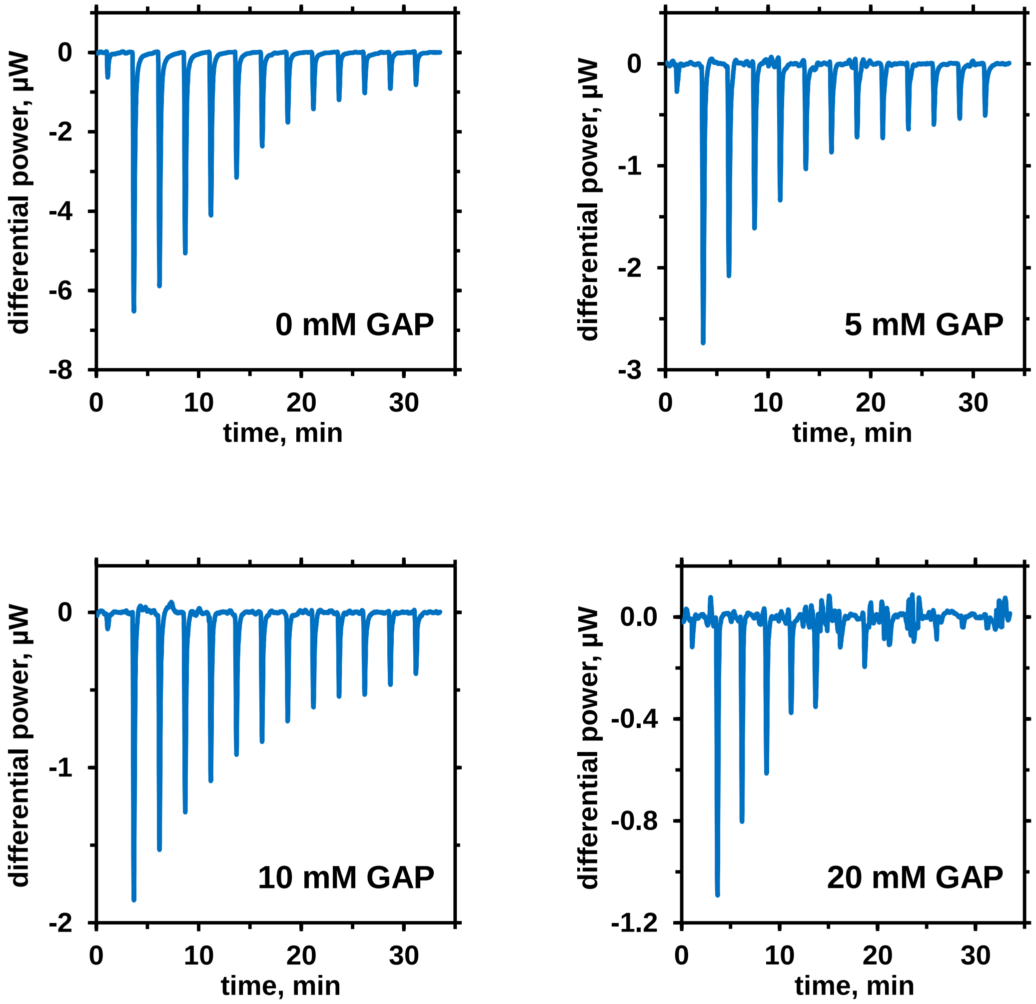


FIGURE 34: Probing the binding affinity of ATP to OEP21-WT in presence of GAP

ATP at a concentration of 1.2 mM in the syringe was titrated into the cell containing 75 µM OEP21-WT in 13 injection steps, and the effect on differential power was measured at 25 °C. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 50 mM LDAO, and different concentrations of GAP (0 mM, 5 mM, 10 mM, 20 mM).

Both were evidenced by the decreasing exothermic heat spikes that correlate with increased GAP concentration. Therefore, GAP at higher concentration bound with low affinity to OEP21-WT, which reduced the exothermic heat development by titration of ATP compared to a setup without GAP. Due to the lower exothermic heat development at higher GAP concentrations, the K_D values and stoichiometries could not be consistently derived from the data.

3.1.7. Characterization of the binding sites by NMR spectroscopy

To further characterize the binding sites of OEP21 for negatively charged metabolites, NMR experiments were conducted. Stepwise titration of ATP to OEP21 and analysis of the ligand-induced chemical shift perturbations in OEP21 allowed the binding site of ATP to be mapped on OEP21, as shown in FIGURE 35 and FIGURE 36.

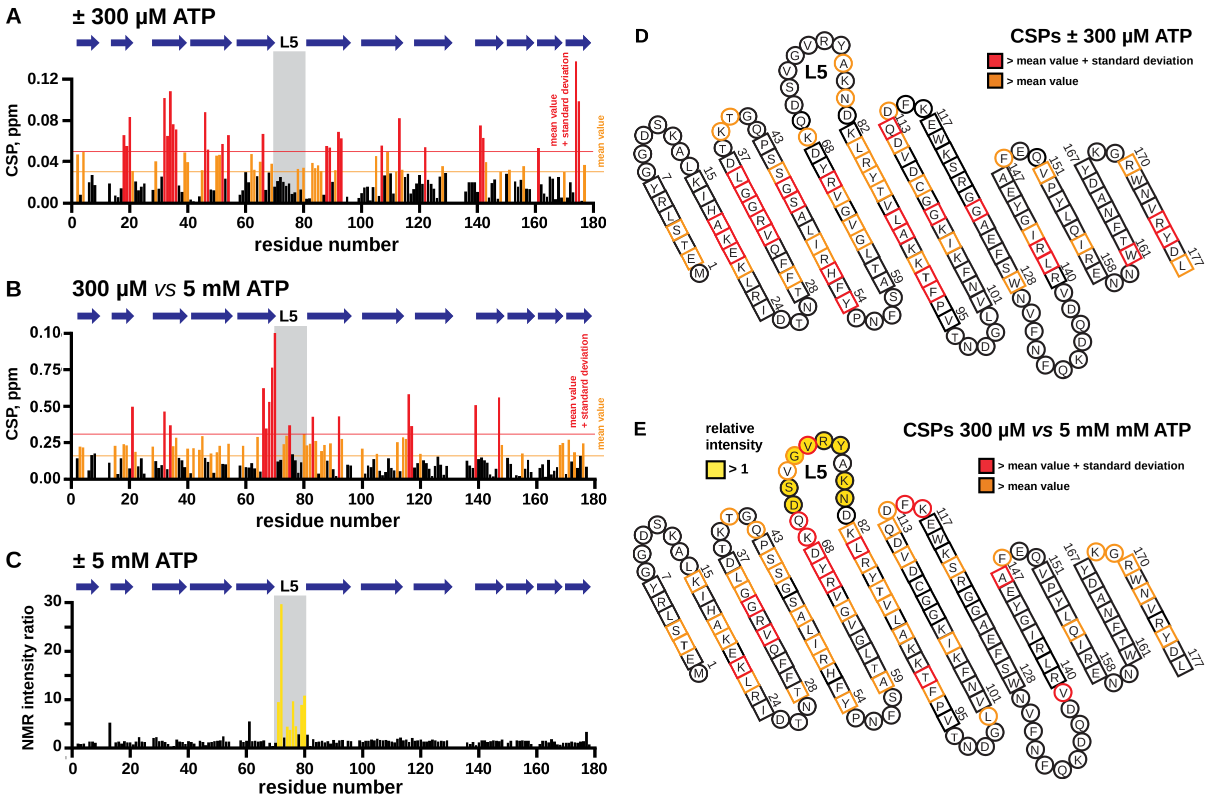


FIGURE 35: NMR titration experiments with ATP

2D-[^{15}N,^{1}H]-TROSY NMR spectra were recorded in the presence of ATP. At an equimolar ratio of protein to ATP (A and D), chemical shift perturbations (CSPs) clustered within the β-barrel pore upon addition of 300 µM ATP. At a high excess of ATP (B and E), chemical shift perturbations were observed in the region around loop five upon the addition of 5 mM ATP compared to 300 µM ATP. In addition to chemical shift perturbations, residues in loop five showed a significant increase in their NMR signal intensity upon addition of 5 mM ATP (C and E), indicating a weak interaction with ATP. Significant chemical shift perturbations (> mean value and > mean value plus standard deviation) at both titration steps are colored in orange and red, respectively. Relative NMR signal intensities > 1 are colored in yellow. All spectra were recorded on a 950 MHz spectrometer with 300 µM ^{2}H,^{15}N-labelled OEP21-WT in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, 7 vol% D_2O, and 350 mM LDAO with 32 scans per increment and 160 complex points in the indirect ^{15}N dimension at 303 K. This figure was provided by Prof. Dr. Franz Hagn.

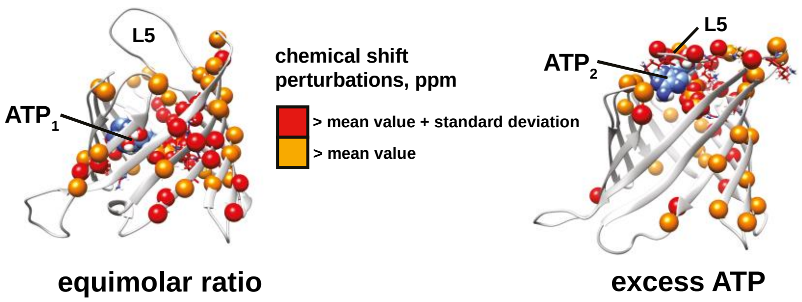
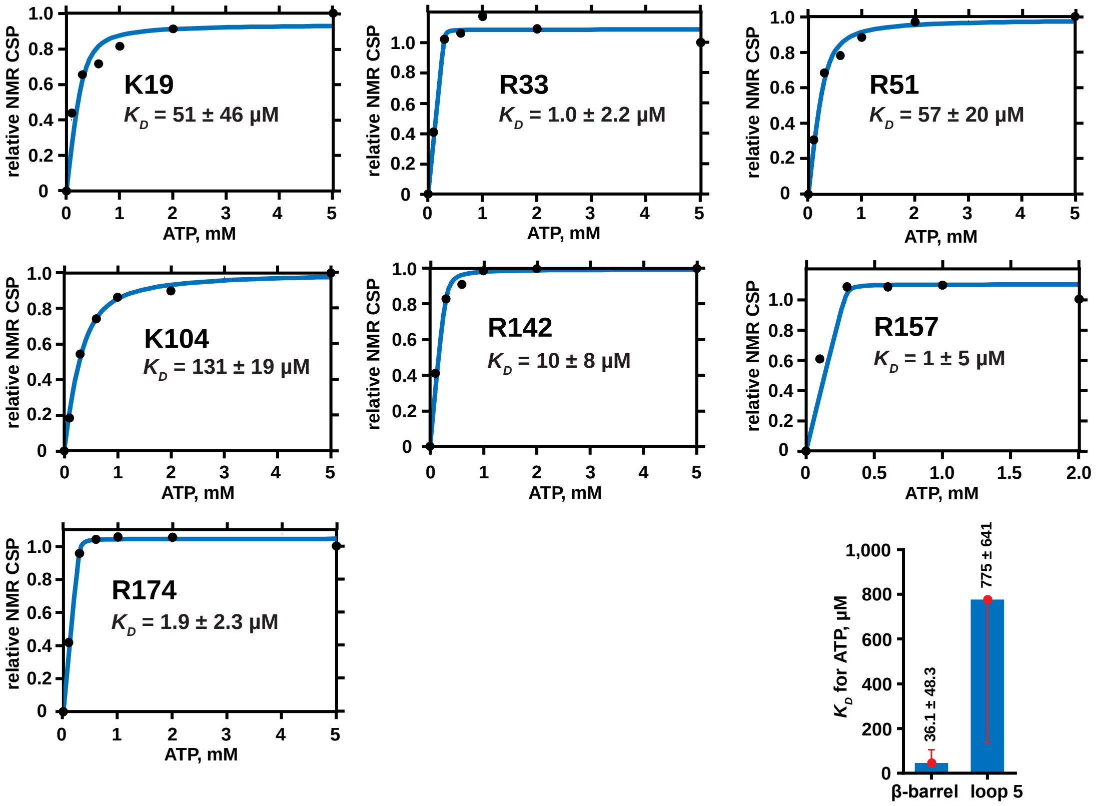


FIGURE 36: NMR chemical shift perturbations induced by ATP

The NMR chemical shift perturbations of amino acid residues obtained from FIGURE 35 were mapped onto a molecular dynamics-based complex structural model of OEP21 and ATP at an equimolar ratio (300 µM) and an excess (5 mM) of ATP. This revealed two binding sites for ATP, an internal binding site with high affinity (ATP$_1$) and a peripheral binding site with low affinity (ATP$_2$). This figure was provided by Prof. Dr. Franz Hagn.

At an equimolar ratio of protein to ATP, the binding site within the β-barrel pore was completely occupied by ATP. At a high excess of ATP, chemical shift perturbations were induced and an increase in NMR signal intensity was observed at another binding site located at the cytoplasmic entrance of the pore that involved loop five. By analysis of the NMR titrations with ATP, K_D values in the low micromolar range were derived for the residues involved in binding within the pore, as shown in FIGURE 37. In contrast, for the peripheral binding site, K_D values of almost 1 mM were determined for the involved residues. Compared with ITC, the high-affinity site inside the pore agreed well with the affinity derived from ITC, whereas the low-affinity peripheral binding site could not be observed with ITC due to sensitivity limitations. Interestingly, deletion of loop five reduced ATP affinity at the peripheral binding site (data not shown).

high-affinity ATP-binding site inside the β-barrel (ATP_1)



low-affinity ATP-binding site at loop 5 (ATP_2)

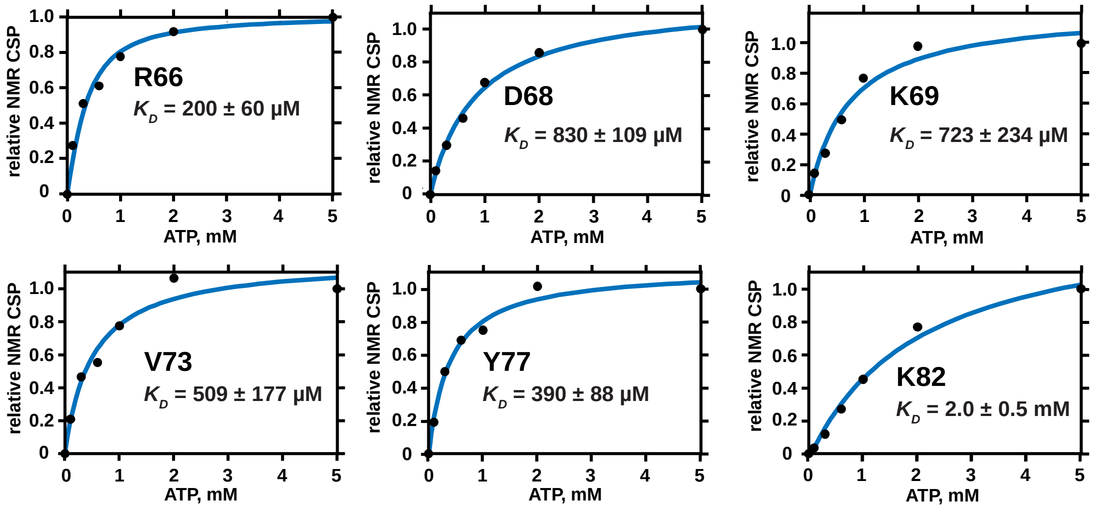


FIGURE 37: K_D for ATP derived from single amino acid residues by NMR titration experiments

The dissociation constant K_D for ATP of single amino acid residues of OEP21 was derived from NMR chemical shift perturbation (CSP) induced by titration with ATP (shown in FIGURE 35). The amino acid residues were located either in the β-barrel pore (upper panel) or in loop five (lower panel), with the bar diagram showing the average K_D values for ATP of the β-barrel and loop five region. These data were provided by Prof. Dr. Franz Hagn.

Analogous to ATP, the binding sites of GAP were mapped on OEP21 by NMR titrations with GAP, as shown in FIGURE 38 and FIGURE 39.

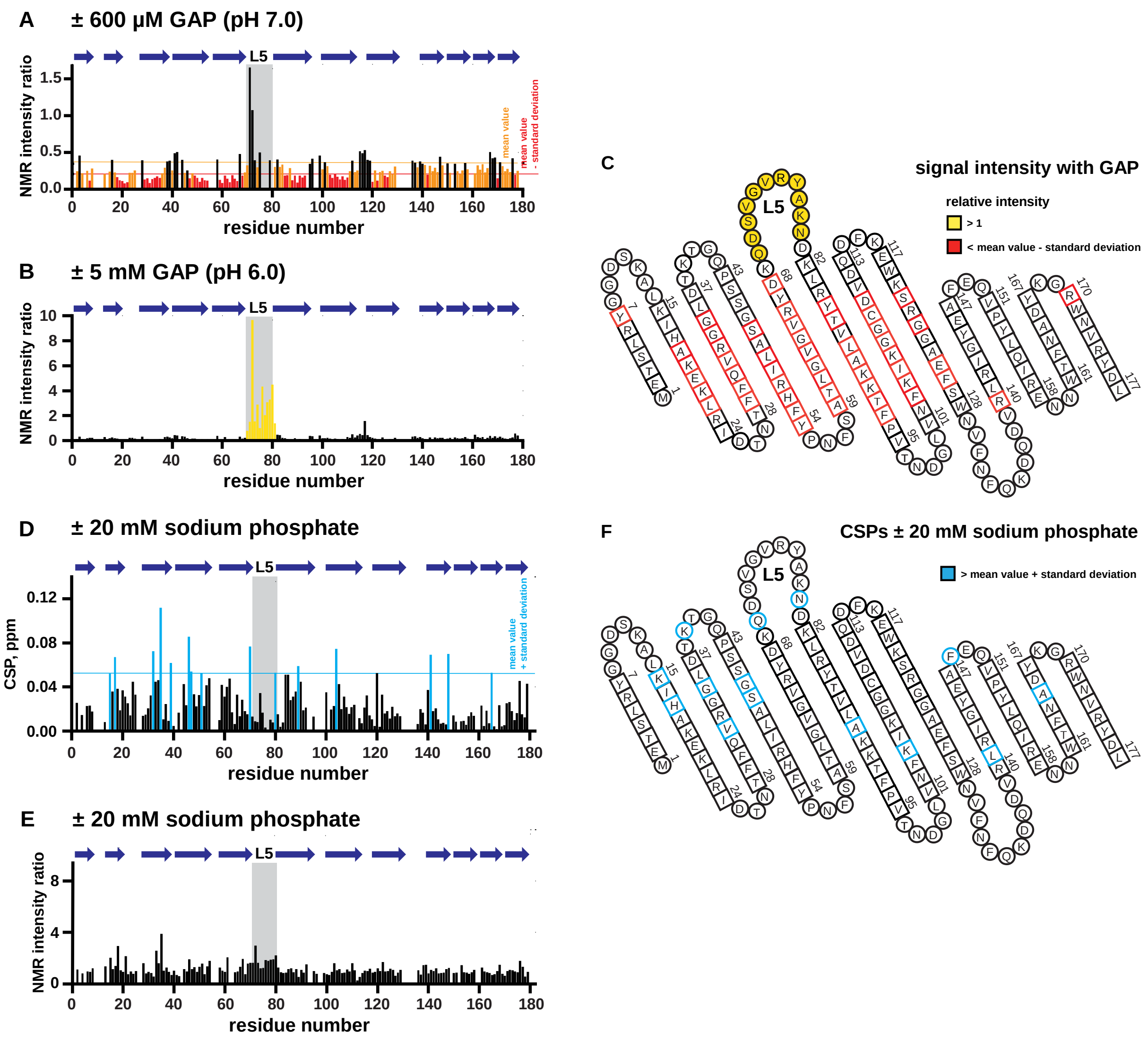


FIGURE 38: NMR titration experiments with GAP and NaPi

2D-[^{15}N,^{1}H]-TROSY NMR spectra were recorded in the presence of GAP and sodium phosphate. Upon addition of 600 µM GAP at pH 7.0 (A and C), the NMR signal intensities of amino acids residues located on the side of the β-barrel that is involved in dimerization and has the highest density of positively charged amino acids decreased compared to those without GAP. Significant decrease in NMR signal intensities (< mean value and < mean value minus standard deviation) upon GAP addition are colored in orange and red, respectively. In contrast, the NMR signal intensities of amino acids residues located in the loop five region increased significantly upon addition of 5 mM GAP at pH 6.0 (B and C), with relative NMR signal intensities > 1 colored in yellow. Upon addition of 20 mM NaPi, minor chemical shift perturbations (CSPs) of some amino acid residues were observed (D and F), with significant chemical shift perturbations (> mean value plus standard deviation) colored in blue. In addition, NaPi did not affect the loop five region like GAP, as shown by only minor changes in NMR signal intensities (E). All spectra were recorded on a 950 MHz spectrometer with ^{2}H,^{15}N-labelled OEP21-WT in a buffer (NaPi pH 6.0 or HEPES pH 7.0) containing 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 7 vol% D_2O with 32 scans per increment and 160 complex points in the indirect ^{15}N dimension at 303 K. A buffer with 20 mM HEPES pH 7.0 was used for the titration experiments with 20 mM NaPi and 600 µM GAP, with samples containing 350 µM / 400 µM protein and 550 mM / 500 mM LDAO, respectively. For the titration experiment with 5 mM GAP, a buffer with 20 mM NaPi pH 6.0 and a sample containing 360 µM protein and 400 mM LDAO was used. This figure was provided by Prof. Dr. Franz Hagn.

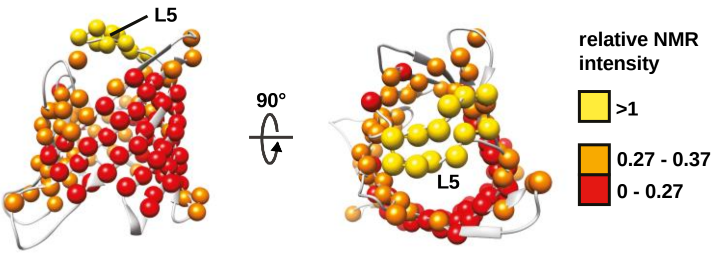


FIGURE 39: NMR signal intensities changes induced by GAP

The relative NMR signal intensities of amino acid residues upon addition of GAP obtained from FIGURE 38 were mapped onto the OEP21 structure. The NMR signals of residues in the β-barrel region are strongly broadened, while loop five showed increased NMR signal intensities, indicating an interaction with GAP. This figure was provided by Prof. Dr. Franz Hagn.

In contrast to ATP, GAP induced line broadening of NMR signals in the β-barrel structure, as shown in FIGURE 40, where the line broadening of the NMR signals was dependent on the GAP concentration.

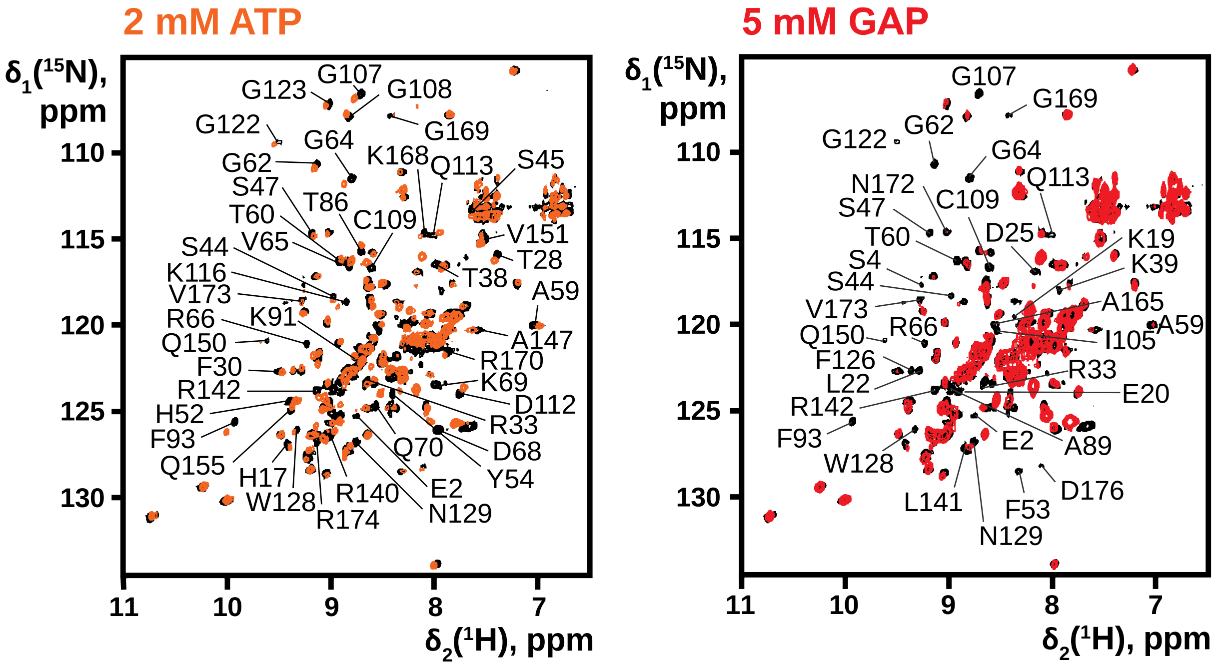


FIGURE 40: Effect of ATP and GAP on the NMR spectrum

2D-[^{15}N,^{1}H]-TROSY NMR spectra of OEP21-WT were recorded without metabolite (black) and in the presence of 2 mM ATP (orange; on the left) or 5 mM GAP (red; on the right). All spectra were recorded on a 950 MHz spectrometer with 300 µM ^{2}H,^{15}N-labelled OEP21-WT in 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, 7 vol% D_2O, and 300 mM LDAO with 48 scans per increment and 192 complex points in the indirect ^{15}N dimension at 303 K. In the presence of GAP, the NMR spectrum showed mainly line broadening of NMR resonances. In contrast, the presence of ATP induced chemical shift perturbations, and only minor line broadening effects were observed. This figure was provided by Prof. Dr. Franz Hagn.

Interestingly, the residues with the most pronounced line broadening effects are located on the side of the β-barrel that is involved in dimerization and has the highest density of positively charged amino acids. The NMR signal intensities of amino acids residues located in the loop five region increased significantly upon addition of GAP, suggesting binding of GAP to loop five. In contrast to GAP, sodium phosphate did not induce line broadening of NMR signals and had little effect on NMR signal intensities. To confirm that the effects of GAP on the OEP21 spectrum were specific and to rule out any adverse effects of GAP on the LDAO micelle, NMR spectra of the unrelated bacterial β-barrel protein OmpX in LDAO, in which no line broadening was observed, were recorded (data not shown). As probed by the line broadening effect on the spectrum, GAP bound to OEP21 with a K_D of 152 µM ± 19 µM, as shown in FIGURE 41.

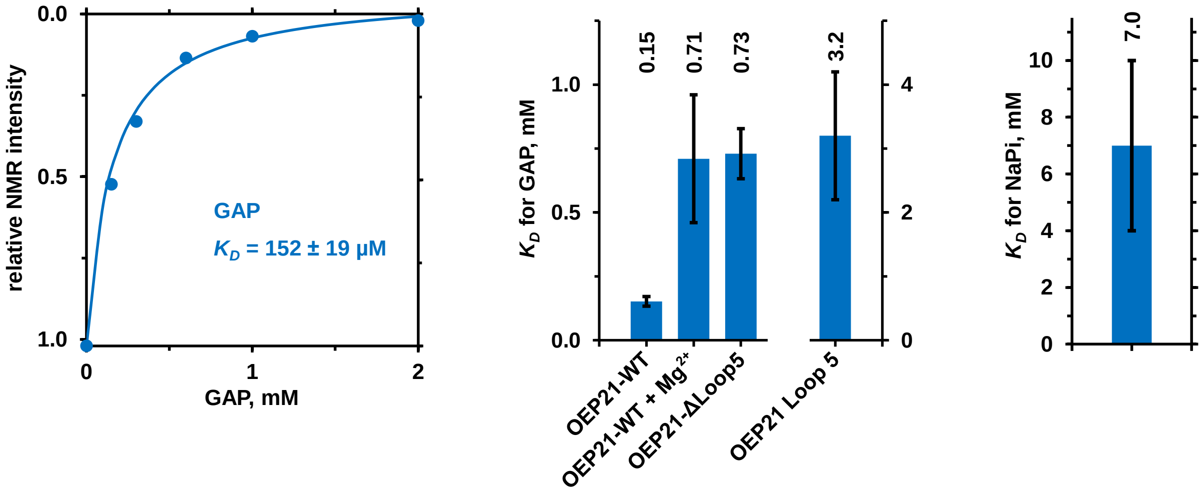


FIGURE 41: K_D for GAP and NaPi derived from NMR experiments

2D-[^{15}N,^{1}H]-TROSY NMR spectra were recorded in the presence of increasing concentrations of GAP to derive K_D values. On the left, the decrease in the relative NMR signal intensity of arginine residue 66 is shown as an example, where a K_D of 152 µM ± 19 µM was derived for GAP. In the middle, K_D values for GAP are summarized for different setups. The affinity for GAP was weakened by the addition of magnesium to 710 µM or the deletion of loop five to 730 µM. Loop five alone interacted only weakly with GAP, as shown by the K_D of 3.2 mM. On the right, a K_D of 7 mM ± 3 mM was derived from chemical shift perturbations for sodium phosphate. All spectra were recorded on a 950 MHz spectrometer with ^{2}H,^{15}N-labelled OEP21-WT or OEP21-ΔLoop5 in 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 7 vol% D_2O at 303 K. For the titration experiments with OEP21-WT and OEP21-ΔLoop5, samples with 400 µM protein and 500 mM / 430 mM LDAO were used, and the spectra were recorded with 48 / 32 scans per increment and 128 / 160 complex points in the indirect ^{15}N dimension, respectively. For the titration experiments in the presence of magnesium, a buffer containing 250 µM $MgCl_2$ and no EDTA was used, with $MgCl_2$ titrated into the sample at the same concentration as GAP. Therefore, a sample with 300 µM OEP21-WT and 430 mM LDAO was used, and the spectrum was recorded with 40 scans per increment and 160 complex points in the indirect ^{15}N dimension. For the titration experiments with NaPi, a sample with 350 µM OEP21-WT and 350 mM LDAO was used, and the spectrum was recorded with 32 scans per increment and 160 complex points in the indirect ^{15}N dimension. This figure was provided by Prof. Dr. Franz Hagn.

As seen previously, higher concentration of GAP resulted in increased NMR signal intensity of residues located in loop five, comparable to ATP. This effect and the effect of line broadening were used to derive affinities of GAP with the β-barrel of OEP21 as well as with loop five and for OEP21-ΔLoop5, which lacks loop five, as shown in FIGURE 41. This suggests that loop five was involved in GAP binding and mediated an approximately fourfold increase in binding affinity. Simultaneous addition of magnesium chloride in the NMR titrations with GAP decreased the binding affinity in a comparable manner. In addition, a K_D of 7 mM ± 3 mM was derived from chemical shift perturbations for sodium phosphate, as shown in FIGURE 41.

Consistent with the NMR data above indicating interaction of loop five with ATP and GAP, amide hydrogen exchange rates in loop five are reduced in complex with GAP and ATP, as shown in FIGURE 42. This demonstrated that loop five is involved in the interaction between OEP21 and metabolites, with stronger effects evident for the higher affinity binder ATP.

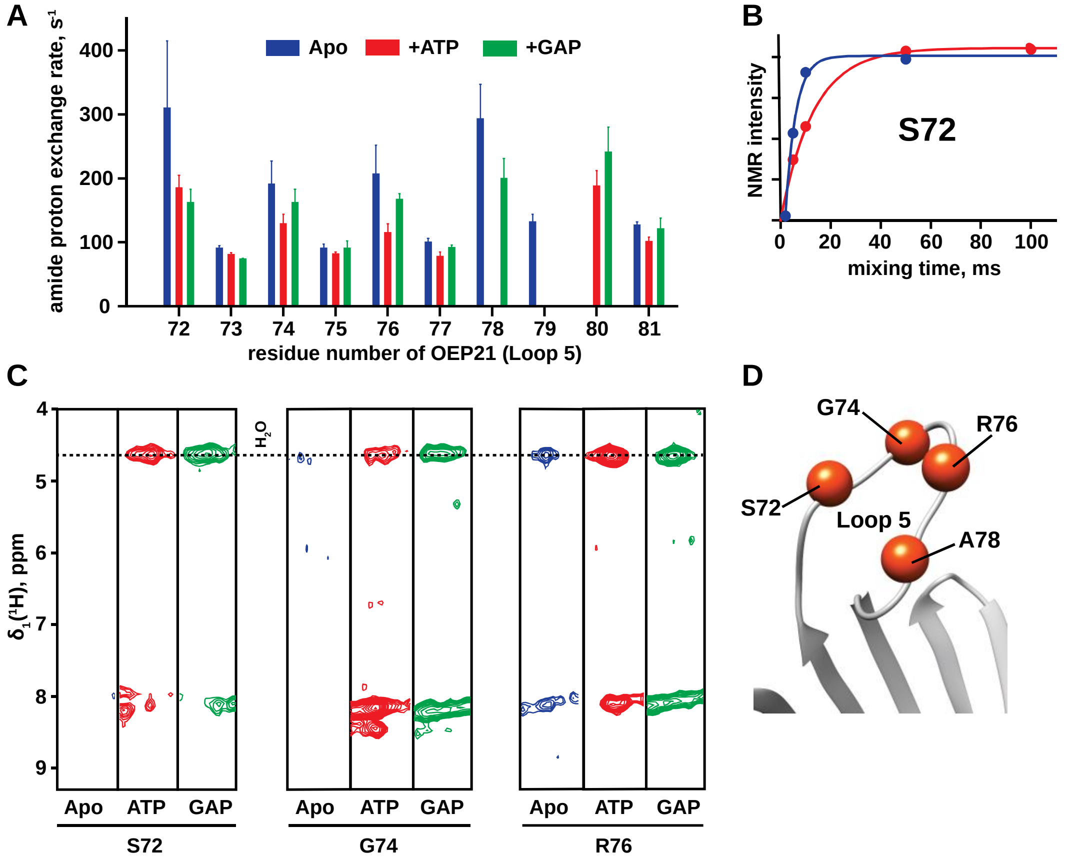


FIGURE 42: Loop five participates in the OEP21 metabolite interaction

Amide proton exchange rates at pH 6.0 for OEP21 loop five residues were determined by CLEANEX NMR experiments [183] in the apo form (blue) or in the presence of 5 mM ATP (red) and 5 mM GAP (green), with reduced values obtained when bound to ATP or GAP (A). At pH 6.0, most of the NMR signals of loop five in the 2D-[^{15}N,^{1}H]-TROSY NMR spectrum can be observed even in the apo form, allowing comparison. The error bars represent the errors obtained by curve fitting. Missing values are due to signal overlap or low peak intensity in the corresponding NMR spectra. In (B), a representative built-up curve of the exchange experiment is shown for residue S72 in the apo state (blue) and in the ATP-bound (red) state. Strips from a 3D-^{15}N-edited-[^{1}H,^{1}H]-NOESY-TROSY experiment (40 ms mixing time) are shown in (C), with diagonal and cross peaks absent or very weak in the apo form but visible when bound to ATP or GAP. The residues in loop five, showing high exchange rates in the apo state and a marked reduction in the ligand-bound states, are shown in (D). All spectra were recorded on a 950 MHz spectrometer with 700 µM ^{2}H,^{15}N-labelled OEP21-WT in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, 7 vol% D_2O, and 330 mM LDAO at 303 K. This figure was provided by Prof. Dr. Franz Hagn.

3.1.8. Generating and verify the OEP21 variants

To investigate the influence of individual amino acids in OEP21 on the ATP affinity, various OEP21 variants were generated in which single positively charged lysine or arginine residues were mutated to alanine. Positions were identified by NMR titrations affected by ATP, and an overview of the mutated residues is shown in FIGURE 43.

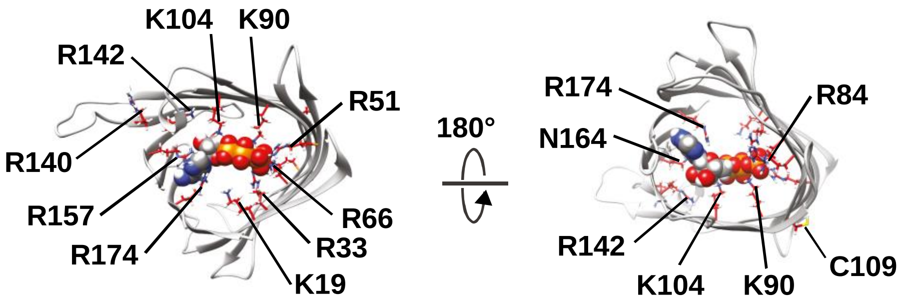


FIGURE 43: Overview of the mutated positively charged residues in OEP21

Single positively charged lysine or arginine residues were mutated to alanine. The positions were identified by NMR titrations affected by ATP.

In addition, a construct lacking loop five (OEP21-ΔLoop5) was generated. The OEP21 variant R66A-N164R has the same positive charge as the wildtype, but the position of an arginine residue in the pore is relocated. The OEP21 variants were expressed, purified, and refolded analogously to OEP21-WT, with all OEP21 variants exhibiting comparable expression levels, refolding behavior, and SEC elution profiles to OEP21-WT. The final purity after SEC of the OEP21 variants was analyzed by SDS-PAGE, and the gels are shown in FIGURE 44, where all variants were obtained highly pure.

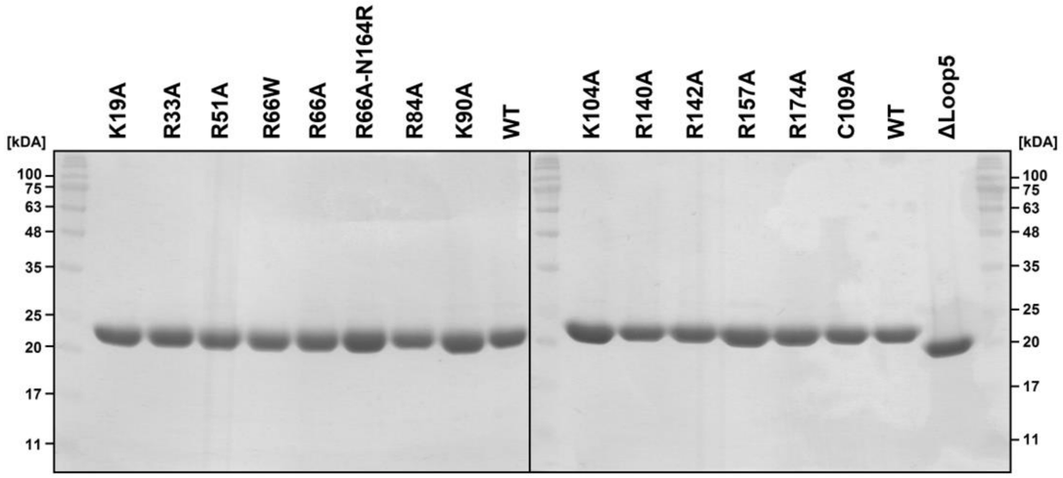


FIGURE 44: Purity of the OEP21 variants

Coomassie-stained SDS-PAGE gels shows the purified OEP21 variants after SEC.

To verify that all OEP21 variants are properly folded and comparable to the wildtype, far-UV CD spectra and thermal melting curves of the OEP21 variants were recorded, as shown in FIGURE 45.

Indeed, the OEP21 variants exhibited all a β-sheet secondary structure with a minimum mean residue ellipticity around 215 nm and comparable thermal stabilities to OEP21-WT. Except for OEP21-ΔLoop5, which exhibited relatively more β-sheet secondary structure due to the lack of random coil region of loop five. Moreover, the thermal stability of OEP21-ΔLoop5 was increased by 10 °C. In conclusion, single mutations of positively charged residues did not significantly affect the overall structure of OEP21, allowing further analysis of the variants.

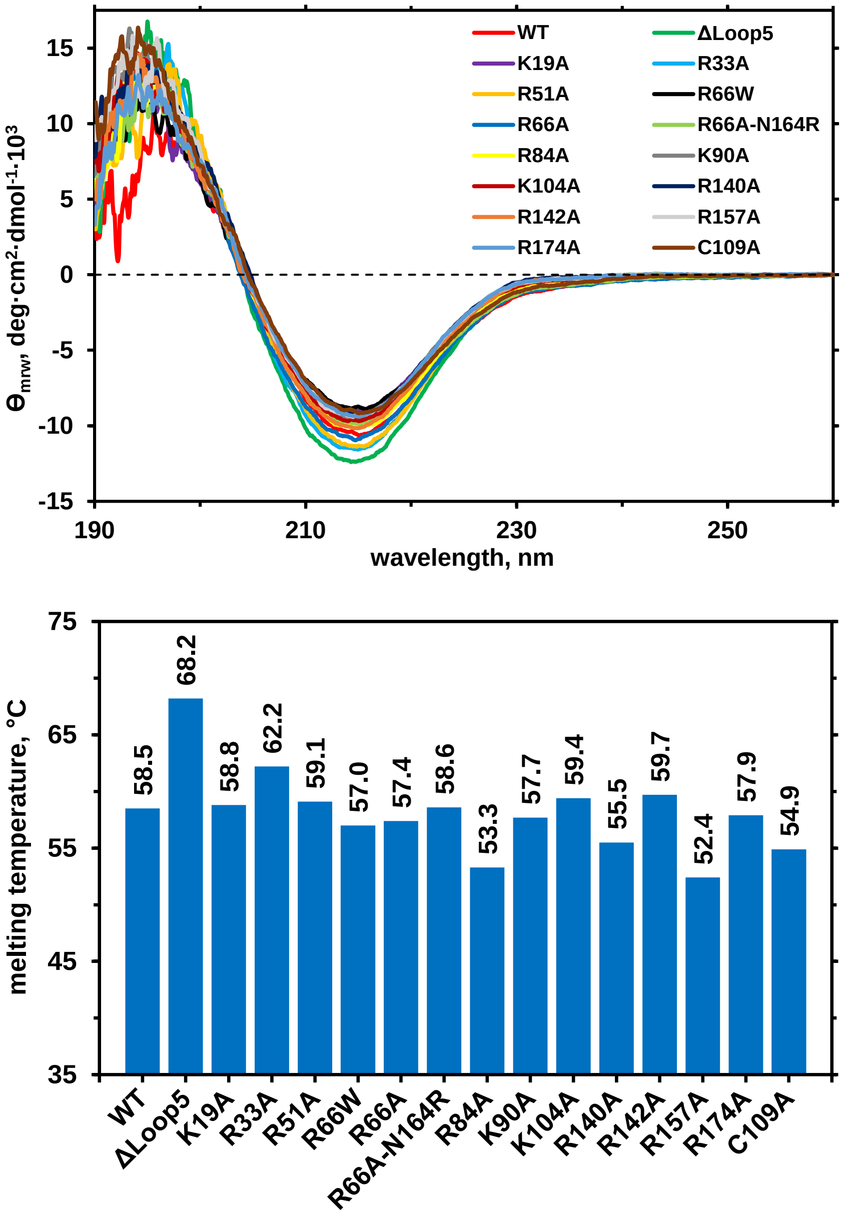


FIGURE 45: Far-UV CD spectra and thermal stabilities of the OEP21 variants

Above, the mean residue weight ellipticity Θ_{mrw} of the OEP21 variants in LDAO from a wavelength of 190 nm to 260 nm at 20 °C is shown. Below, the thermal melting points of the OEP21 variants are summarized, where the thermal melting behavior was observed by heating up from 20 °C to 100 °C at a wavelength of 215 nm. All spectra were recorded with 10 µM protein in 10 mM NaPi pH 7.0, 0.5 mM DTT, and 0.1 wt% LDAO.

In addition, an OEP21 construct lacking the His$_{10}$-tag (OEP21-ΔHis) was used to exclude possible interfering effects of the His$_{10}$-tag on the experiments. OEP21-ΔHis could be obtained in a comparable manner to OEP21-WT, as shown in Figure 46 based on the SEC chromatogram and final purity. OEP21-ΔHis showed slightly more impurities compared to OEP21-WT, which could be due to the use of cation exchange chromatography instead of Ni-NTA affinity chromatography.

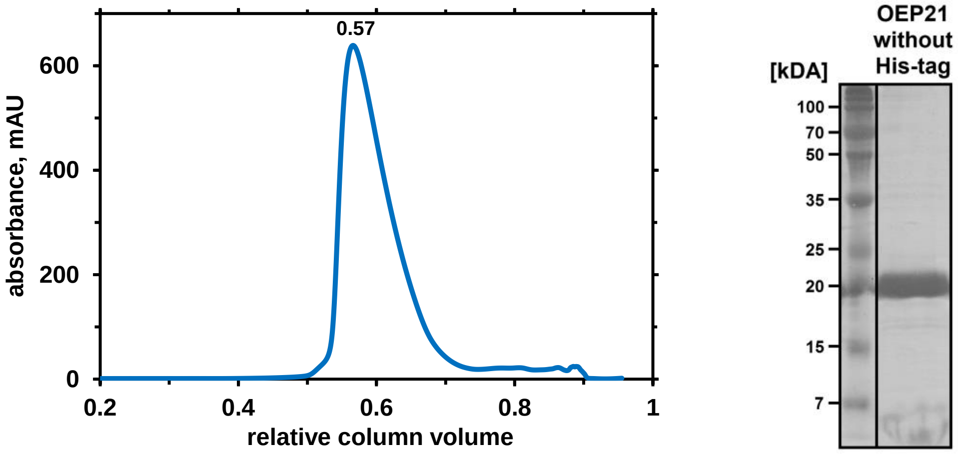


Figure 46: SEC chromatogram of refolded OEP21-ΔHis

On the left, refolded OEP21-ΔHis was applied onto a Superdex 200 Increase 10/300 GL SEC column. The running buffer consisted of 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 0.1 wt% LDAO. The main peak eluted at 0.57 CV, as indicated by the maximum absorbance measured at 280 nm (blue curve). On the right, the purity of the main peak is verified by a Coomassie-stained SDS-PAGE gel.

In addition, the CD-detected secondary structure content and thermal stability were compared to OEP21-WT, as shown in Figure 47. OEP21-ΔHis exhibited slightly less β-sheet secondary structure elements than OEP21-WT, which might be due to the higher amount of impurities. Interestingly, the thermal stability of OEP21-ΔHis was about 3.5 °C higher than that of OEP21-WT, which could be due to several reasons, such as batch-dependent fluctuations, the protein-to-detergent ratio, the type and amount of impurities, or the His$_{10}$-tag itself. However, the charge-dependent stabilization of OEP21-ΔHis in the presence of GAP and ATP agreed well with that of OEP21-WT, considering the generally higher thermal stability of the apo form.

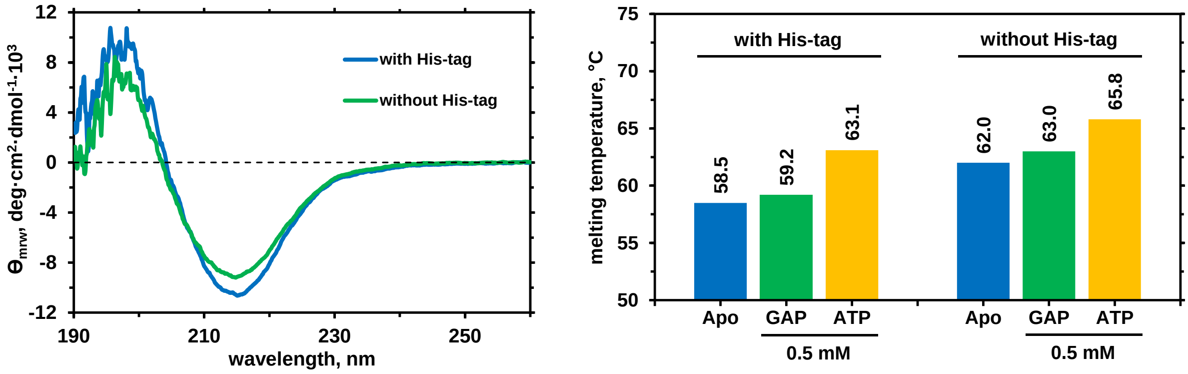


FIGURE 47: Far-UV CD spectrum and thermal melting of OEP21-ΔHis

On the left, the mean residue weight ellipticity Θ_{mrw} of OEP21-WT (blue) and OEP21-ΔHis (green) in LDAO micelles at a wavelength from 190 nm to 260 nm at 20 °C is shown. On the right, the thermal melting behavior of OEP21-WT and OEP21-ΔHis by heating up from 20 °C to 100 °C at a wavelength of 215 nm in the apo form or in the presence of 0.5 mM GAP and 0.5 mM ATP is summarized. All spectra were recorded with 10 µM protein in 10 mM NaPi pH 7.0, 0.5 mM DTT, and 0.1 wt% LDAO.

3.1.9. Determination of the binding affinity of the OEP21 variants

To investigate the effect of mutated positively charged residues on the binding affinity for ATP, ITC and FP experiments were performed. First, the K_D values for the OEP21 variants were determined by ITC, as summarized in FIGURE 48.

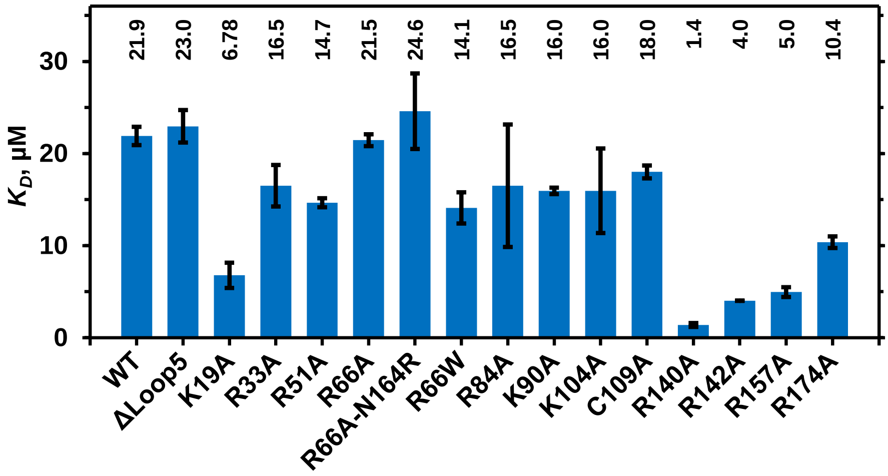


FIGURE 48: K_D values of the OEP21 variants for ATP determined by ITC

ATP at a concentration of 1.2 mM in the syringe was titrated into the cell containing 75 µM of the respective OEP21 variant in 13 injection steps, and the effect on differential power was measured at 25 °C. The buffer used contained 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 50 mM LDAO was used. The resulting K_D values of the OEP21 variants for ATP have been summarized above, with K_D values and error bars representing mean values and standard deviations derived from two independent ITC experiments.

The derived K_D values for the OEP21 variants were all in a low micromolar range, between 1 µM to 25 µM. Because of the fluctuating K_D values obtained by ITC between several experiments, it could not be assumed with certainty that the differences were caused by the mutations. Therefore, FP experiments were conducted with fluorescently labeled MANT-ATP to derive K_D values for the OEP21 variants. FIGURE 49 shows an example of the effect on FP of stepwise titration of OEP21-WT to MANT-ATP. From these data, a K_D of 182 nM was determined for MANT-ATP. In contrast, the binding affinity for ATP determined by ITC was significantly weaker with a K_D of 21.9 µM, which may reflect increased affinity in the presence of the hydrophobic MANT group besides the use of different methods. As control, LDAO was titrated to MANT-ATP as shown in FIGURE 49, and no effect on FP was observed even up to 100 mM LDAO (data not shown).

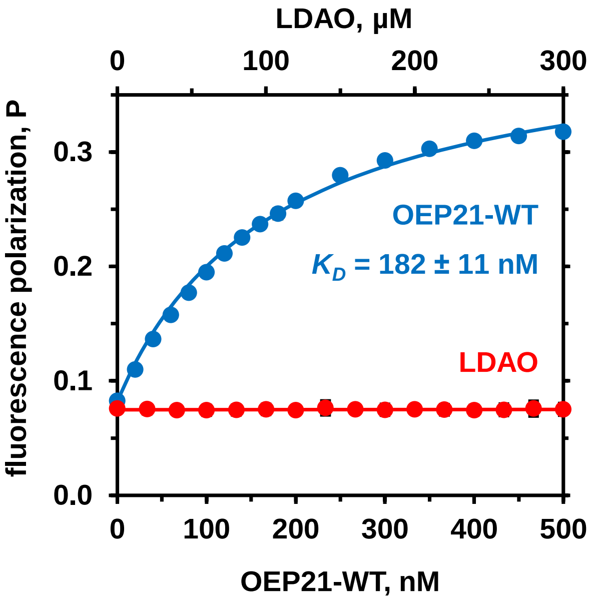


FIGURE 49: Exemplarily FP titration experiment with OEP21-WT to MANT-ATP

FP of stepwise titrated OEP21-WT (blue) or LDAO (red, control) to 200 nM MANT-ATP in 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 0.1 wt% LDAO at 25 °C was measured. A K_D of 182 nM ± 11 nM was derived from the increase in FP by addition of OEP21-WT.

To verify that MANT-ATP binds in the same manner as ATP, ATP was titrated as a competitor to a preformed complex of OEP21-WT and MANT-ATP, as shown in FIGURE 50. Indeed, ATP was able to competitively displace MANT-ATP from OEP21, with a half maximal inhibitory concentration (IC_{50}) of 4.8 µM, demonstrating that MANT-ATP was suitable for FP experiments.

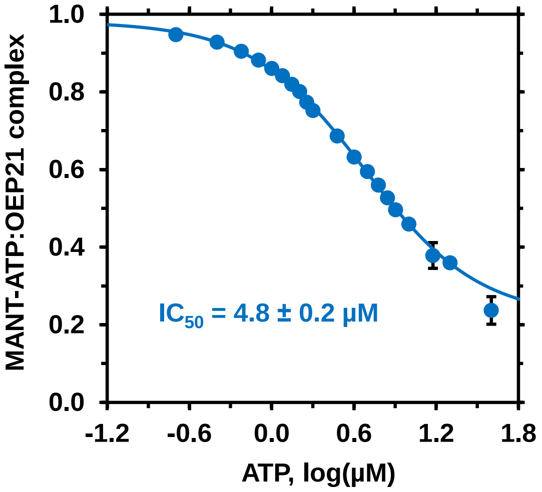


FIGURE 50: Competition of ATP with MANT-ATP

FP of stepwise titrated ATP to a preformed complex of 1 µM MANT-ATP and 1 µM OEP21-WT in 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 0.1 wt% LDAO at 25 °C was measured. FP values were normalized (MANT-ATP:OEP21-WT complex) and an IC_{50} of 4.8 µM ± 0.2 µM was derived for competitor ATP.

Analogous to the K_D determination for OEP21-WT, binding affinities were determined for the other OEP21 variants, with the raw data illustrated in FIGURE 51 and the derived K_D values summarized in FIGURE 52.

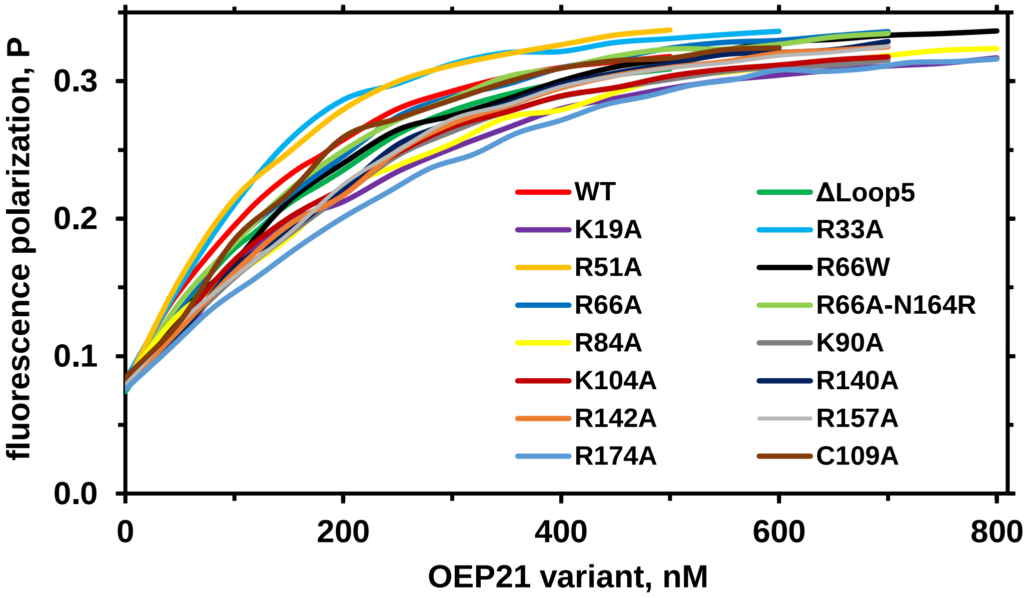


FIGURE 51: Change in FP by titration of the OEP21 variants

FP of stepwise titrated protein of the respective OEP21 variant to 200 nM MANT-ATP in 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 0.1 wt% LDAO at 25 °C was measured. For clarity, the single injection points were connected via a curve.

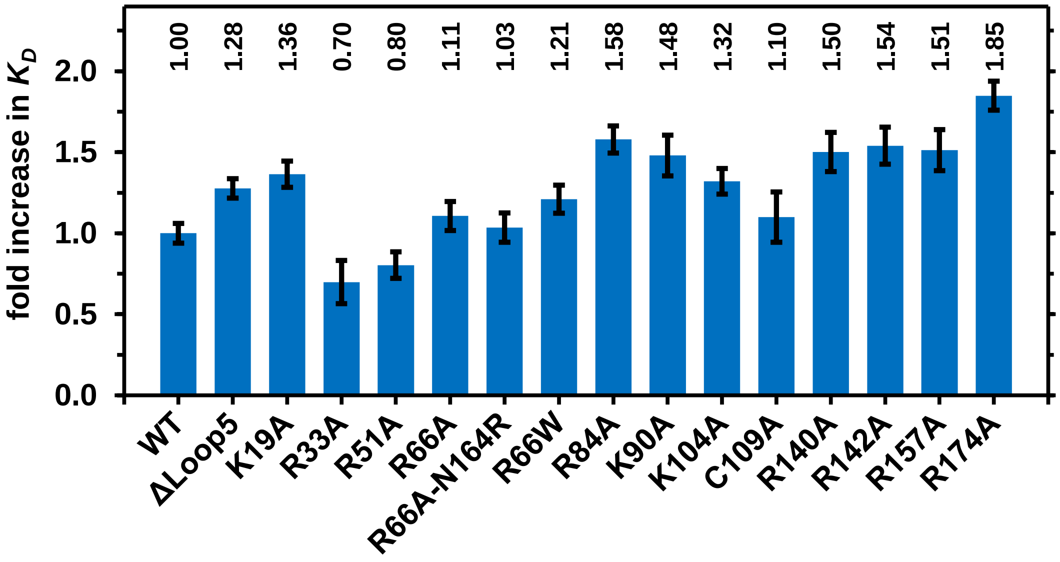


FIGURE 52: Binding affinities for the OEP21 variants determined with FP

FP of stepwise titrated protein of the respective OEP21 variant to 200 nM MANT-ATP in 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 0.1 wt% LDAO at 25 °C was measured. K_D values were derived from the increase in FP by addition of OEP21, normalized using OEP21-WT as a reference, and summarized above. The mean values of three independent FP experiments were used to determine the K_D values, with error bars representing the standard deviations of the fitted K_D values.

The binding affinities of the OEP21 variants were comparable to the wildtype (182 nM) and ranged from 127 nM to 336 nM for R33A and R174A, respectively. Consistent with the ATP titration experiments by NMR, deletion of loop five did not affect the high-affinity binding site within the β-barrel pore, as shown by the comparable K_D of 232 nM for OEP21-ΔLoop5. It follows that the removal of individual positive charges could be compensated by other positively charged residues in the proximity. This suggests that the binding mechanism could be rather promiscuous and depends mainly on the bulk electrostatic properties of the OEP21 pore.

3.1.10. Binding of metabolites is based on bulk electrostatic properties

Further FP-based binding experiments were performed in the presence of increased sodium chloride concentrations to confirm the assumption that the binding mechanism could be rather promiscuous and mainly based on bulk electrostatics properties. The results are summarized in FIGURE 53.

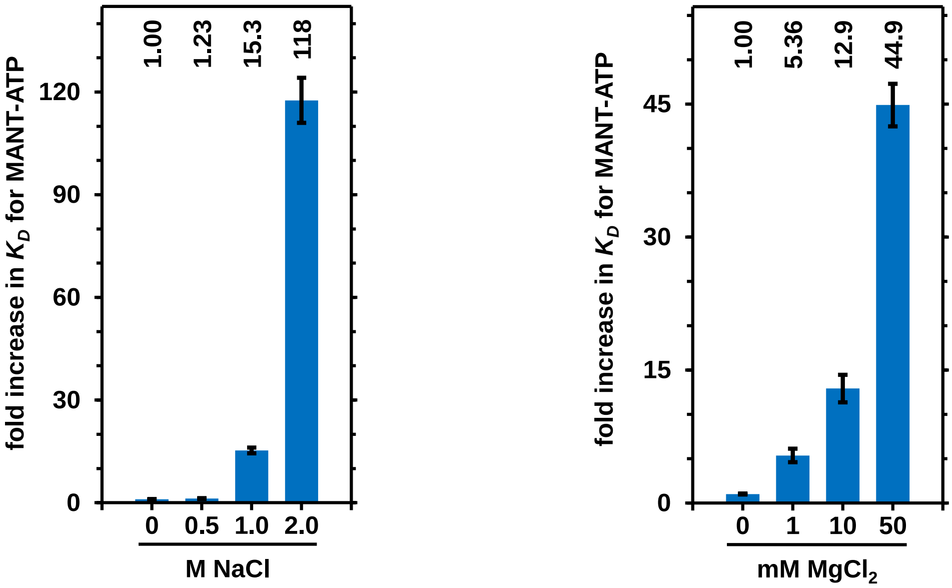


FIGURE 53: Effect of NaCl and $MgCl_2$ on the binding of MANT-ATP to OEP21

FP of stepwise titrated OEP21-WT to 200 nM MANT-ATP in 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 0.1 wt% LDAO at 25 °C was measured. Titration was performed in the presence of different concentrations of NaCl (left) and $MgCl_2$ (right), using a buffer without EDTA in the presence of $MgCl_2$. K_D values were derived from the increase in FP by addition of OEP21, normalized using OEP21-WT without additional NaCl or $MgCl_2$ as a reference, and summarized above. The mean values of three independent FP experiments were used to the determine the K_D values, with error bars representing the standard deviations of the fitted K_D values.

Because increased salt concentration can disturb the electrostatic interactions between protein and MANT-ATP, decreased binding affinity would be the expected consequence. Indeed, the binding affinity decreased by 118-fold at a concentration of 2 M NaCl, confirming the assumption of a binding mechanism based on bulk electrostatics properties. Since magnesium is known to bind to ATP with a K_D of 50 µM ± 10 µM [184], the negative charge density of ATP is reduced, suggesting a lower binding affinity. To verify this, FP-based binding experiments were performed in the presence of increasing $MgCl_2$ concentrations, as summarized in FIGURE 53. This assumption was confirmed, as a 45-fold decrease in the binding affinity of MANT-ATP was observed in the presence of 50 mM $MgCl_2$. FIGURE 54 shows the binding affinity of $MgCl_2$ for ATP, GAP, or NaPi as measured by ITC. This clearly demonstrates that $MgCl_2$ bound to ATP with a K_D of 54 µM, which is in excellent agreement with the reported K_D of 50 µM ± 10 µM [184]. Moreover, the binding of $MgCl_2$ to GAP or NaPi was much weaker with a K_D of 9 mM.

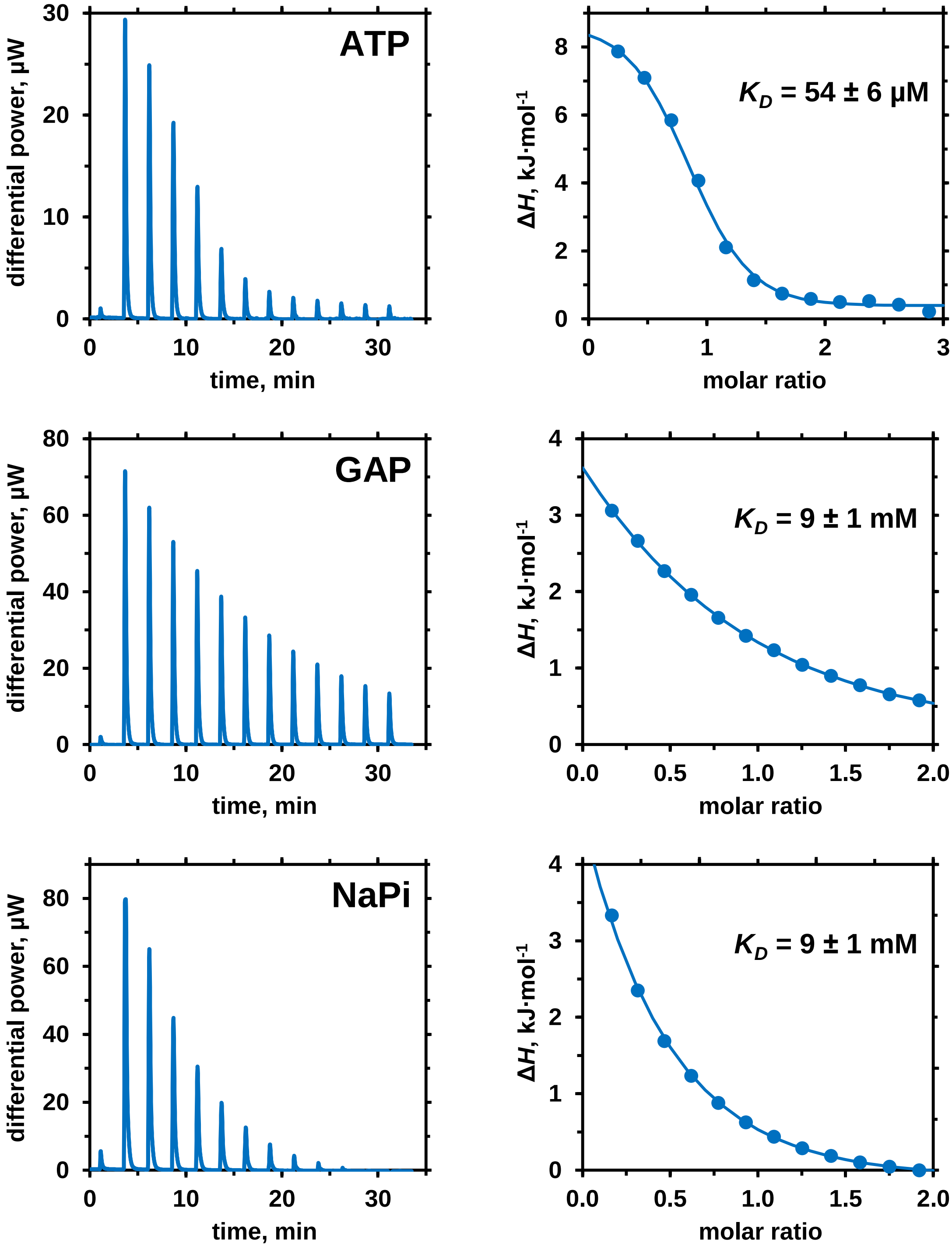


FIGURE 54: Binding affinities of $MgCl_2$ to ATP, GAP, or NaPi

$MgCl_2$ was titrated to ATP (above), GAP (center), or sodium phosphate (below) in 13 injection steps and the effect on differential power was measured at 25 °C using a buffer containing 10 mM HEPES pH 7.0 and 20 mM NaCl. Thus, the concentration in the cell was 1 mM for ATP, 10 mM for GAP, or 15 mM for NaPi, while the concentration of $MgCl_2$ in the syringe was 15 mM, 100 mM, or 150 mM, respectively. On the right, the relative enthalpy change over the molar ratio ($MgCl_2$ to ATP/GAP/NaPi) of each metabolite is illustrated, which was obtained from the raw data of differential power on the left. A K_D of 54 µM ± 6 µM for ATP and a K_D of 9 mM ± 1 mM for GAP and NaPi were obtained for the binding affinities to $MgCl_2$.

To better understand the conditions under which transport of triosephosphates through the OEP21 pore might occur, further FP binding experiments were performed in the presence of GAP, GAP plus $MgCl_2$, or NaPi. For this purpose, the respective metabolites were titrated stepwise to a preformed complex of MANT-ATP and OEP21-WT or OEP21-ΔLoop5, as illustrated in FIGURE 55.

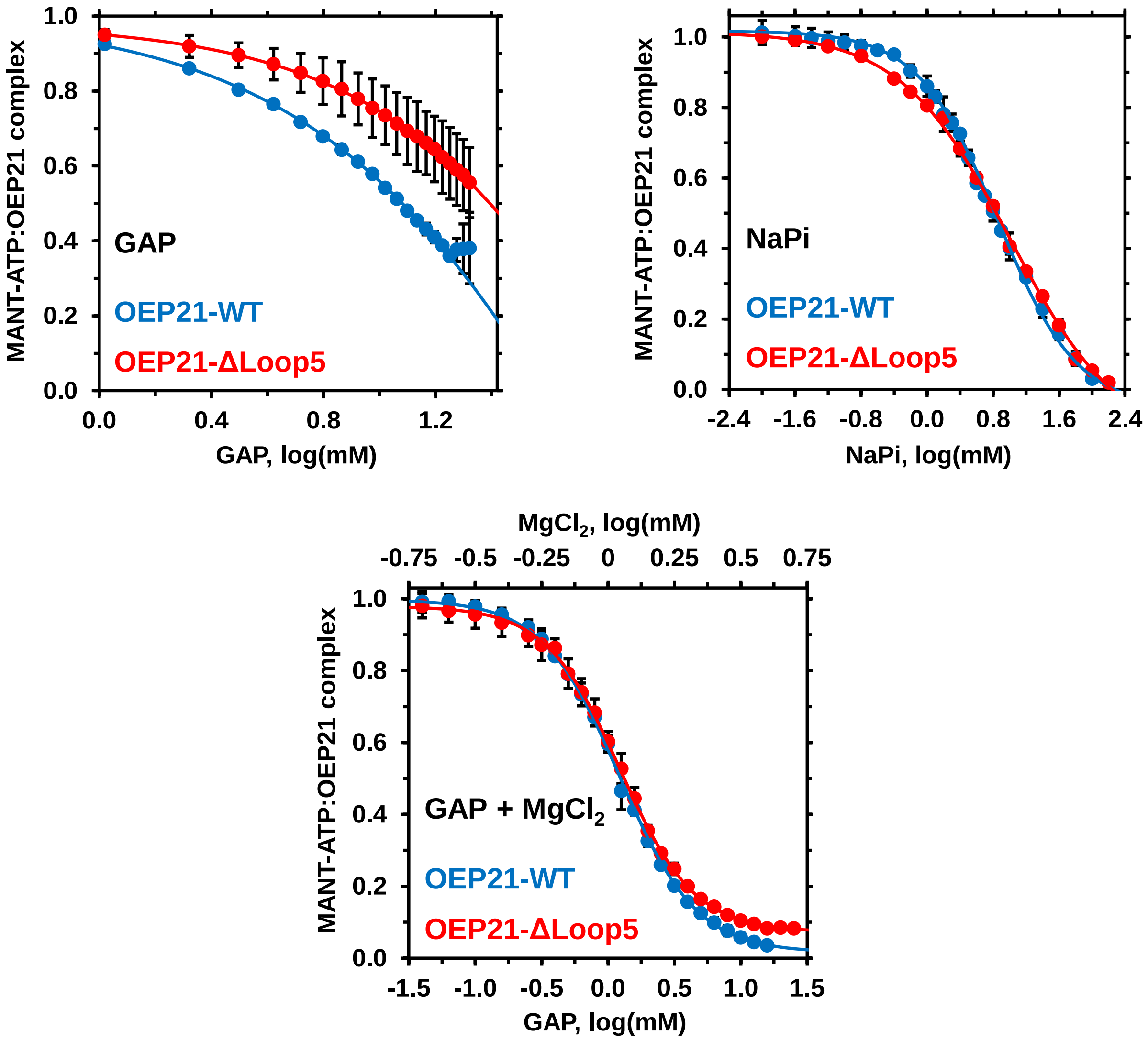


FIGURE 55: Competition experiments with GAP, NaPi, or GAP + $MgCl_2$

FP of stepwise titrated GAP (above left), NaPi (above right), or GAP + $MgCl_2$ (below) to a preformed complex of 1 µM MANT-ATP and 1 µM OEP21-WT (blue) or OEP21-ΔLoop5 (red) in 10 mM HEPES pH 7.0, 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 0.1 wt% LDAO at 25 °C was measured. No EDTA was used in the presence of $MgCl_2$. Instead, 100 µM $MgCl_2$ was added to the buffer before GAP + $MgCl_2$ were titrated. FP values were normalized (MANT-ATP:OEP21-WT complex and MANT-ATP:OEP21-ΔLoop5 complex, respectively), and IC_{50} values were derived for competitors GAP, NaPi, and GAP + $MgCl_2$, as shown in FIGURE 56. The indicated data points and error bars represent mean values and standard deviations derived from three independent FP experiments, except for the titration with GAP, for which two independent FP experiments were conducted.

By adding GAP to the preformed complex, dissociation of MANT-ATP by GAP, which acted as a competitor, was observable by decreasing FP. Determination of the IC_{50} value resulted in 12 mM for OEP21-WT and 27 mM for OEP21-ΔLoop5, as shown in FIGURE 56, indicating less effective competition by GAP for OEP21-ΔLoop5.

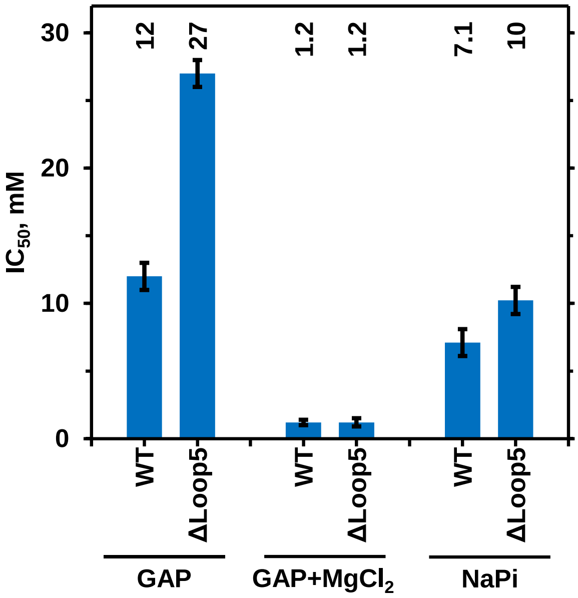


FIGURE 56: IC_{50} values for the competitors GAP, NaPi, and GAP + $MgCl_2$

Derived from FIGURE 55, the IC_{50} values for the competitors GAP, NaPi, and GAP + $MgCl_2$ were determined for OEP21-WT and OEP21-ΔLoop5. The error bars represent the standard deviations of the fitted IC_{50} values.

Since magnesium is released from the thylakoid lumen of chloroplasts into the stroma upon illumination [185] and magnesium transporters have been reported in the inner envelope of chloroplasts [186], it is very likely that increased concentrations of magnesium are present in the intermembrane space under light conditions. In FIGURE 54, the binding of $MgCl_2$ to ATP was confirmed, whereas the binding to GAP or NaPi was much weaker. Therefore, the negative charge density of ATP was reduced and resulted in a lower binding affinity for OEP21, as demonstrated in FIGURE 53. The presence of magnesium indicates a much more efficient displacement of ATP by GAP. This could be verified as shown in FIGURE 55 and FIGURE 56, where both constructs had a significantly lower IC_{50} value of 1.2 mM when GAP and $MgCl_2$ were added simultaneously. In addition, NaPi was tested as a competitor for ATP, as shown in FIGURE 55, where the phosphate concentration in plant cells varies between 1 mM and 10 mM [187]. Furthermore, NaPi was able to dissociate MANT-ATP from OEP21 with an IC_{50} of 7.1 mM for OEP21-WT and 10 mM for OEP21-ΔLoop5, consistent with the lower efficiency of GAP for OEP21-ΔLoop5.

The relatively high GAP concentrations required for ATP dissociation can be explained by partial non-specific binding to the LDAO micelle, as shown in FIGURE 57 by 1D NMR spectra.

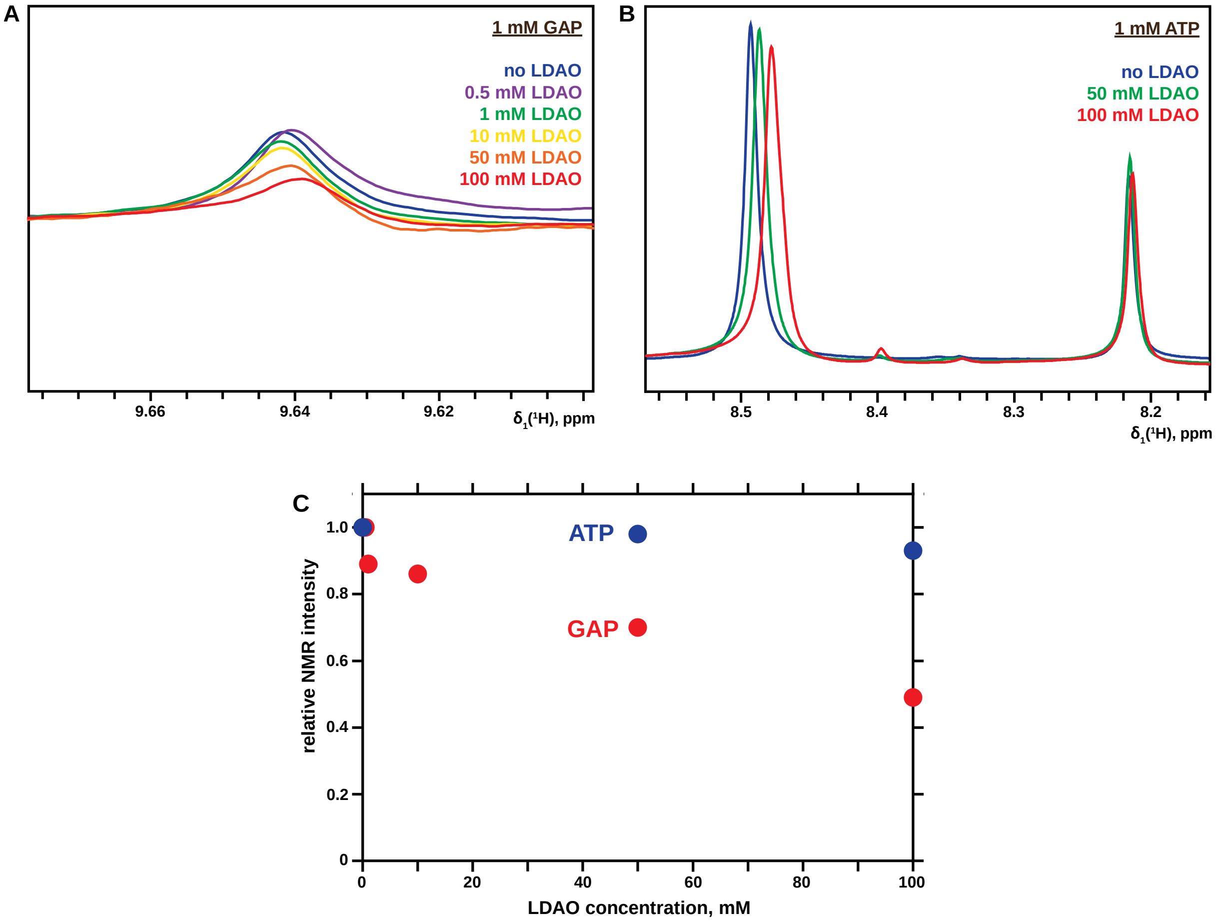


FIGURE 57: Probing the interaction of GAP and ATP with LDAO detergent micelles

^{1}H NMR spectra of the proton of the aldehyde moiety of GAP (^{1}H: 9.64 ppm) upon increasing LDAO concentration are superimposed in (A), where the decrease in signal intensity indicates a non-specific interaction. Similarly, ^{1}H NMR spectra of the protons in the nucleobase of ATP (^{1}H: 8.48 ppm) upon addition of LDAO are shown in (B), where no clear signal decrease is observed for ATP, indicating that ATP does not interact with the detergent micelle. The relative NMR peak intensities derived from the data in (A) and (B) are summarized in (C). All spectra were recorded in the presence of different LDAO concentration on an 800 MHz spectrometer with 1 mM GAP or 1 mM ATP in 10 mM NaPi pH 7.0, 20 mM NaCl, and 7 vol% D_2O at 303 K. This figure was provided by Prof. Dr. Franz Hagn.

In addition, the competition of ATP and GAP to OEP21 was investigated by NMR spectroscopy, as shown in FIGURE 58. For this purpose, 2D-[^{15}N,^{1}H]-TROSY experiments of ^{2}H,^{15}N-labeled OEP21-WT or OEP21-ΔLoop5 in complex with 1 mM ATP were recorded in the presence of increasing concentrations of GAP. As can be seen in FIGURE 40, the OEP21 spectrum showed strong line broadening of the NMR resonances in the presence of GAP.

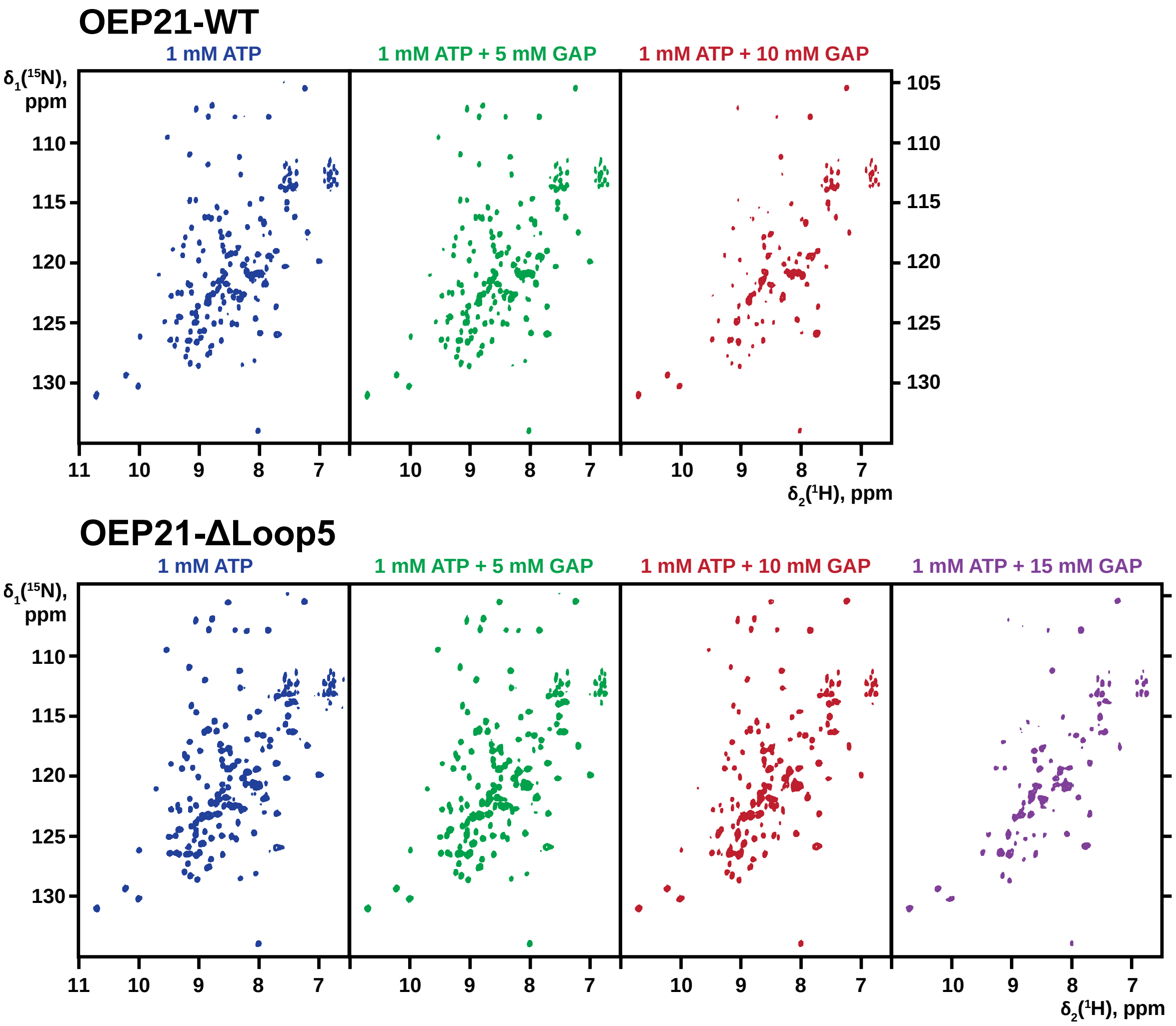


FIGURE 58: Competitive binding mode of ATP and GAP observed by NMR

2D-[^{15}N,^{1}H]-TROSY NMR spectra of OEP21-WT (above) and OEP21-ΔLoop5 (below) in complex with 1 mM ATP were recorded in the presence of increasing concentrations of GAP (5 mM, 10 mM, 15 mM) at 303 K. All spectra were recorded on a 950 MHz spectrometer with 400 µM ^{2}H,^{15}N-labelled protein in 20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT, and 7 vol% D_2O with 32 scans per increment and 160 complex points in the indirect ^{15}N dimension. The LDAO concentration was around 500 mM for OEP21-WT and 450 mM for OEP21- ΔLoop5. The OEP21 spectra showed strong line broadening of NMR resonances in the presence of higher GAP concentrations, indicating displacement of ATP by GAP, with higher GAP concentrations required in the case of OEP21-ΔLoop5 than OEP21-WT. This figure was provided by Prof. Dr. Franz Hagn.

Similarly, GAP was able to displace ATP from OEP21 at higher GAP concentrations, as observed by the appearance of strong line broadening of the NMR resonances. In agreement with the results of the FP experiments, higher GAP concentrations are required for ATP displacement in the case of OEP21-ΔLoop5 compared to OEP21-WT. These data suggest that phosphorylated metabolites use the same positively charged binding surface of OEP21 and that dissociation of the high-affinity binder ATP can be achieved at cellular solute concentration levels.

3.2. Bna-Tic40

3.2.1. Protein production and purification

The purity of the purchased proteins BnaA10-Tic40 and BnaC9-Tic40 was verified by SDS-PAGE, as shown in FIGURE 59, indicating highly pure protein samples. The BnaA10 construct consisted of 153 amino acids, whereas BnaC9 had only 151 amino acids. To compare the two proteins, a sequence alignment was performed between their amino acid sequences, which is shown in TABLE 28. A total of fifteen divergent positions were found between the two proteins.

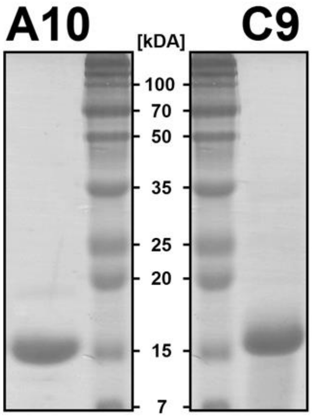


FIGURE 59: Purity of BnaA10 and BnaC9

Coomassie-stained SDS-PAGE gels of BnaA10-Tic40 and BnaC9-Tic40 protein purchased from Crelux GmbH (Martinsried, GER) show their purity and approximate molecular weight indicated by a protein standard marker. BnaA10 and BnaC9 have a theoretical molecular weight of 17,640 g·mol^{-1} and 17,522 g·mol^{-1}, respectively.

TABLE 28: Sequence alignment between BnaA10 and BnaC9

The amino acid sequences of BnaA10-Tic40 and BnaC9-Tic40 are aligned, with the numbering of amino acids shifted by two positions between the two. Deviations are colored red.

	305	315	325	335	345	355
A10	GGAGLSVEAL	EKMMEDPTVQ	KMVYPHLPEE	MRNPETFKWM	LKNPQYRQQL	QDMLNNMSGS
C9	--AGGSVEVL	EKMIEDPTFQ	KMLYPHLPEE	MRNPETFKWV	RKNPQYRQQL	QDMLNNMSES
	305	313	323	333	343	353
	365	375	385	395	405	415
A10	GEWDKRMTET	LKNFDLNSPE	VKQQFDQIGL	TPEEVISKIM	ENPDVAMAFQ	NPRVQAALME
C9	GEWDKRMTET	LKNFDPNSPE	VKQGFDQLGL	TPEEVISKIM	ENPDVSMAFQ	NPRVEAALMD
	363	373	383	393	403	413
	425	435	445	455		
A10	CSENPMNIMK	YQNDKEVMDV	FNKISQLFPG	MTG		
C9	CSENPMNIMK	YQNDKEVMDV	FNKISQLFPG	LTG		
	423	433	443	453		

3.2.2. Circular dichroism spectroscopy

To gain insight about the secondary structure of BnaA10 and BnaC9, far-UV spectra of both proteins were recorded, as shown in FIGURE 60. Both variants showed a spectrum representing an α-helical secondary structure, as a pure α-helical structure has a characteristic double minimum at 208 nm and 222 nm and a stronger maximum at 192 nm [176],[174]. In addition, the thermal melting behavior of both proteins was observed. No transition indicative of a melting temperature was observed for either variant. Interestingly, the α-helical secondary structure was retained after thermal melting up to 100 °C.

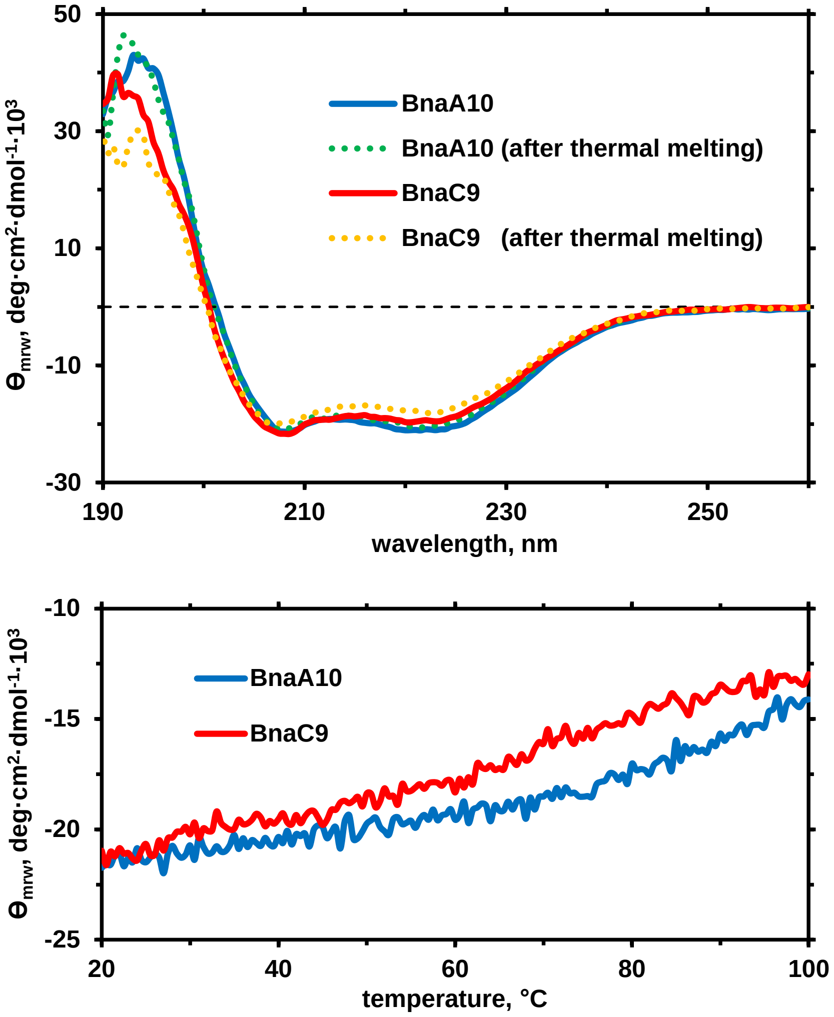


FIGURE 60: Far-UV CD spectra and thermal melting behavior of BnaA10 and BnaC9

The mean residue weight ellipticity Θ_{mrw} of BnaA10-Tic40 (blue) and BnaC9-Tic40 (red) at 20 °C and a wavelength from 190 nm to 260 nm is shown above. After thermal melting up to 100 °C, far-UV CD spectra of BnaA10 (green dotted) and BnaC9 (orange dotted) were recorded again. Shown below is the mean residue weight ellipticity Θ_{mrw} of BnaA10-Tic40 (blue) and BnaC9-Tic40 (red) during heating up from 20 °C to 100 °C at a wavelength of 222 nm. All spectra were recorded with 10 µM protein in 10 mM NaPi pH 6.0, 20 mM NaCl, 0.5 mM EDTA, and 2 mM β-mercaptoethanol.

Fractional helicities of 49 % and 45 % were determined for BnaA10 and BnaC9, respectively. EQUATION 15 was used for the calculation, with Θ_{mrw} at 222 nm was -20,879 deg·cm^2·dmol^{-1} and -19,358 deg·cm^2·dmol^{-1} for BnaA10 and BnaC9, respectively.

As can be seen from the NMR spectroscopy experiments in chapter 3.2.3, BnaA10 was selected for further experiments because its spectral quality was clearly better. It was reported that the presence of 10 % TFE was used for NMR structure determination of the C-terminal NP domain of Tic40 [50], so it was expected that the addition of TFE would further improve the spectra quality. Fluorinated alcohols and especially TFE are widely used to stabilize conformations of peptide fragments [188, 189]. To examine the effect of TFE on the secondary structure of BnaA10, far-UV CD spectra were recorded with different TFE concentrations, as shown in FIGURE 61.

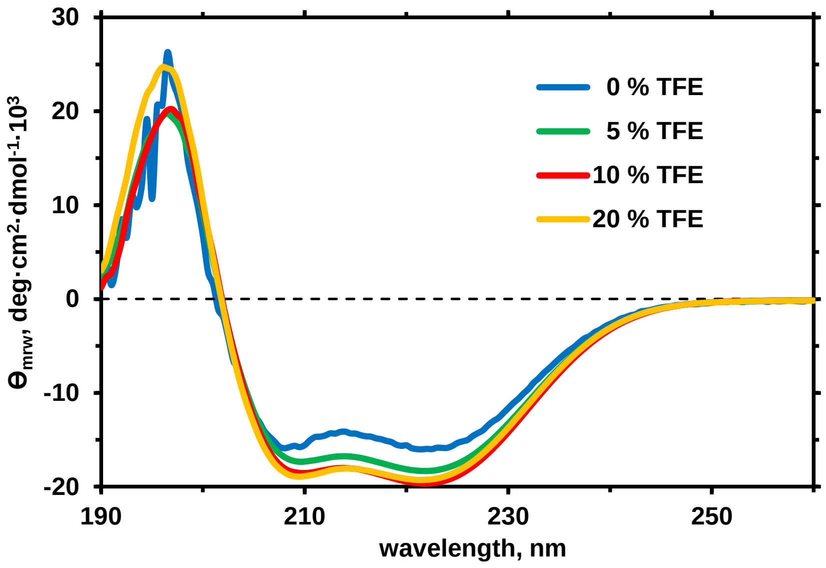


FIGURE 61: Far-UV CD spectra of BnaA10 with different TFE concentrations

The mean residue weight ellipticity Θ_{mrw} of BnaA10-Tic40 is shown at a wavelength from 190 nm to 260 nm, with spectra recorded with different concentrations of TFE: 0 vol% (blue), 5 vol% (green), 10 vol% (red), and 20 vol% (orange). All spectra were recorded with 20 µM protein in 10 mM NaPi pH 6.0, 20 mM NaCl, 0.5 mM EDTA, and 2 mM β-mercaptoethanol at 20 °C. Θ_{mrw} at 222 nm was 15,958 deg·cm^2·dmol^{-1} for 0 vol%, 18,330 deg·cm^2·dmol^{-1} for 5 vol%, 19,647 deg·cm^2·dmol^{-1} for 10 vol%, and 19255 deg·cm^2·dmol^{-1} for 20 vol%.

In summary, increasing TFE concentration resulted in increased fractional helicity. The fractional helicities determined using EQUATION 15 were 36 %, 42 %, 46 %, and 45 % for 0 vol%, 5 vol%, 10 vol%, and 20 vol% TFE, respectively. Therefore, a TFE concentration of 10 vol% was considered sufficient for further analysis.

3.2.3. NMR spectroscopy of BnaA10 and BnaC9

Initial 2D-[^{15}N,^{1}H]-TROSY NMR spectra of BnaA10 and BnaC9 were recorded at different temperatures (303 K, 313 K, 323 K) and different pH values (6.0, 7.0, 8.0) to identify a setup with satisfactory spectral quality, as shown in FIGURE 62.

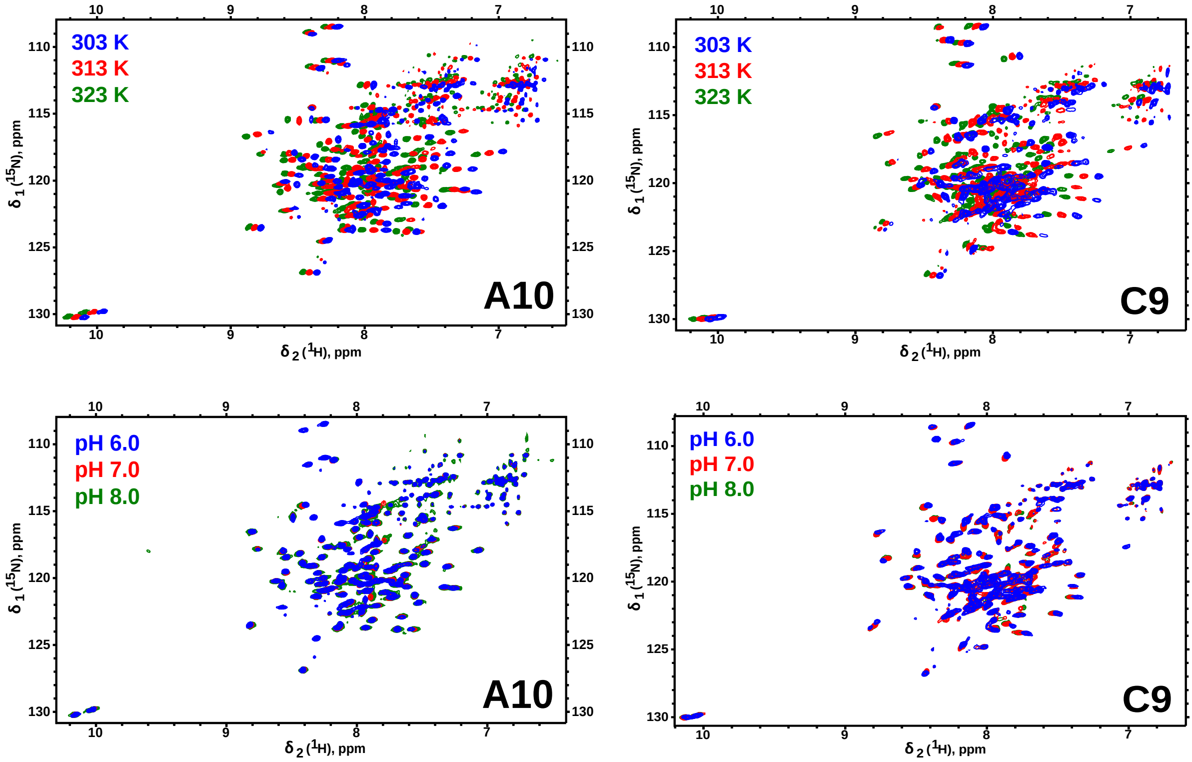


FIGURE 62: NMR spectra of BnaA10 and BnaC9 at different temperatures and pH values

2D-[^{15}N,^{1}H]-TROSY NMR spectra of BnaA10-Tic40 (left) and BnaC9 (right) at 303 K (blue), 313 K (red), and 323 K (green) are shown above. Shown below are NMR spectra of BnaA10-Tic40 (left) and BnaC9 (right) at pH 6.0 (blue), 7.0 (red), and 8.0 (green). All spectra were recorded on an 800 MHz spectrometer with 200 µM ^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, and 7 vol% D_2O with 40 scans per increment and 256 complex points in the indirect ^{15}N dimension.

Since the NMR spectra recorded at 303 K and at pH 6.0 had the best spectral quality, the spectra of BnaA10 and BnaC9 were compared under these conditions, which is shown in FIGURE 63. In summary, the spectral quality of both constructs was not satisfactory for more detailed NMR analysis, with BnaC9 exhibiting poor spectral quality.

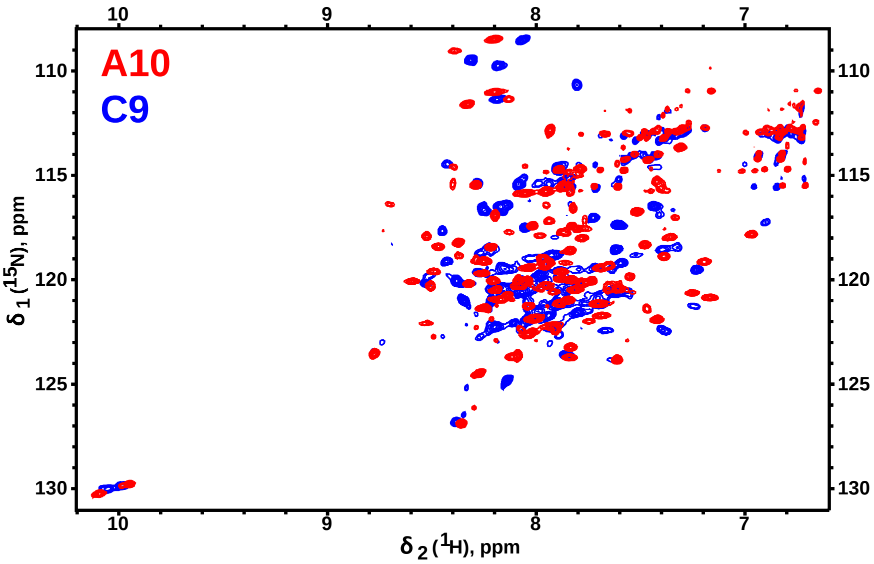


FIGURE 63: Comparison of the NMR spectra of BnaA10 and BnaC9

Comparison of the 2D-[^{15}N,^{1}H]-TROSY NMR spectra from BnaA10-Tic40 (red) and BnaC9 (blue). Both spectra were recorded on an 800 MHz spectrometer with 200 µM ^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, and 7 vol% D_2O at 303 K with 40 scans per increment and 256 complex points in the indirect ^{15}N dimension.

To improve the spectral quality, TFE was added to the NMR samples. FIGURE 64 compares the 2D-[^{15}N,^{1}H]-TROSY NMR spectra of BnaA10 and BnaC9 in the presence of 10 vol% TFE.

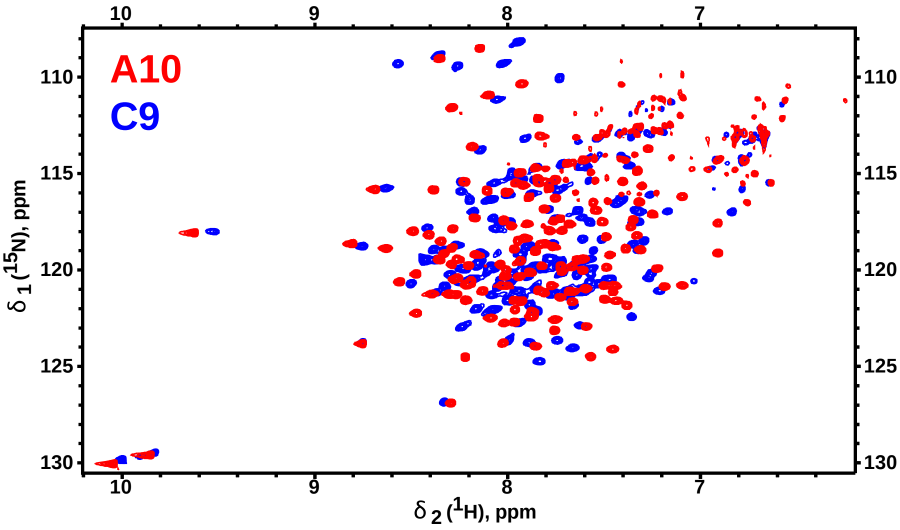


FIGURE 64: Comparison of the NMR spectra of BnaA10 and BnaC9 with 10 % TFE

Comparison of the 2D-[^{15}N,^{1}H]-TROSY NMR spectra from BnaA10-Tic40 (red) and BnaC9 (blue). Both spectra were recorded on a 600 MHz spectrometer with 400 µM ^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, 7 vol% D_2O, and 10 vol% TFE-*d*$_3$ at 303 K with 32 scans per increment and 256 complex points in the indirect ^{15}N dimension.

As expected, the addition of TFE significantly improved the spectral quality of both constructs. In agreement with the results obtained without TFE, the spectral quality of BnaA10 was better than that of BnaC9. In the case of BnaA10, the positions and intensities of the signals are very sensitive to changes in structure and dynamics. Comparison of the NMR spectra of BnaA10 and BnaC9 revealed many changes in signal positions and signal intensities originating from only fifteen amino acids deviating between both constructs, indicating altered structure and dynamics. Since the spectral quality of BnaC9 was not sufficient for a more detailed NMR analysis, BnaA10 was selected for further experiments.

In addition, 2D-[^{15}N,^{1}H]-TROSY NMR spectra of BnaA10 were recorded in the presence of various TFE concentrations, as shown in FIGURE 65, where already 10 vol% TFE were sufficient to improve the NMR spectral quality significantly.

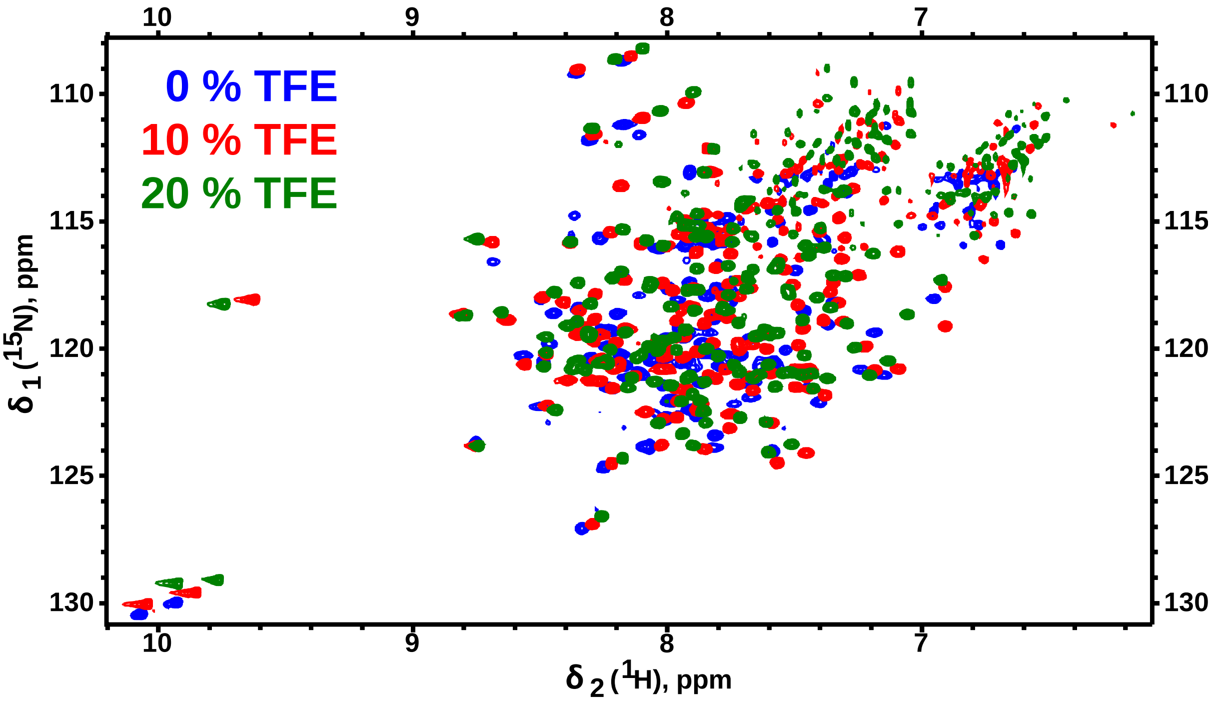


FIGURE 65: NMR spectra of BnaA10 at different TFE concentrations

Comparison of the 2D-[^{15}N,^{1}H]-TROSY NMR spectra of BnaA10-Tic40 in the presence of 0 vol% (blue), 10 vol% (red), and 20 vol% TFE-*d*$_3$ (green). All spectra were recorded on a 600 MHz spectrometer with 400 µM ^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, and 7 vol% D_2O at 303 K with 32 scans per increment and 256 complex points in the indirect ^{15}N dimension.

In general, systemic TFE-induced shift changes lead to the gradual formation of helical structures that follow the trend of a monotonic upfield drift in the hydrogen dimension [50, 190].

3.2.4. NMR structural studies on BnaA10-Tic40

To investigate the structural state of BnaA10, a set of 3D triple resonance experiments consisting of HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, and HN(CO)CACB were recorded in the presence of 10 vol% TFE-d_3 to obtain sequence-specific resonance assignments. FIGURE 66 shows assigned resonances based on the 2D-[^{15}N,^{1}H]-HSQC NMR spectrum.

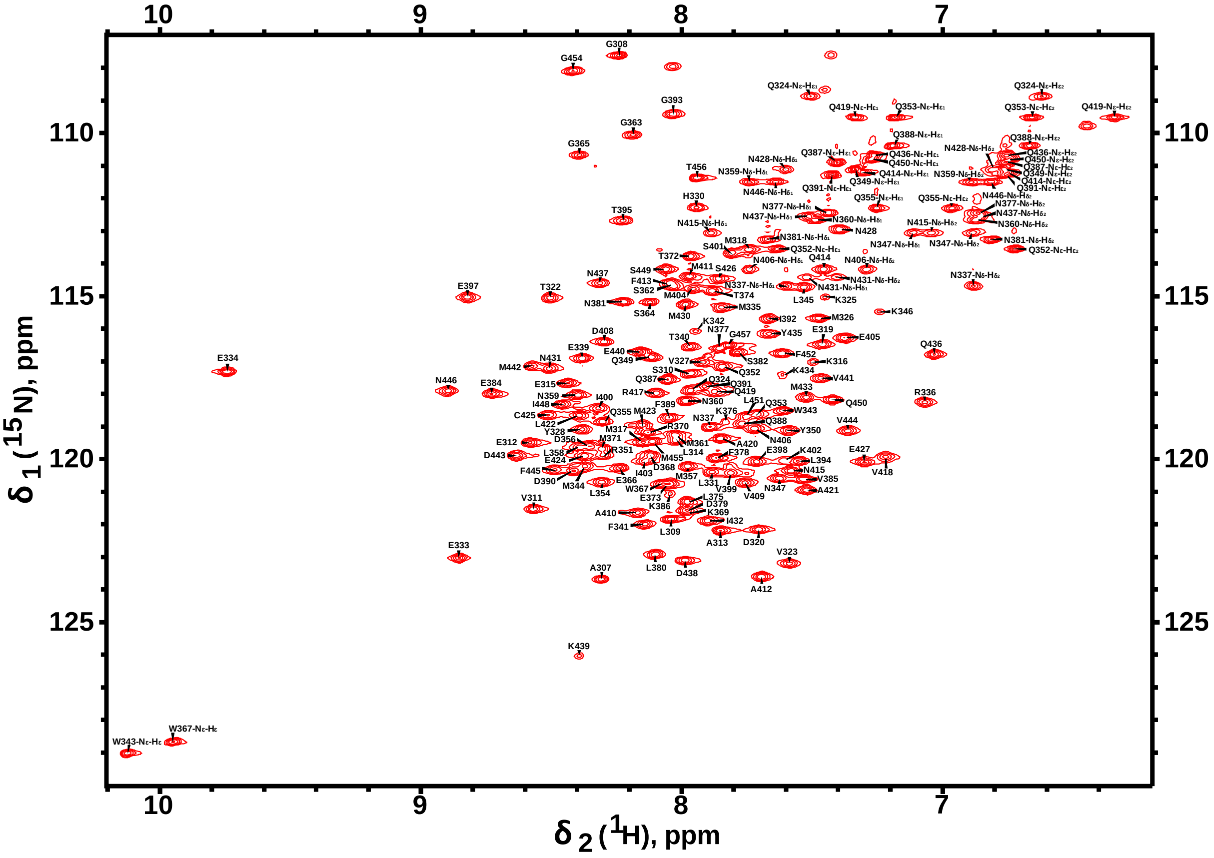


FIGURE 66: Assignment of the 2D-[^{15}N,^{1}H]-HSQC NMR spectrum of BnaA10

The 2D-[^{15}N,^{1}H]-HSQC NMR spectrum of BnaA10-Tic40 with assigned resonances is shown. The spectrum was recorded on a 600 MHz spectrometer with 1.1 mM ^{13}C,^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, 7 vol% D_2O, and 10 vol% TFE-d_3 at 308 K with 8 scans per increment and 256 complex points in the indirect ^{15}N dimension.

The side chain amide resonances of asparagine and glutamine were assigned using the HN(CO)CACB experiment. The ^{13}C chemical shift information obtained from backbone resonance assignment was used to identify secondary structure elements in BnaA10, as shown in FIGURE 67.

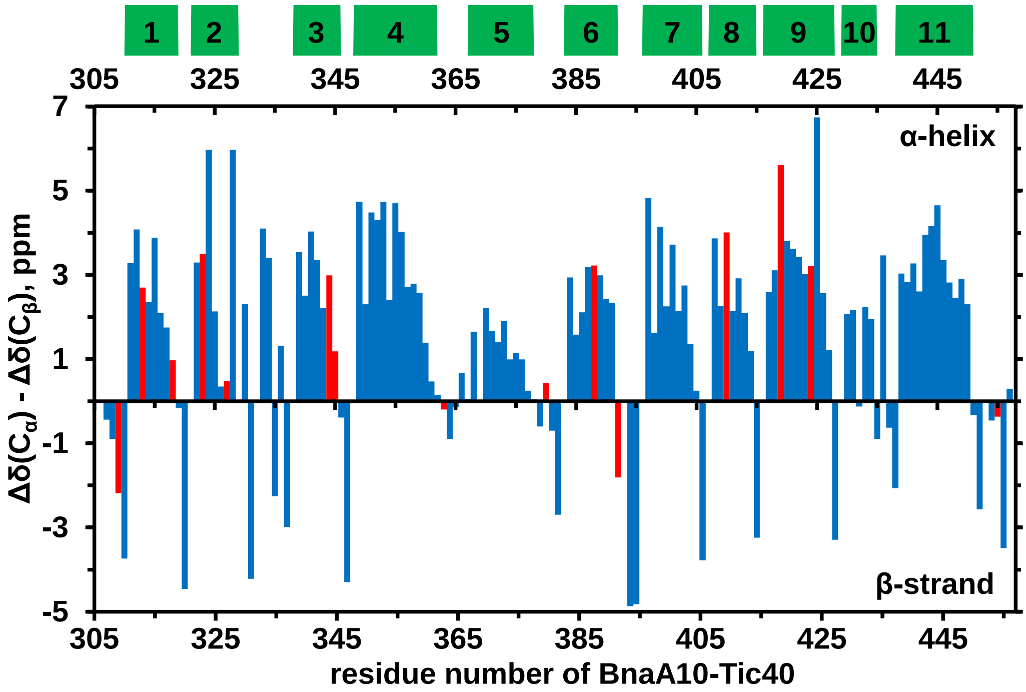


FIGURE 67: Secondary structure estimation of BnaA10 using ^{13}C chemical shifts

The secondary structure of BnaA10-Tic40 was estimated using ^{13}C secondary chemical shift information obtained from backbone resonance assignment. $\Delta\delta(C_\alpha) - \Delta\delta(C_\beta)$ was calculated according to EQUATION 6, where a positive value corresponds to an α-helical and a negative value to a β-sheet secondary structure. Therefore, eleven α-helical secondary structure elements were identified, visualized by the green boxes above. Positions that exhibited amino acid sequence variations between BnaA10 and BnaC9 are highlighted in red.

However, only α-helical secondary structure elements were identified, which is consistent with CD spectroscopy data. Further resonances were assigned using CC(CO)NH, H(CCO)NH, HNHA, and HNHB experiments. In addition, diastereotopic H_β protons were assigned if possible. FIGURE 68, FIGURE 69, and FIGURE 70 show sections of the [^{13}C,^{1}H]-HSQC spectrum with assigned carbon side chain resonances. TABLE 36 in the appendix lists all chemical shifts of the assigned resonances.

NOE contacts were derived from a set of 3D-NOESY spectra consisting of HNH-NOESY, NNH-NOESY, CNH-NOESY, HCH-NOESY, and CCH-NOESY experiments. Unfortunately, no 3D structure determination could be performed due to missing or overlapping long-range NOE contacts.

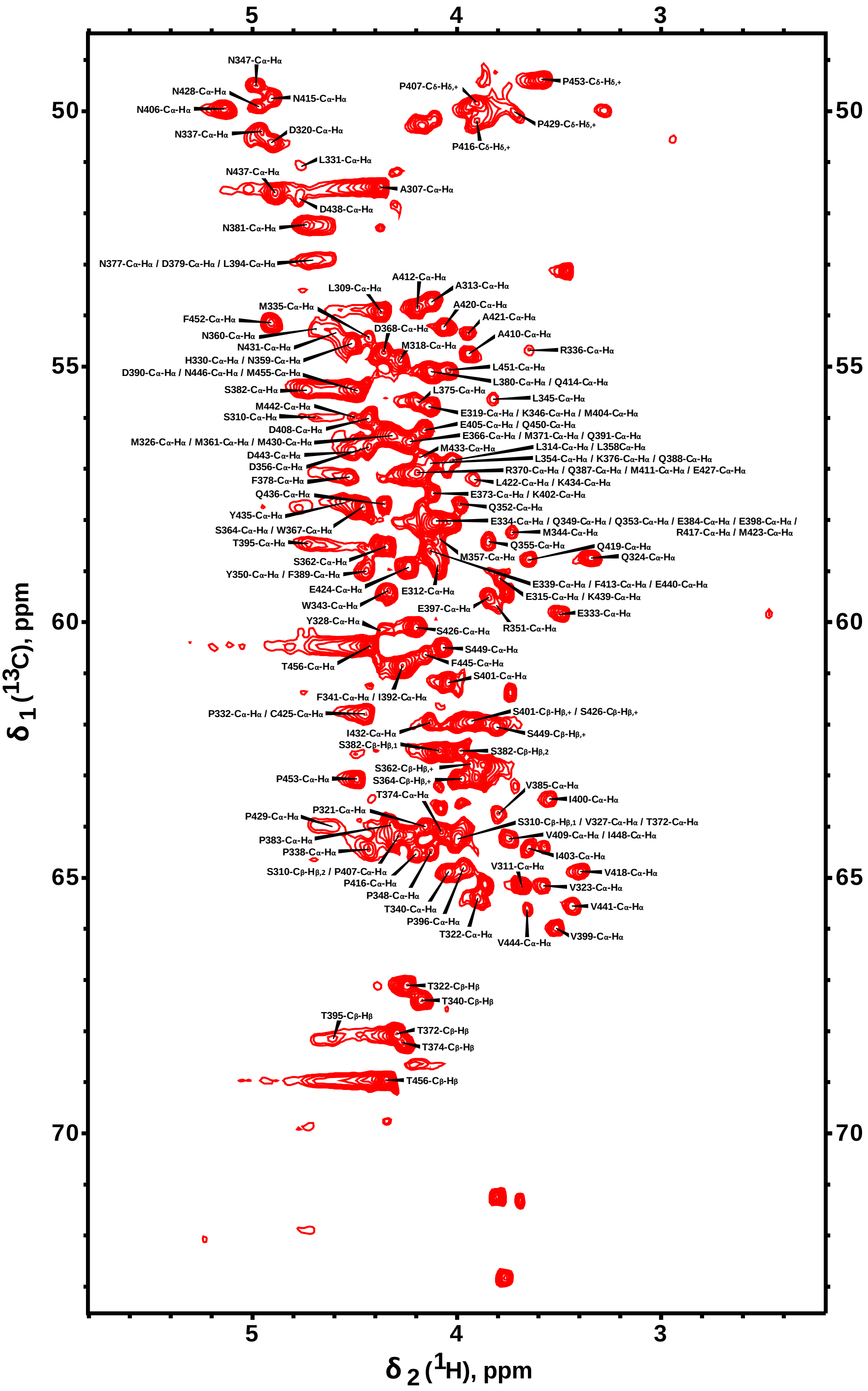


FIGURE 68: Assignment of the 2D-[^{13}C,^{1}H]-HSQC NMR spectrum of BnaA10, Part 1

The 2D-[^{13}C,^{1}H]-HSQC NMR spectrum of BnaA10-Tic40 with assigned resonances is shown. The spectrum was recorded on a 600 MHz spectrometer with 1.1 mM ^{13}C,^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, 7 vol% D_2O, and 10 vol% TFE-*d*$_3$ at 308 K with 8 scans per increment and 320 complex points in the indirect ^{13}C dimension.

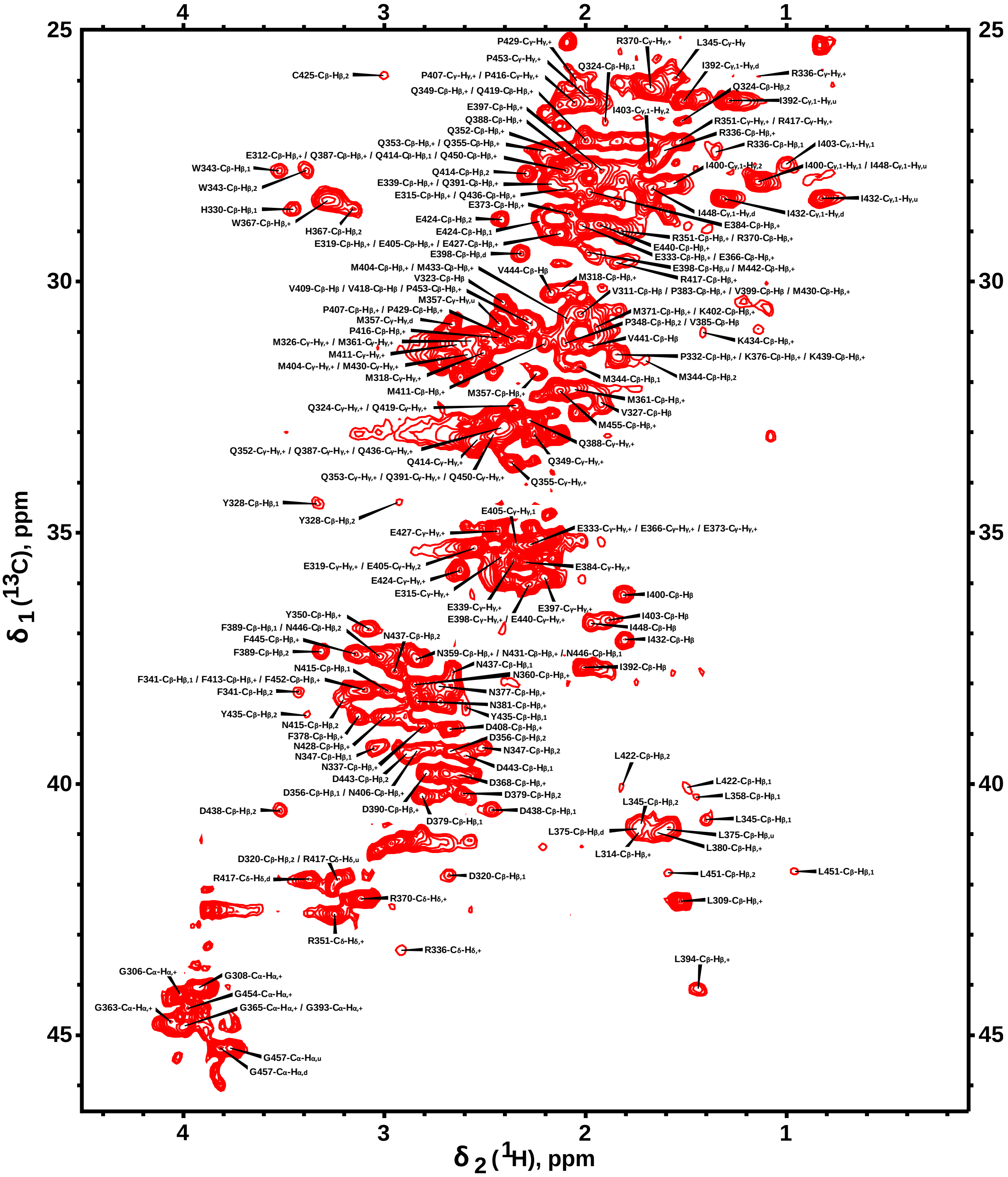


FIGURE 69: Assignment of the 2D-[^{13}C,^{1}H]-HSQC NMR spectrum of BnaA10, Part 2

The 2D-[^{13}C,^{1}H]-HSQC NMR spectrum of BnaA10-Tic40 with assigned resonances is shown. The spectrum was recorded on a 600 MHz spectrometer with 1.1 mM ^{13}C,^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, 7 vol% D$_2$O, and 10 vol% TFE-*d*$_3$ at 308 K with 8 scans per increment and 320 complex points in the indirect ^{13}C dimension.

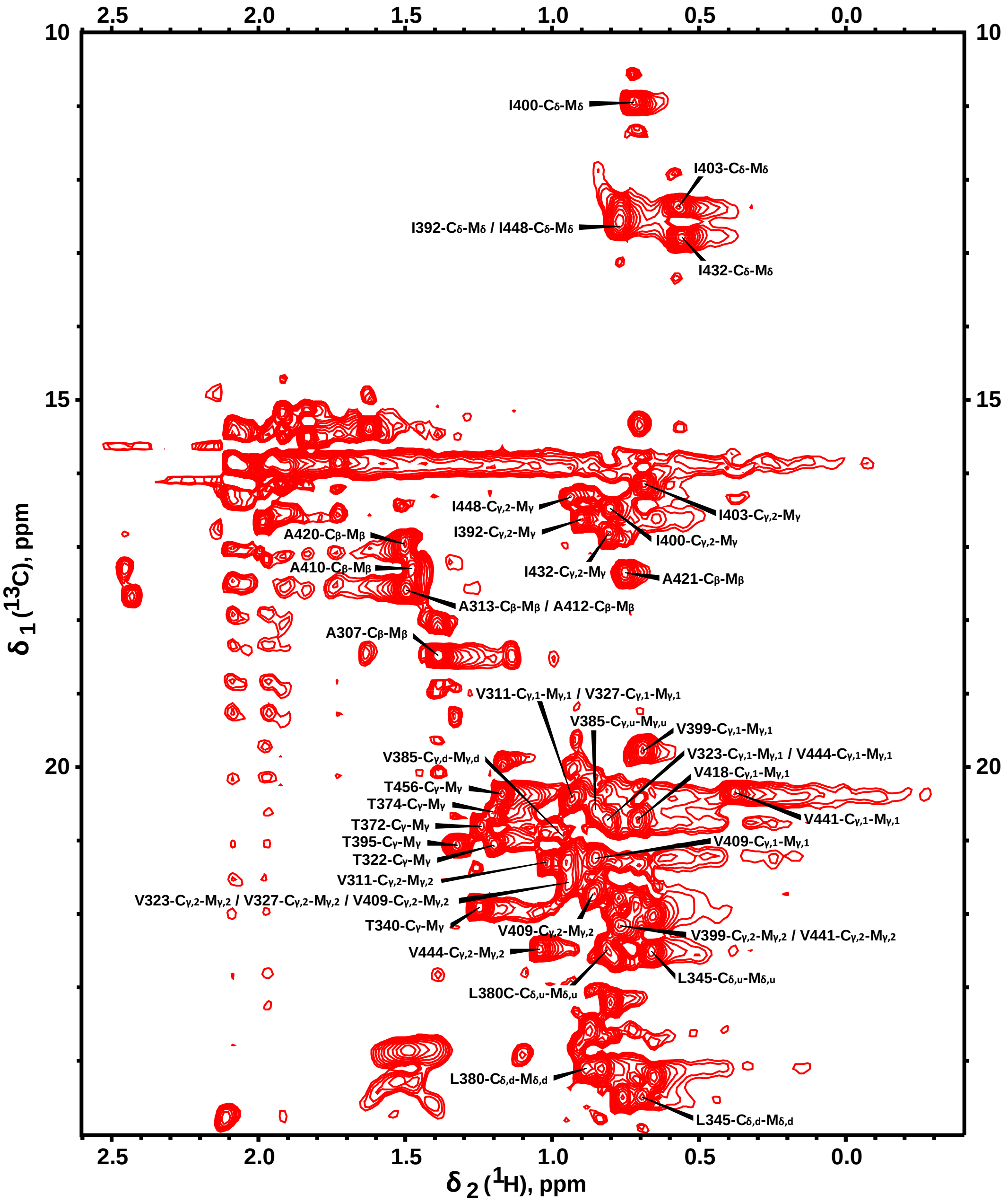


FIGURE 70: Assignment of the 2D-[^{13}C,^{1}H]-HSQC NMR spectrum of BnaA10, Part 3

The 2D-[^{13}C,^{1}H]-HSQC NMR spectrum of BnaA10-Tic40 with assigned resonances is shown. The spectrum was recorded on a 600 MHz spectrometer with 1.1 mM ^{13}C,^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, 7 vol% D_2O, and 10 vol% TFE-*d*$_3$ at 308 K with 8 scans per increment and 320 complex points in the indirect ^{13}C dimension.

3.3. ADAM17 and iRhom2

3.3.1. Protein production and purification of ADAM17-TM

Three different constructs were generated to produce the transmembrane domain of ADAM17, and all constructs were co-expressed with Get3-MBP. The amino acid sequences can be found in the appendix in TABLE 32 and TABLE 33, respectively. Since Get3 contains an enormous hydrophobic groove [114], the idea was that the transmembrane domain of ADAM17 could be embedded within Get3, allowing soluble expression of the transmembrane domain.

The first construct ADAM17-TM-TEV-Z-tag contained an N-terminal His$_{10}$-tag followed by a Z-tag and a sequence for TEV cleavage between the Z-tag and the C-terminal ADAM17-TM. Indeed, co-expression of Get3-MBP enabled the expression of ADAM17-TM, as clearly shown in FIGURE 71 by two intense bands on the SDS-PAGE gel, indicating a high expression level of ADAM17-TM-TEV-Z-tag and Get3-MBP.

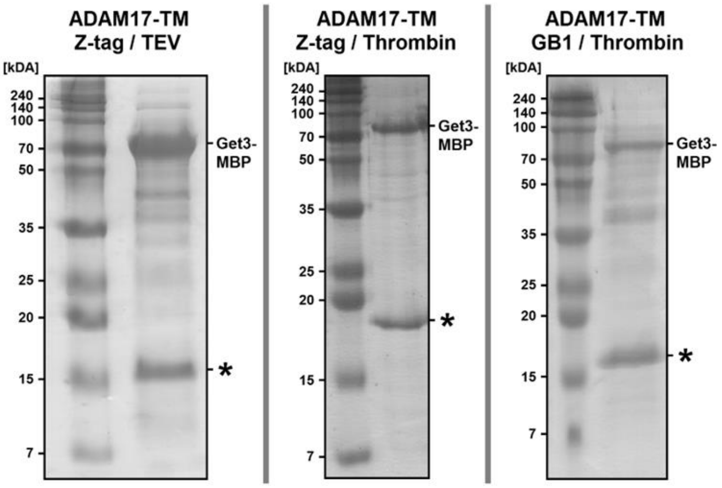


FIGURE 71: Expression of the ADAM17-TM constructs

Coomassie-stained SDS-PAGE gels of ADAM17-TM constructs expressed by *E. coli* BL21(DE3) in LB medium 4 h after induction with 1 mM IPTG at 37 °C are shown, while Get3-MBP was co-expressed. On the left, the expression of ADAM17-TM combined with a Z-tag and a sequence for TEV cleavage is shown. In the middle, the expression of the construct containing a sequence for thrombin cleavage instead of one for TEV is shown. On the right, the expression of ADAM17-TM combined with GB1 and a sequence for thrombin cleavage is shown. The corresponding protein band of the construct is marked by an asterisk.

To purify soluble expressed ADAM17-TM, Ni-NTA affinity chromatography was performed in the presence of DPC, as shown exemplarily in FIGURE 72. This procedure yielded unsatisfactory results because Get3-MBP could not be separated from ADAM17-TM via Ni-NTA affinity chromatography.

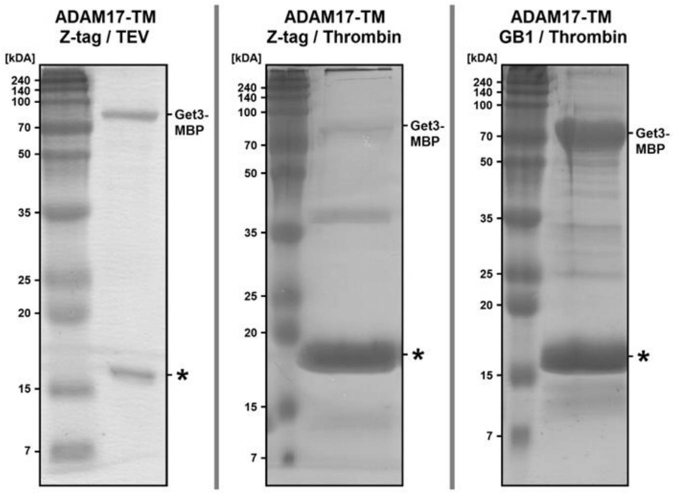


FIGURE 72: ADAM17-TM constructs purified with Ni-NTA affinity chromatography

Coomassie-stained SDS-PAGE gels of the ADAM17-TM constructs purified by Ni-NTA affinity chromatography are shown. From left to right, the elution fractions of ADAM17-TM-TEV-Z-tag, ADAM17-TM-Z-tag, and ADAM17-TM-GB1 are shown, while a buffer containing 50 mM Tris pH 8.0, 500 mM NaCl, 5 mM BME, 500 mM imidazole, and 0.2 wt% DPC was used to elute the proteins from the Ni-NTA columns.

The second construct, ADAM17-TM-Z-tag, contained a sequence for thrombin cleavage instead one for TEV because thrombin is more suitable for cleavage in the presence of detergents such as DPC. The expression and purification of ADAM17-TM-Z-tag was comparable to that with the TEV cleavage site, as shown in FIGURE 71 and FIGURE 72. However, after elution of the protein from Ni-NTA affinity chromatography, another protein band of approximately 40 kDa was present. This could be a proteolytically digested version of Get3-MBP, whereas Get3 without MBP has a molecular weight of 39.4 kDa.

In addition, SEC of ADAM17-TM-Z-tag / Get3-MBP in the presence of DPC showed the formation of high molecular weight species greater than 500 kDa, as shown in FIGURE 73. Analysis of these high molecular weight species by SDS-PAGE revealed that ADAM17-TM-Z-tag and Get3-MBP occurred together, whereas ADAM17-TM was most likely embedded within Get3. Therefore, DPC was apparently unable to dissociate the transmembrane domain of ADAM17 from Get3 and solubilize ADAM17-TM. Interestingly, the enormous size of ADAM17-TM-Z-tag / Get3-MBP could not be explained by the sizes of the individual molecules, because ADAM17-TM-Z-tag and Get3-MBP had a molecular weight of 20.8 kDa and 86.3 kDa, respectively. Considering that Get3 occurs as a homodimer [114], the expected size of the complex is still smaller than the observed size, indicating the formation of larger assemblies of ADAM17-TM-Z-tag with Get3-MBP. To exclude the Z-tag as an interfering factor, ADAM17-TM-Z-tag was cleaved by thrombin and SEC was performed, as shown in FIGURE 74.

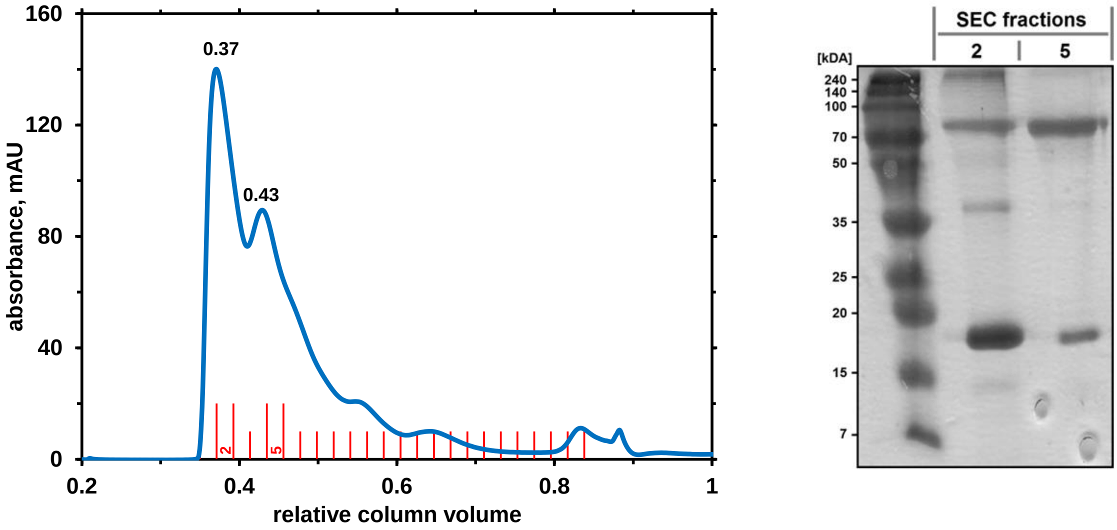


FIGURE 73: SEC chromatogram of ADAM17-TM-Z-tag

On the left, ADAM17-TM-Z-tag was applied on a Superdex 200 Increase 10/300 GL SEC column, where the SEC running buffer contained 20 mM NaPi pH 6.0, 50 mM NaCl, 1 mM EDTA, 5 mM DTT, and 0.1 wt% DPC. The first major peak eluted at 0.37 CV, which corresponded to the void volume of the SEC column. The second major peak eluted at 0.43 CV, corresponding to a molecular weight of approximately 536 kDa for this species. On the right, a Coomassie-stained SDS-PAGE gel of fractions two and five is shown, demonstrating the presence of ADAM17-TM-Z-tag (20.8 kDa) and Get3-MBP (86.3 kDa) in addition to minor impurities.

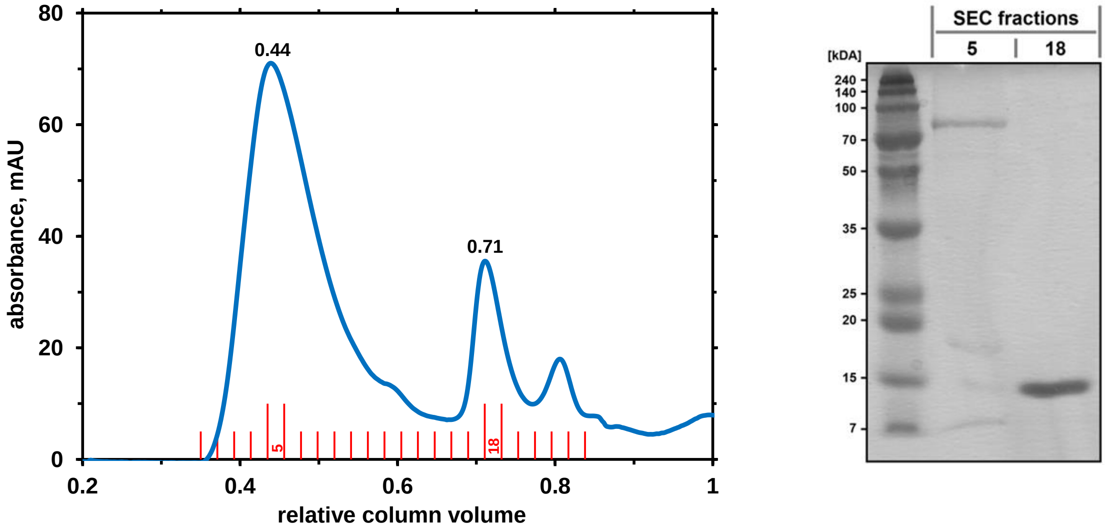


FIGURE 74: SEC chromatogram of thrombin-cleaved ADAM17-TM-Z-tag

On the left, thrombin-cleaved ADAM17-TM-Z-tag was applied on a Superdex 200 Increase 10/300 GL SEC column, and the SEC running buffer contained 20 mM NaPi pH 6.0, 50 mM NaCl, 1 mM EDTA, 5 mM DTT, and 0.1 wt% DPC. The three major peaks eluted at 0.44 CV, 0.71 CV, and 0.81 CV, corresponding to molecular weights of approximately 483 kDa, 30.1 kDa, and 10.8 kDa, respectively. On the right, a Coomassie-stained SDS-PAGE gel of the fractions five and eighteen is shown. Fraction five indicated the presence of ADAM17-TM-Z-tag (20.8 kDa) and its cleavage products (15.7 kDa and 5.05 kDa) as well as Get3-MBP (86.3 kDa), whereas fraction eighteen contained only the Z-tag (15.7 kDa).

Analysis of the SEC fractions by SDS-PAGE revealed that the Z-tag could be separated from the ADAM17-TM / Get3-MBP complex by SEC. In addition, thrombin cleavage was not complete because non-cleaved ADAM17-TM-Z-tag was present in fraction five in addition to small amounts of the Z-tag. Comparing the intensities of the protein bands on the SDS-PAGE gel of thrombin-cleaved Z-tag in fraction eighteen and ADAM17-TM in fraction five, it is noticeable that the corresponding intensity of ADAM17-TM was significantly weaker, indicating a loss of ADAM17-TM. Since the estimated sizes from SEC were inconclusive, an additional electrospray ionization mass spectrometry (ESI-MS) analysis was performed with the fractions five and eighteen. In the case of fraction eighteen, a molecular weight of 15,617 $g \cdot mol^{-1}$ was determined, which agrees perfectly with the theoretical one of 15,619 $g \cdot mol^{-1}$ for the His_{10}-Z-tag without the first methionine. Also, the obtained molecular weight of 5,045 $g \cdot mol^{-1}$ in fraction five could be unambiguously assigned to ADAM17-TM, with a theoretical value of 5,047 $g \cdot mol^{-1}$. In addition to ADAM17-TM, other species with molecular weights of 2,717 $g \cdot mol^{-1}$, 3,864 $g \cdot mol^{-1}$, and 18,370 $g \cdot mol^{-1}$ were found in fraction five that could not be reasonably assigned and possibly represented degradation products of Get3-MBP.

Since GB1 can easily refold after denaturation with e. g. GdmCl, the third construct ADAM17-TM-GB1 with GB1 instead of the Z-tag was generated. The expression and purification of ADAM17-TM-GB1 was comparable to the previous constructs, as shown in FIGURE 71 and FIGURE 72. FIGURE 75 shows the purification of ADAM17-TM-GB1 via Ni-NTA affinity chromatography in a DPC-containing buffer visualized by SDS-PAGE, where the elution fraction contained ADAM17-TM-GB1 and Get3-MBP in addition to various minor impurities. After dialysis of the elution fraction, ADAM17-TM-GB1 was proteolytically digested with thrombin. Following digestion, a second Ni-NTA affinity chromatography was conducted, separating GB1 containing the His_{10}-tag from ADAM17-TM. Therefore, ADAM17-TM embedded within Get3-MBP was found in the flow-through and wash fractions.

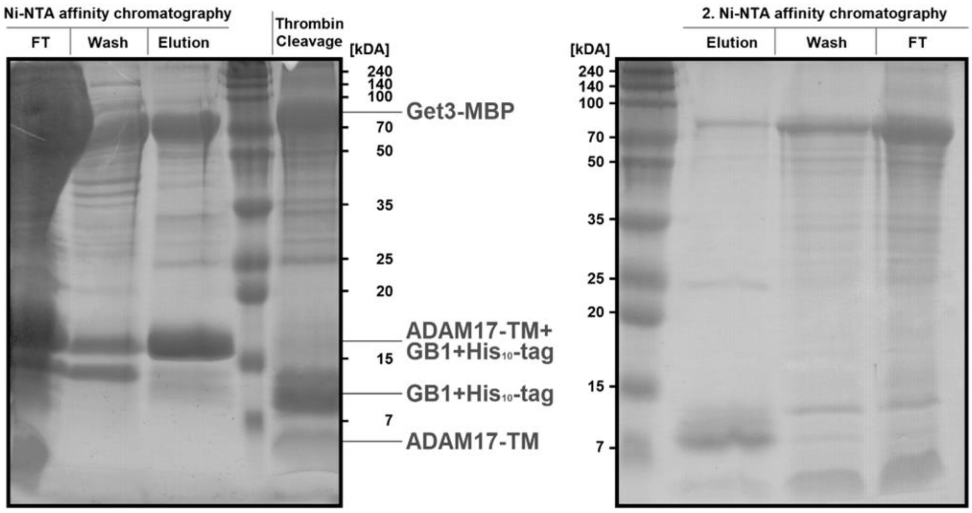


FIGURE 75: Purification of ADAM17-TM-GB1 via Ni-NTA affinity chromatography

On the left, a Coomassie-stained SDS-PAGE gel of ADAM17-TM-GB1 purified via Ni-NTA affinity chromatography is shown, where a buffer containing 50 mM Tris pH 8.0, 500 mM NaCl, 5 mM BME, 10 mM imidazole, and 0.2 wt% DPC was used. In the first step, the applied protein sample was allowed to flow through the column (FT) after incubation with the Ni-NTA slurry for 30 min. The column was then washed with buffer before the proteins were eluted from the column with a buffer containing a total of 500 mM imidazole. The elution fraction was then dialyzed against a buffer containing 20 mM Tris pH 8.0, 50 mM NaCl, 5 mM BME, and 5 mM $CaCl_2$ before ADAM17-TM-GB1 was cleaved by thrombin. On the right, a Coomassie-stained SDS-PAGE gel shows purification of thrombin-cleaved ADAM17-TM-GB1 via a second Ni-NTA affinity chromatography, using the same buffer and procedure as above. Because the His_{10}-tag was N-terminal of GB1, ADAM17-TM was found in the flow-through and wash fractions where GB1 was bound to the column.

ADAM17-TM-GB1 before thrombin cleavage and after thrombin cleavage, followed by a second Ni-NTA affinity chromatography, was applied on SEC, as shown in FIGURE 76 and FIGURE 77, respectively. The SEC chromatogram of ADAM17-TM-GB1 was comparable to that of ADAM17-TM-Z-tag, as the major species was approximately 536 kDa in size and contained predominantly ADAM17-TM-GB1 and Get3-MBP. Because Get3-MBP could not be separated from thrombin-cleaved ADAM17-TM-GB1, the SEC chromatogram of this sample showed the presence of both proteins, and the ratio of Get3-MBP appeared to decrease with an increasing elution volume. In addition, various minor impurities were present in all SEC fractions.

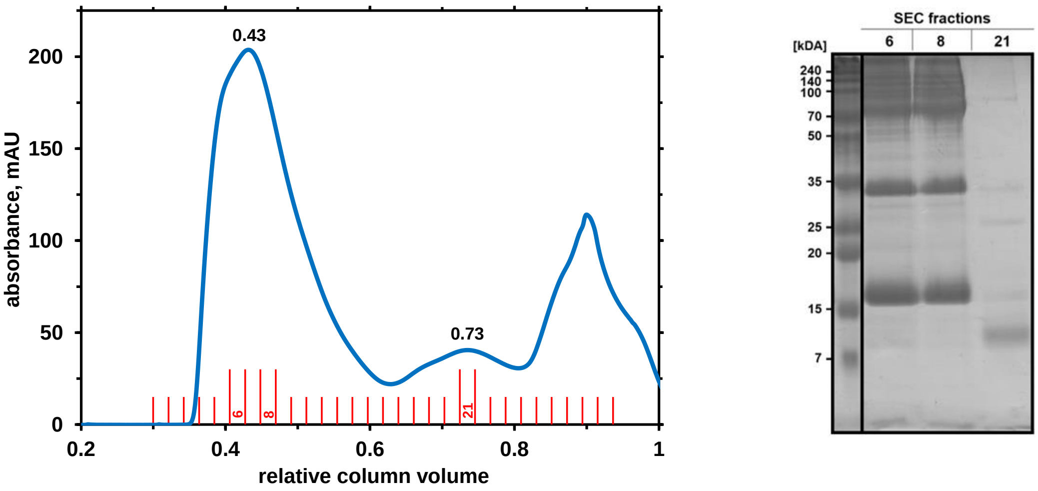


FIGURE 76: SEC chromatogram of ADAM17-TM-GB1

On the left, ADAM17-TM-GB1 was applied on a Superdex 200 Increase 10/300 GL SEC column, where the SEC running buffer contained 20 mM NaPi pH 6.0, 50 mM NaCl, 1 mM EDTA, 5 mM DTT, and 0.1 wt% DPC. The major peak eluted at 0.43 CV and a minor peak eluted at 0.73 CV, corresponding to molecular weights of approximately 536 kDa and 24.5 kDa, respectively. On the right, a Coomassie-stained SDS-PAGE gel of the fractions six, eight and twenty-one is shown, indicating the presence of ADAM17-TM-GB1 (14.9 kDa) and Get3-MBP (86.3 kDa) in addition to some impurities in the fractions six and eight, while fractions twenty-one appeared to contain cleaved His$_{10}$-GB1 (9.52 kDa).

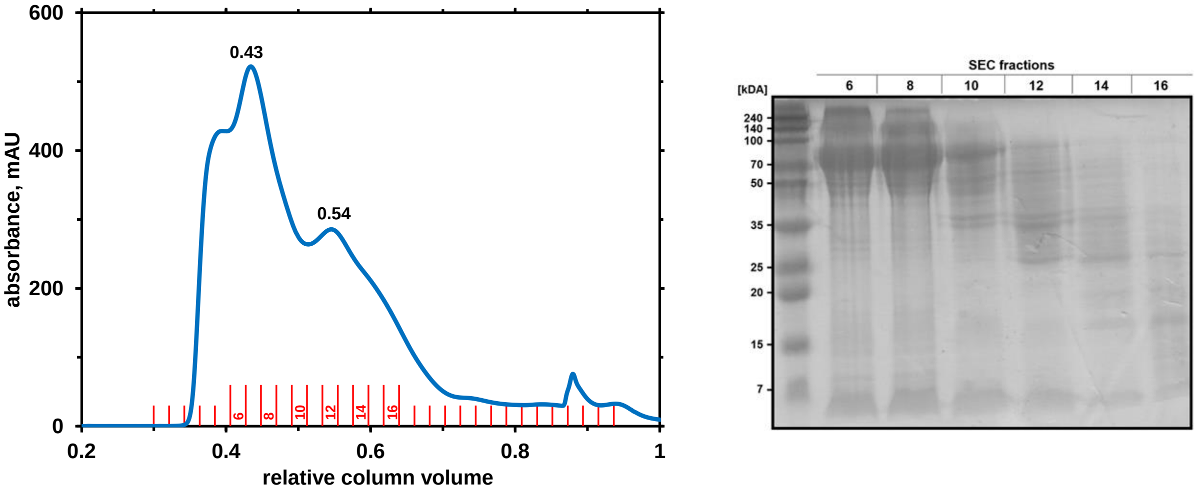


FIGURE 77: SEC chromatogram of thrombin-cleaved ADAM17-TM-GB1

On the left, thrombin-cleaved ADAM17-TM-GB1 was applied on a Superdex 200 Increase 10/300 GL SEC column, where the SEC running buffer contained 20 mM NaPi pH 6.0, 50 mM NaCl, 1 mM EDTA, 5 mM DTT, and 0.1 wt% DPC. The two major peaks eluted at 0.43 CV and 0.54 CV, corresponding to molecular weights of approximately 536 kDa and 173 kDa, respectively. On the right, a Coomassie-stained SDS-PAGE gel of the fractions six, eight, ten, twelve, fourteen and sixteen is shown, indicating ADAM17-TM (5.34 kDa) and Get3-MBP (86.3 kDa) among various impurities in all fractions.

Interestingly, all three constructs exhibited different running behavior on SDS-PAGE, as seen in FIGURE 71 and FIGURE 72. For example, ADAM17-TM-TEV-Z-tag appeared much smaller on the gel with around 16 kDa than its theoretical 21.0 kDa and ADAM17-TM-Z-tag also appeared smaller with around 18 kDa than its theoretical 20.8 kDa. ADAM17-TM-GB1, on the other hand, appeared to be slightly larger with around 16 kDa than its theoretical 14.9 kDa. Since all constructs had an N-terminal His$_{10}$-tag before the Z-tag or GB1 and bound to the Ni-NTA columns, the N-terminal part of the proteins was not affected by proteolytic cleavage. In addition, the presence of Get3-MBP suggested that the C-terminal ADAM17-TM was not truncated because Get3 can bind the transmembrane domain of ADAM17. In conclusion, the differences in running behavior during SDS-PAGE cannot be explained by truncated constructs.

To achieve complete separation of ADAM17-TM and Get3, numerous different purification conditions were tested (data not shown), all of which resulted in unsatisfactory separation of the two proteins. For example, ADAM17-TM-GB1 was purified via Ni-NTA affinity chromatography under denaturing conditions using either 6 M GdmCl or 8 M urea instead of 0.2 wt% DPC, and Get3-MBP still could not be washed off during Ni-NTA affinity chromatography. Because Get3 contains an ATPase subdomain [114], Ni-NTA affinity chromatography was performed in the presence of either 5 mM ATP or 5 mM ADP, which did not improve the separation of ADAM17-TM from Get3. The same result was obtained when 5 mM $MgCl_2$ or 5 mM $ZnCl_2$ were present during Ni-NTA affinity chromatography, where zinc is involved in the dimerization of Get3 [114]. In conclusion, ADAM17-TM could be produced at a high expression level by co-expression of Get3, but separation of the two proteins remained challenging.

3.3.2. Protein production and purification of iRhom2-TM1

To produce the first transmembrane domain of iRhom2, two different constructs were generated with either a Z-tag or GB1, and both constructs were co-expressed with Get3-MBP. The amino acid sequences can be found in the appendix in TABLE 32 and TABLE 34, respectively. Co-expression of Get3-MBP indeed allowed expression of iRhom2-TM1, as clearly shown in FIGURE 78 by two intense bands on the SDS-PAGE gel, indicating a high expression level of iRhom2-TM1-Z-tag or iRhom2-TM1-GB1 and Get3-MBP. To purify soluble expressed iRhom2-TM1, Ni-NTA affinity chromatography was performed in the presence of DPC, as shown exemplarily in FIGURE 78. This process resulted in an unsatisfactory outcome, as Get3-MBP could not be separated from iRhom2-TM1-Z-tag / iRhom2-TM1-GB1 via Ni-NTA affinity chromatography.

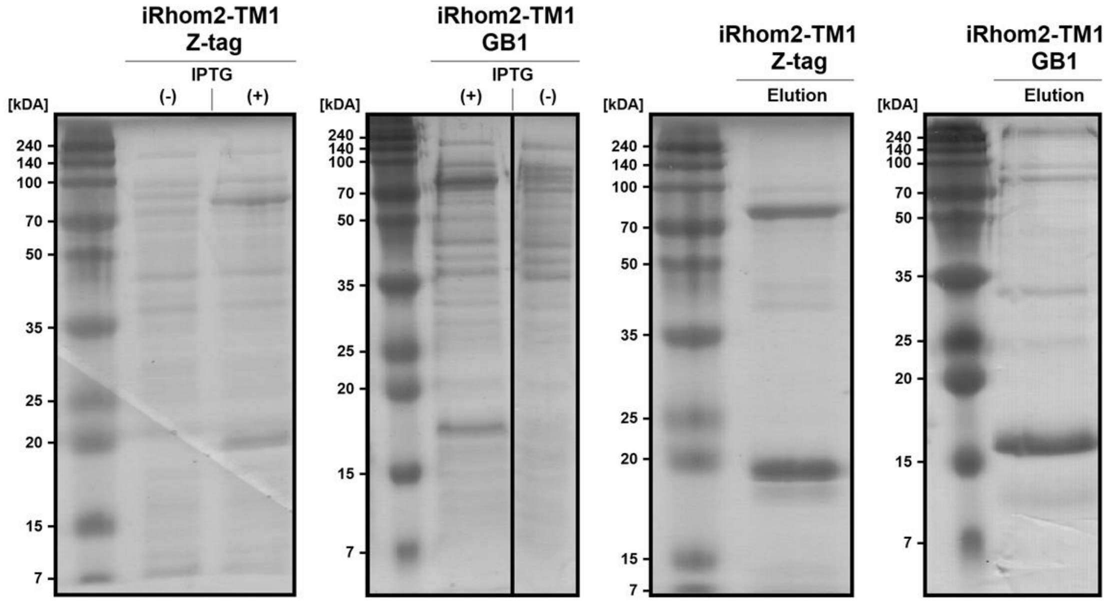


FIGURE 78: Expression and purification of the iRhom2-TM1 constructs

On the left, Coomassie-stained SDS-PAGE gels of iRhom2-TM1 constructs (Z-tag or GB1) expressed by *E. coli* BL21(DE3) in LB medium 4 h after induction with 1 mM IPTG at 37 °C are shown, while Get3-MBP was co-expressed. On the right, Coomassie-stained SDS-PAGE gels of the iRhom2-TM1 (Z-tag or GB1) constructs purified by Ni-NTA affinity chromatography are shown, where a buffer containing 50 mM Tris pH 8.0, 500 mM NaCl, 5 mM BME, 500 mM imidazole, and 0.2 wt% DPC was used to elute the proteins from the Ni-NTA columns.

Interestingly, iRhom2-TM1-Z-tag appeared smaller on the gel with around 19 kDa than its theoretical 22.8 kDa, although this difference cannot be explained by truncated constructs. Subsequent SEC of non-cleaved and thrombin-cleaved iRhom2-TM1 constructs showed the formation of high molecular weight species (data not shown), comparable to the results obtained with ADAM17-TM. SDS-PAGE showed that iRhom2-TM1-Z-tag / iRhom2-TM1-GB1 and Get3-MBP could not be separated with SEC and appeared together, whereas iRhom2-TM1 was most likely embedded within Get3. To achieve complete separation of iRhom2-TM1 and Get3, numerous different purification conditions were tested in analogy to ADAM17-TM (data not shown), all of which resulted in unsatisfactory separation of the two proteins. However, iRhom2-TM1 could be produced at a high expression level by co-expression of Get3.

3.3.3. Protein production of iRhom2-7TM

To produce all seven transmembrane domains of iRhom2 in one step, iRhom2-7TM was produced as a fusion protein with a thrombin-cleavable α_5 integrin fragment, with the protein expected to accumulate in inclusion bodies due to fusion with the integrin fragment. The amino acid sequence of this construct can be found in the appendix in TABLE 35 and its expression in *E. coli* BL21(DE3) at 37 °C is shown in FIGURE 79. The SDS-PAGE gel revealed that iRhom2-7TM-Integrin with a molecular weight of 60 kDa was not expressed, as the protein expression pattern was the same before and after induction with IPTG. Subsequent purification of the inclusion bodies confirmed this observation, as iRhom2-7TM-Integrin was not detectable in the elution fraction of the Ni-NTA affinity chromatography (data not shown).

Because iRhom2-7TM was not producible with integrin as a fusion protein, iRhom2-7TM was fused to a thrombin-cleavable Z-tag and the construct was co-expressed with Get3-MBP, whereas amino acid sequences can be found in the appendix in TABLE 32 and TABLE 35, respectively. FIGURE 79 shows the expression of iRhom2-7TM-Z-tag / Get3-MBP in *E. coli* BL21(DE3) at 37 °C, where the protein expression pattern was the same before and after induction with IPTG, indicating no expression of the construct. This was verified by subsequent purification via Ni-NTA affinity chromatography, which yielded an elution fraction that did not contain iRhom2-7TM-Z-tag (data not shown).

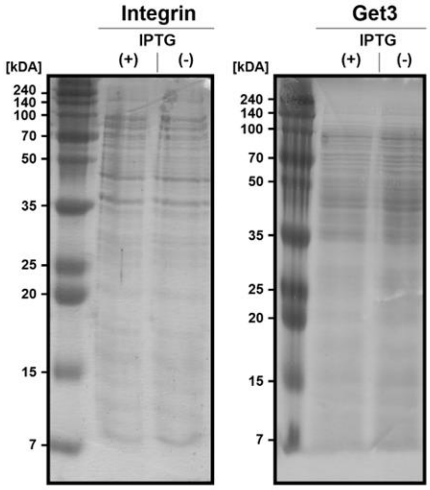


FIGURE 79: Expression of the iRhom2-7TM constructs

Coomassie-stained SDS-PAGE gels shows protein expression at 37 °C of iRhom2-7TM combined with a thrombin-cleavable integrin (left) or with a thrombin-cleavable Z-tag (right) by *E. coli* BL21(DE3) in LB medium before induction (-) and 4 h after induction with 1 mM IPTG (+). In the case of the construct with a Z-tag, Get3-MBP was also co-expressed.

4. DISCUSSION

4.1. OEP21

4.1.1. OEP21 as a β-barrel membrane pore with a positively charged interior

Expression of OEP21-WT as inclusion bodies in *E. coli* BL21(DE3) and purification of His_{10}-tagged OEP21-WT via Ni-NTA affinity chromatography under denaturing conditions yielded a large amount of already highly pure OEP21-WT (FIGURE 15), allowing further experiments. Refolding of OEP21-WT in LDAO micelles as a membrane mimetic proved to be the right choice, as final purification by SEC yielded a homogenous protein sample (FIGURE 16) and the far-UV spectrum of OEP21-WT in LDAO micelles showed the expected β-sheet secondary structure for OEP21 (FIGURE 18). This was also confirmed by secondary ^{13}C chemical shift information showing twelve β-strand regions of varying lengths for OEP21 (FIGURE 19), which is in contrast to a predicted sequence-based secondary structure with eight β-strand regions [43], but in good agreement with the prediction of the AlphaFold databank [191, 192]. Therefore, OEP21 appeared to adapt a native conformation in LDAO micelles, which allowed obtaining a high-resolution structure of OEP21-WT by NMR spectroscopy, showing that OEP21 assembled a well-defined β-barrel pore with an overall cone-like shape (FIGURE 20). The pore has a highly positively charged inner surface (FIGURE 20) to which negatively charged metabolites such as phosphate, triosephosphates, and ATP can bind, allowing translocation through the pore. Translocation of these negatively charged metabolites has been previously demonstrated by electrophysiology experiments [43, 78]. In addition, the cone-shaped structure of OEP21 may indicate directional metabolite flow. Determination of the orientation of OEP21 in the chloroplast outer envelope demonstrated that loop five is oriented toward the cytosol, with the wide opening of OEP21 oriented in the outer envelope toward the inter membrane space (FIGURE 21), which may facilitate channeling of metabolites into the cytosol. This confirms the outward-rectifying properties proposed for OEP21-mediated metabolite flux [43]. The translocation process of triosephosphates and its directionality are most likely driven by a metabolite concentration gradient across the outer envelope membrane, since OEP21 does not consume ATP or any other form of direct chemical energy [43, 71, 78].

4.1.2. Oxidation-dependent oligomerization of OEP21

As indicated by the early SEC retention volume of the OEP21/LDAO complex, OEP21 forms higher oligomers in LDAO micelles (FIGURE 16). Chemical crosslinking experiments with OEP21 in liposomes or in LDAO confirmed the appearance of higher oligomers (FIGURE 23). However, the LDAO-to-protein seems to be critical because the tendency of OEP21 to form oligomers depends on the LDAO concentration, with a higher LDAO concentration allowing the appearance of smaller species of OEP21/LDAO besides the formation of higher oligomeric species (FIGURE 17, FIGURE 23). Interestingly, at a low LDAO concentration, i.e., in the oligomeric state of OEP21, the β-barrel structure of the monomers appeared to be slightly compressed, as shown by NMR experiments (FIGURE 24). This is also indicated by the reduced binding affinity from ITC experiments, which was likely due to impaired accessibility of ATP to oligomeric OEP21 at a low LDAO concentration (FIGURE 30, FIGURE 31). However, the presence of nucleotides such as ATP, which is present in plant cells at a concentration of about 1–2 mM [182], can stabilize the pore in an active form and prevent conformational changes associated with oligomer formation (FIGURE 24).

Moreover, OEP21 can appear as dimeric form meditated by a cysteine residue that is oriented toward the membrane and well-positioned to form a disulfide bridge under oxidizing conditions, as shown by NMR and oxidation experiments (FIGURE 25, FIGURE 26). The strong tendency of OEP21 to form dimers and oligomers may be even more pronounced *in vivo*, as the protein content in the outer envelope membrane is estimated to be one-quarter of the total weight [193], which is about fifteen times higher than the protein concentration used in the crosslinking experiments. Further experiments showed that the tendency of OEP21 to oligomerize under oxidizing conditions is even more prominent when disulfide-bridged dimers are present (data not shown), suggesting an important influence on oligomerization. In addition, the presence of GAP or especially ATP at a low detergent concentration resulted in fewer oligomers (data not shown), supporting the effect of ATP in the NMR measurements (FIGURE 24). This oxidation-dependent oligomerization of OEP21 in response to oxidative stress may be a critical switch for plant cells in executing the hypersensitive response, a type of programmed cell death to defend against pathogen infections in plants [194]. In addition, ATP levels have been shown to play a role in disease resistance and to decrease dramatically during the hypersensitive response [195].

4.1.3. OEP21 binds to metabolites via promiscuous electrostatic interactions

To characterize in detail the mechanism of metabolite binding by OEP21, various experiments were performed. CD-detected thermal melting experiments (FIGURE 27) demonstrated that negatively charged metabolites can bind to OEP21, as evidenced by increased stability of OEP21, with the stabilizing effect depending on the number of charges in the metabolite. ITC experiments were used to determine the binding affinity of various negatively charged metabolites to OEP21 (FIGURE 32), and the binding affinity increased with the number of negative charges in the metabolite, which agrees well with the results of the CD-detected thermal melting experiments and the trend of electrophysiological experiments [43, 78]. But the absolute values of binding affinity obtained from ITC experiments differ significantly from the previously reported values derived from electrophysiological experiments. For example, ITC experiments revealed a K_D of 11.9 µM for the nucleoside triphosphate ATP (FIGURE 28), whereas electrophysiological experiments suggested a K_D in the range of 100 – 200 µM for ATP, depending on the experimental setup [43, 78]. However, ITC experiments allowed to obtain more reliable binding affinities for various metabolites than electrophysiological experiments, which was confirmed by NMR titration experiments that revealed a K_D of 36.0 µM for the high-affinity ATP-binding site inside the β-barrel (FIGURE 37). Since the binding affinity for ATP to OEP21 is in the low micromolar range and ATP is present in plant cells at a concentration of about 1 – 2 mM [182], OEP21 appears to be present in the outer envelope mainly in an ATP-bound state. ATP-induced chemical shift perturbations in OEP21 during NMR titration experiments allowed mapping of the ATP-binding site on OEP21 (FIGURE 35, FIGURE 36). Interestingly, OEP21 exhibits a second ATP-binding site with low affinity and a K_D of 775 µM, located at the cytoplasmic entrance of the pore that involves loop five. This is consistent with electrophysiology experiments where a second ATP-binding site was also identified, although the reported binding affinity in the electrophysiology experiments was much lower with a K_D of 15 – 18 mM for ATP [43, 78]. This low-affinity peripheral binding site for ATP could not be observed with ITC due to sensitivity limitations, so ITC experiments resulted in approximately equimolar stoichiometry (FIGURE 28). Due to the limitations of ITC, no reliable K_D value could be obtained for the monophosphorylated GAP (FIGURE 33), suggesting binding with a low affinity. Using NMR titration experiments, a K_D of 152 µM for GAP was derived by line broadening of the NMR signals (FIGURE 41), and the binding sites of GAP could be mapped on OEP21 (FIGURE 38, FIGURE 39). The low binding affinity for GAP is consistent with CD and ITC experiments, which demonstrated that the binding affinity increases with the number of negative charges in the metabolite, and also with electrophysiology experiments in which K_D values in the higher micromolar range were obtained for GAP, namely 170 / 900 µM [78] or 1.5 mM [43], depending on the experimental setup.

Interestingly, the effect of ATP and GAP on OEP21 is quite different, as shown by the respective NMR spectra (FIGURE 40). ATP titrations showed specific NMR chemical shift perturbations that allowed deriving two well-defined ATP-binding sites (FIGURE 36), whereas GAP titrations induced a concentration-dependent line broadening of NMR signals in the complete β-barrel structure (FIGURE 39). Further NMR experiments in which sodium phosphate was titrated to OEP21 (FIGURE 38) or GAP was titrated to the unrelated bacterial β-barrel protein OmpX in LDAO resulted in no line broadening of NMR signals and demonstrated that the line broadening effects of GAP on the OEP21 spectrum were specific. However, in contrast to ATP, GAP partially interacted non-specific with LDAO detergent micelles (FIGURE 57).

In addition to ATP, GAP was shown to bind to loop five, as the NMR signal intensities of amino acids residues located in the loop five region increased significantly upon addition of GAP (FIGURE 38, FIGURE 39). For binding of GAP to loop five of OEP21, a K_D of 3.2 mM was derived from the NMR titration experiments (FIGURE 41), consistent with a lower binding affinity compared with ATP for loop five and compared with the higher binding affinity of the β-barrel for GAP. Moreover, OEP21-ΔLoop5 showed lower binding affinity with a K_D of 730 µM for GAP compared with OEP21-WT (FIGURE 41), demonstrating the importance of loop five for binding negatively charged metabolites. The interaction of loop five with metabolites was confirmed by NMR experiments using amide hydrogen exchange rates (FIGURE 42) and indicated by the slightly reduced proteolytic cleavage of loop five in the presence of ATP or GAP (FIGURE 22). However, a second binding site for GAP involving loop five has not been discussed in previous electrophysiology experiments [43, 78]. Binding of negatively charged metabolites to OEP21 has been demonstrated, raising the question of how the interplay of different metabolites might affect binding events to OEP21. Since ATP was suspected to act as an inhibitor of triosephosphate transport across OEP21 [43], the competition of GAP with ATP was investigated using ITC experiments. Indeed, GAP was found to bind to OEP21 and compete with ATP in a concentration-dependent manner (FIGURE 34). Moreover, ITC experiments in the presence of sodium phosphate (FIGURE 32) confirmed that negatively charged metabolites acted as competitors for ATP. This was further confirmed by FP-based binding experiments with MANT-ATP in the presence of increasing sodium chloride concentrations (FIGURE 53), which demonstrated that the binding mechanism is rather promiscuous and mainly based on bulk electrostatics properties. In addition, the FP-derived binding affinity of OEP21 variants in which single positively charged lysine or arginine residues were mutated to alanine (FIGURE 43) was comparable to that of the wildtype (FIGURE 51, FIGURE 52). This demonstrated that the removal of individual positive charges was compensated by other positively charged residues in the proximity, confirming that metabolite binding to OEP21 is promiscuous and dominated by electrostatic interactions.

Thereby, all OEP21 variants were comparable to the wildtype as confirmed by the comparable expression level, refolding behavior, SEC elution profile, final purity (FIGURE 44), and CD-detected secondary structure content and thermal stability (FIGURE 45). However, as shown by the ITC-derived binding affinities for OEP21 variants (FIGURE 48), FP experiments prove to be a more reliable method for determining binding affinities. Since FP is related to the tumbling property of a fluorophore, which is reduced by binding to a larger molecule such as OEP21, the readout is more direct than ITC, which measures relative heat production that can be affected by other processes besides binding. Determination of binding affinities by FP proved to be a robust assay, as shown by the titration of LDAO to MANT-ATP, with LDAO alone causing no change in FP (FIGURE 49), and the fact that the additional fluorescently labeled MANT group did not hinder binding to OEP21 (FIGURE 49), as well as MANT-ATP and ATP bound to OEP21 in a comparable manner, as shown by competition experiments with ATP (FIGURE 50). In contrast to the ITC-derived binding affinity for ATP (K_D: 21.9 µM, FIGURE 48), the binding affinity determined by FP was significantly higher (K_D: 182 nM, FIGURE 49), which may reflect increased affinity in the presence of the hydrophobic MANT group besides the use of different methods. However, the application of FP-based binding experiments is limited by the availability of fluorescently labeled analogs. Nevertheless, ITC is a useful technique, although the binding affinities determined here, e.g., for various metabolites (FIGURE 32), may be inaccurate, but this should not affect the overall trend.

ITC was also a useful method for determining the binding affinities of magnesium chloride to ATP, GAP, or sodium phosphate. The results show that $MgCl_2$ bound to ATP with a K_D of 54 µM (FIGURE 54), which is in excellent agreement with the reported K_D of 50 µM ± 10 µM [184], whereas the binding of $MgCl_2$ to GAP or NaPi was much weaker (K_D: 9 mM, FIGURE 54). Interestingly, magnesium appears to play a crucial role in binding to OEP21, as shown by ITC, FP, and NMR experiments. In addition, it is very likely that increased concentrations of magnesium are present in the intermembrane space under light conditions, since magnesium is released from the thylakoid lumen of chloroplasts into the stroma upon illumination [185] and magnesium transporters have been reported in the inner envelope of chloroplasts [186]. Due to the relatively strong binding of $MgCl_2$ to ATP, the negative charge density of ATP is reduced, suggesting a reduced binding affinity of ATP to OEP21. Indeed, FP-based binding experiments in the presence of increasing $MgCl_2$ concentrations demonstrated a significant concentration-dependent decrease in the binding affinity of MANT-ATP to OEP21 (FIGURE 53). Interestingly, the presence of $MgCl_2$ during NMR titrations with GAP also significantly decreased the binding affinity of GAP to OEP21 (FIGURE 41). All these findings provided the basis for the FP-based competition experiments, which demonstrated the ability of OEP21 to bind various metabolites depending on their binding affinity / negative charge density and their relative concentration, considering the presence of other possible competitors.

Here, it was shown that GAP (IC_{50}: 12 mM) and NaPi (IC_{50}: 7.1 mM) can act as competitors for ATP, whereas the presence of $MgCl_2$ indicates much more efficient competition by GAP (IC_{50}: 1.2 mM, FIGURE 55, FIGURE 56) due to the reduced charge density of ATP. Consistent with these results, NMR-based experiments showed competition between ATP and GAP for OEP21 (FIGURE 58). However, the relatively high GAP concentrations required for ATP dissociation, compared to the concentrations required with sodium phosphate, were in contradiction with the derived binding affinities of GAP (K_D: 152 µM) and NaPi (K_D: 7 mM). This could be explained by partial non-specific binding to the detergent micelle (FIGURE 57), but the influences are not yet entirely clear.

In summary, all these novel insights into the binding mechanism and interaction between different metabolites provide the basis for a better understanding of the functional role of OEP21 as a transport channel in the outer envelope.

4.1.4. Size-selective metabolite transport through OEP21

Based on the results presented here on the binding mechanism and interplay of different metabolites, Dr. Umut Günsel experimentally investigated channel functionality by metabolite transport assays using OEP21 proteoliposomes. For this purpose, liposomes filled with metabolites were subjected to SEC, and the loss in metabolite content was subsequently quantified, as shown in FIGURE 80 (A). In the assay deadtime (approximately four minutes), both ATP and GAP were able to pass through the OEP21 membrane channel, as shown in FIGURE 80 (B) and (C). Interestingly, the content of metabolites inside liposomes decreased with an increasing concentration of liposome-inserted OEP21 (data not shown), indicating an OEP21-dependent flux outside the liposomes. In addition, the effect of 5 mM ATP inside and outside the liposomes on GAP translocation was investigated, and GAP transport was not hindered but was actually enhanced by the presence of ATP (FIGURE 80 B), suggesting that ATP plays an activating rather than an inhibitory role and probably stabilizes the channel in an open state. Consistent with reduced binding affinity in the presence of $MgCl_2$ and NaCl (FIGURE 53), GAP translocation was reduced upon addition of $MgCl_2$ (FIGURE 80 B) and with increasing NaCl concentrations (data not shown). A similar result was obtained for the translocation of ATP, where the presence of $MgCl_2$ decreased the translocation efficiency (FIGURE 80 C). In addition, OEP21-ΔLoop5 resulted in higher efficiency in both ATP and GAP translocation compared with OEP21-WT. Interestingly, ATP and GAP appeared to permeate simultaneously through the OEP21 channel, although they showed a competing binding mode during FP-based experiments.

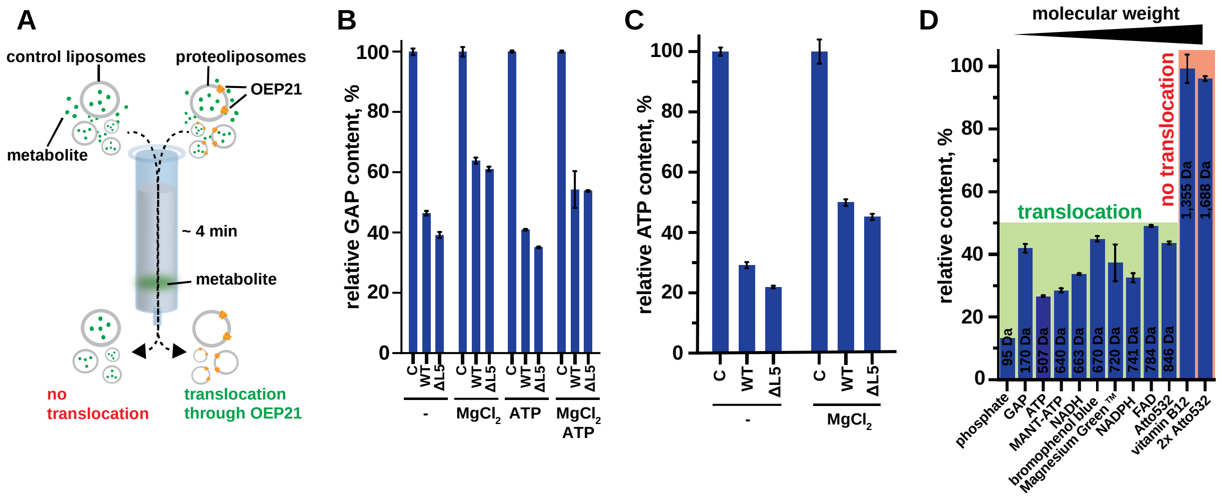


FIGURE 80: Size-selective metabolite transport through OEP21

The setup of the translocation assay based on SEC with control liposomes and proteoliposomes assembled with OEP21 is shown in (A). Therefore, the liposomes filled with metabolites were subjected to SEC, and the loss in metabolite content was subsequently quantified. The assay deadtime was approximately four minutes using PD-10 desalting columns packed with Sephadex G-25 resin (GE Healthcare / Chicago, USA). The relative GAP content quantified after SEC for control liposomes (C) and proteoliposomes assembled with OEP21-WT (WT) or OEP21-ΔLoop5 (ΔL5) is shown in (B). In addition, GAP content was quantified using liposomes prepared in the presence of 10 mM $MgCl_2$, or / and 5 mM ATP, with SEC conducted under the same conditions. Analogous to (B), (C) shows relative ATP content and (D) shows the relative content for various metabolites of different sizes. Liposomes were prepared according to FIGURE 21 in the presence of 2 mM ATP, where the molecule to be tested was added to the liposomes just before the three freeze-thaw cycles and fifteen passes through a 0.2-µm-filter using a mini-extruder. The concentrations of the metabolites used were: 2 mM GAP, 2 mM ATP, 200 µM MANT-ATP, 2 mM NADH, 140 µM bromophenol blue, 100 µM Magnesium Green^{TM} (Thermo Fisher Scientific / Waltham, USA), 2 mM NADPH, 1.5 mM flavin adenine dinucleotide (FAD), 30 µM Atto532 (ATTO-TEC GmbH / Siegen, GER), or 2.5 mM vitamin B12. The monomeric Atto532 (846 Da) was generated by incubating Atto532-maleimide (ATTO-TEC GmbH / Siegen, GER) with two times the excess of BME, while the dimeric (2x) Atto532 (1,688 Da) was generated by mixing DTT and Atto532-maleimide in a molar ratio of 1:2 for at least 2 h at room temperature. Liposome samples were then applied to PD-10 desalting columns equilibrated with 20 mM HEPES pH 7.5 and 250 mM NaCl. The content of each molecule inside the collected liposomes was quantified in the presence of 0.2 vol% Triton X-100 to disintegrate the liposomes. GAP content was determined via NADH fluorescence (excitation: 340 nm, emission: 450 nm) generated by an enzyme-coupled NADH assay in the presence of 1 mM NAD^+, 10 mM potassium phosphate pH 7.5, and 0.02 $g \cdot L^{-1}$ GAP dehydrogenase. The following metabolites were determined using fluorescence detection with the respective excitation and emission wavelengths: NADH and NADPH (340 nm / 450 nm), Magnesium Green^{TM} (506 nm /546 nm), FAD (450 nm / 520 nm), and Atto532 (532 nm / 572). ATP and MANT-ATP content were determined using Invitrogen^{TM} ATP Determination Kit (Thermo Fisher Scientific / Waltham, USA). Bromophenol blue and vitamin B12 contents were determined by absorbance measurement at 590 nm and 363 nm, respectively. These data were provided by Dr. Umut Günsel.

This raised the question of the selectivity of OEP21, and the size limit of small molecules for channel permeation was investigated using translocation assays with molecules of increasing molecular weight (FIGURE 80 D). The results revealed that molecules up to a molecular weight of about 1 kDa can pass through the channel to a similar extent, while larger molecules such as vitamin B12 or the dimeric fluorescent dye '2x Atto532' cannot pass through the pore of OEP21. This clearly shows that the specificity of metabolite translocation across OEP21 is quite broad, but still has a sharp size limit. Further experiments investigating the leakiness of OEP21 proteoliposomes for positively charged Na^+ ions (data not shown) revealed that translocation of even small cations is clearly slower than that of much larger but negatively charged molecules, where complete release is achieved after only a few minutes.

Considering all the data presented here, a functional model of OEP21 could be constructed, as shown in FIGURE 81.

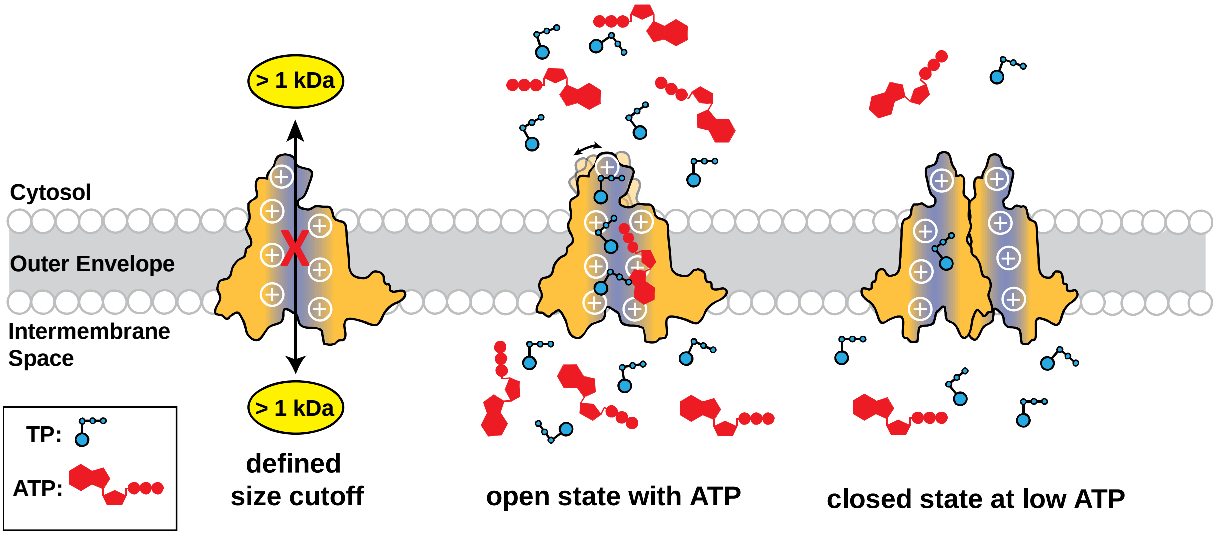


FIGURE 81: Functional model of OEP21

A schematic representation of OEP21 functionality is shown in (E), where OEP21 has a substrate cutoff of approximately 1 kDa. At physiological ATP concentrations, OEP21 is in its open state, whereas under oxidative stress and low ATP conditions, enhanced OEP21 oligomerization leads to pore closure. This figure was provided by Prof. Dr. Franz Hagn.

In the apo state, OEP21 appears to be quite dynamic, allowing for conformational changes, such as squeezing the β-barrel in the oligomeric state. This shape change could lead to closure of OEP21, as indicated by the binding and transport experiments here and also suggested for the mitochondrial β-Barrel porin voltage-dependent anion channel (VDAC) [196]. Interestingly, ATP probably leads to thermodynamic stabilization of the pore shape and keeps the channel in an active and permeable state, with ATP present in higher concentrations in cells.

Furthermore, it was shown that the relatively strong binding affinity of ATP does not prevent the binding and translocation of other phosphorylated metabolites, such as GAP, the primary product of photosynthesis. Because the transport assay presented has limited temporal resolution, it was not possible to investigate whether ATP can slow the kinetics of GAP translocation, but endpoint analysis after approximately four minutes and competition experiments indicated that simultaneous translocation is possible. However, ATP may have a regulatory role in OEP21 functionality to enable triosephosphate transport under conditions where triose phosphate utilization, i.e., sucrose synthesis, is not limiting and phosphate concentration is high enough to enable efficient photosynthesis and ATP production. When triose phosphate utilization is limited, ATP levels decreases significantly, and further export of triosephosphates must be restricted to prevent phosphate depletion in chloroplasts, which can be very dangerous for the plant cell [197].

As shown, OEP21 facilitates size-dependent translocation of preferentially negatively charged molecules, suggesting that the flux of larger molecules is most likely directed by the larger members of the OEP family, such as OEP24, OEP37, and OEP40. This is also supported by proteomics data from plants with different levels of photosynthetic activity, e.g., C3 versus C4 plants, showing that the abundance of individual members of the OEP family is shifted toward the larger pores when photosynthesis is more efficient [76]. Therefore, not only the size limit but also the total metabolite transport capacity might be increased for larger OEPs. In contrast to the outer mitochondrial membrane, where only three VDAC isoforms mediate metabolite exchange [198, 199], the situation in chloroplasts appears to be more variable and fine-tuned, with greater variation in metabolite flux in chloroplasts between different plants, requiring a variable set of OEPs. Possible functional redundancy between specific members of the OEP family could be the main reason for the lack of a phenotype under normal growth conditions, where, for example, OEP37 knock-outs in *Arabidopsis thaliana* did not result in a phenotype [75]. In addition, a knock-down of OEP21 in *Arabidopsis thaliana* also showed no phenotype (data not shown). However, the extent to which the individual OEP members contribute to the overall properties of the outer envelope membrane is poorly understood to date and requires further investigations in the future.

4.1.5. The role of loop five of OEP21

Deletion of loop five does not appear to affect the overall structure of OEP21, as indicated by homology modeling of soybean OEP21 lacking loop five (data not shown), CD-detected secondary structure content (FIGURE 45), NMR experiments (FIGURE 58), and FP-based binding affinity (FIGURE 52). However, OEP21-ΔLoop5 showed a more efficient GAP translocation as observed in the liposome-based transport assays (FIGURE 80 B, C).

Consistent with this finding, higher GAP concentrations are required for ATP displacement in the case of OEP21-ΔLoop5 compared with OEP21-WT, as shown by FP-based and NMR-based competition experiments (FIGURE 56, FIGURE 58). This can be explained by the involvement of cytosolic loop five in the formation of an external binding site for negatively charged metabolites before they are released into the cytosol. Moreover, loop five of OEP21 is absent in many plants, whereas most other parts of the protein have high sequence conservation (data not shown), implying that loop five is not essential for OEP21 functionality. This suggests that loop five may also be involved in fine-tuning metabolite flux in certain plants such as pea and clover, since their genome is quite large and evolves faster than that of other Leguminosae [200]. Interestingly, sequence alignment analysis shows that loop five likely arose from duplication of a peptide stretch in the adjacent β-barrel (61–71: GVGVRYDKQD, 72–81: SVGVRYAKND).

4.2. Bna-Tic40

Male-sterile lines play an important role in plant breeding to maintain the vigor of hybrids [1]. The question now was whether a functional relationship between the fertile BnaA10-Tic40 and the non-fertile BnaC9-Tic40 could be derived from structural data.

The proteins BnaA10-Tic40 and BnaC9-Tic40 exhibited high purity (FIGURE 59), which allowed further experiments to be performed to characterize both proteins. As a first step, the secondary structure of BnaA10 and BnaC9 was investigated using far-UV spectra (FIGURE 60), and the spectra for both proteins show a high proportion of α-helical secondary structure, which agreed well with the predictions. The absence of a clear transition during CD-detected thermal melting experiments (FIGURE 60) for BnaA10 and BnaC9 indicated a less compact folding and rather dynamic structure elements of both proteins. This was also indicated by the low spectral quality of initial 2D-[^{15}N,^{1}H]-TROSY NMR spectra of both proteins (FIGURE 62). Although BnaA10 and BnaC9 are quite comparable, with only about 10 % difference in their amino acid sequences (TABLE 28), the impact on structure appeared to be rather large, as BnaA10 exhibited more α-helical secondary structure (FIGURE 60) and better NMR spectral quality (FIGURE 62, FIGURE 63). In conclusion, the structure of BnaA10 appeared to be more defined and less dynamic than that of BnaC9. However, the NMR spectral quality of BnaA10 was also not satisfactory for more detailed NMR analysis. Therefore, the addition of TFE seemed to be the right choice because TFE apparently stabilized conformations of peptide fragments [188, 189], as shown by the enhanced formation of α-helical secondary structure elements (FIGURE 61) and the significantly improved NMR spectral quality of BnaA10 (FIGURE 65).

The presence of 10 vol% TFE already seems to be sufficient to stabilize peptide conformations of BnaA10, as the addition of 20 vol% TFE did not increase the α-helical secondary structure content or improve the NMR spectral quality of BnaA10 compared with 10 vol% TFE. This is in line with the reported NMR structure of the C-terminal NP domain of Tic40 determined in the presence of 10 vol% TFE [50]. Comparison of the 2D-[^{15}N,^{1}H]-TROSY NMR spectra of BnaA10 and BnaC9 in the presence of 10 vol% TFE (FIGURE 64) shows that the spectral quality of both constructs improved significantly upon the addition of TFE, with BnaA10 still exhibiting better spectral quality, indicating a more defined structure than that of BnaC9.

The improved spectral quality of BnaA10 in the presence of TFE enabled the recording of 3D triple resonance experiments and the determination of sequence-specific resonance assignments (FIGURE 66, FIGURE 68 – FIGURE 70, TABLE 36). However, the 3D-NOESY spectra revealed that long-range NOE contacts were absent or not resolved, preventing reasonable structure determination with the recorded NMR spectra. The structure determination of BnaA10 could be allowed by recording further NOESY experiments under optimized sample conditions or the use of a more specific labeled protein. Examination of only the structural unknown Tic40 fragment (residues 305 – 395) would reduce the number of NMR signals in the 3D-NOESY spectra and would probably allow the definition of important NOE contacts. But the absence of important NOE contacts could indicate that the overall structure is still quite plastic, even if the individual α-helical elements were stabilized by the addition of TFE. However, further experiments could still not allow structure determination by NMR spectroscopy if the protein itself is too dynamic and therefore lacks long-range NOE contacts.

Nevertheless, the use of the AlphaFold databank [191, 192] allowed to derive a structural model of the structural unknown Tic40 fragment (residues 305 – 395), as shown in FIGURE 82. The deposited structural model of Tic40 from *Arabidopsis thaliana*, which shows high sequence identity with BnaA10-Tic40 shown in TABLE 29, agrees very well with the α-helical secondary structural elements identified by ^{13}C chemical shift information (FIGURE 67). This structural model in FIGURE 82 was combined with the published NMR structure of the NP-repeat domain of Tic40 from *Arabidopsis thaliana* [50] (residues 396 – 457, FIGURE 7). The NMR structure overlaid very well with the structural model predicted by AlphaFold as well as with the α-helical secondary structural elements determined here experimentally using ^{13}C chemical shift information. An additional α-helical domain (helix six) was detected between the N-terminal and C-terminal domain using ^{13}C chemical shift information, which is also visible in the structural model predicted by AlphaFold. The elongated shape of the structural model and the absent of larger hydrophobic clusters that would be expected for a compactly folded protein are consistent with the relatively low NMR spectral quality.

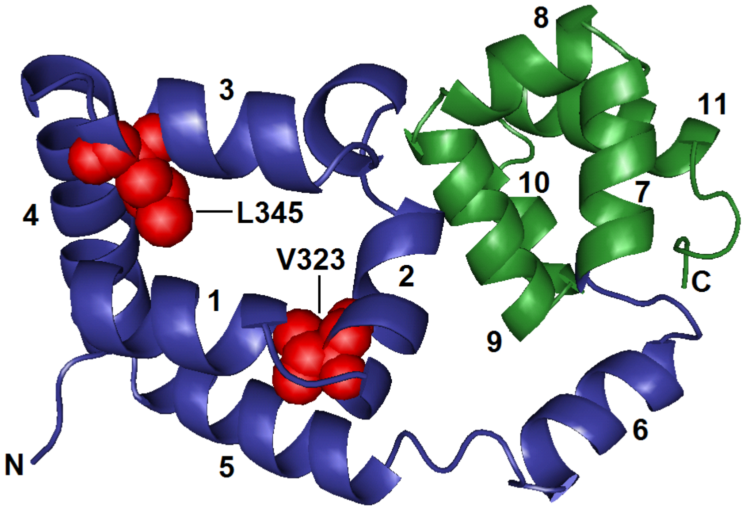


FIGURE 82: Structural model of BnaA10-Tic40

The structural model of the unknown N-terminal Tic40 fragment (residues 305 – 395, colored in blue) was obtained from the structure of *Arabidopsis thaliana* Tic40 (UniProt accession code: Q9FMD5) predicted by the AlphaFold databank (https://alphafold.ebi.ac.uk) [191, 192]. In addition, the structural model was combined with the published NMR structure of the C-terminal NP-repeat domain of Tic40 from *Arabidopsis thaliana* [50] (residues 396 – 457, colored in green), as shown in FIGURE 7. Amino acids residues are numbered according to BnaA10, with two deviating positions of BnaC9 (V323F and L345R) highlighted in red. The eleven α-helical secondary structure elements identified using ^{13}C secondary chemical shift information, as shown in FIGURE 67, agree well with the predicted N-terminal structural model and the C-terminal NMR structure. This procedure was conducted by Prof. Dr. Franz Hagn and the figure was provided by him.

TABLE 29: Sequence alignment between BnaA10 and the sequence of *Arabidopsis thaliana*

The amino acid sequences of BnaA10-Tic40 and the corresponding sequence of *Arabidopsis thaliana* are aligned, with the numbering of amino acids shifted by ten positions between the two. Deviations are colored red.

	305	315	325	335	345	355
BnaA10	GGAGLSVEAL	EKMMEDPTVQ	KMVYPHLPEE	MRNPETFKWM	LKNPQYRQQL	QDMLNNMSGS
A. thaliana	GGPGLSVEAL	EKMMEDPTVQ	KMVYPYLPEE	MRNPETFKWM	LKNPQYRQQL	QDMLNNMSGS
	295	305	315	325	335	345
	365	375	385	395	405	415
BnaA10	GEWDKRMTET	LKNFDLNSPE	VKQQFDQIGL	TPEEVISKIM	ENPDVAMAFQ	NPRVQAALME
A. thaliana	GEWDKRMTDT	LKNFDLNSPE	VKQQFNQIGL	TPEEVISKIM	ENPDVAMAFQ	NPRVQAALME
	355	365	375	385	395	405
	425	435	445	455		
BnaA10	CSENPMNIMK	YQNDKEVMDV	FNKISQLFPG	MTG		
A. thaliana	CSENPMNIMK	YQNDKEVMDV	FNKISQLFPG	MTG		
	415	425	435	445		

Interestingly, the two amino acids F321 and R343 of BnaC9 (V323 and L345 in BnaA10, respectively), which were essential for interaction with BnChimera, a proposed male sterility gene [1], are located in the N-terminal domain. Position 323 in BnaA10 is oriented toward the interior, while position 345 appears to be partially exposed to the surface, as shown in FIGURE 82. This suggests that these two mutated positions in BnaC9 most likely affected the folding state of the N-terminal domain, which is supported by the observed lower NMR spectral quality of BnaC9. Therefore, BnaC9 appears to exhibit a higher degree of protein flexibility and weakening of the folding state. In conclusion, binding between BnaC9 and the BnChimera to restore fertility depends either on interactions mediated by the specific amino acids F321 and R343 or, more globally, on the less compact folding state of BnaC9, which could facilitate the exposure of hydrophobic surfaces.

4.3. ADAM17 and iRhom2

Tail-anchored proteins mediate many important biochemical activities in almost every cell membrane, so their structural and functional feature are of great interest. Because tail-anchored proteins consist of an N-terminal cytosolic domain anchored to the intracellular membrane by a single C-terminal transmembrane domain, recombinant protein production and subsequent purification are challenging. Therefore, co-expression of the ATPase Get3 should allow soluble expression of tail-anchored proteins, with the helical transmembrane domain binding deeply within the composite hydrophobic groove of dimeric Get3 (FIGURE 10).

The soluble protein expression of the transmembrane domain of ADAM17 and the first transmembrane domain iRhom2 was enabled by co-expression of Get3-MBP (FIGURE 71, FIGURE 78). Therefore, the hydrophobic groove of Get3 was capable to embed both transmembrane domains. This allowed high protein expression levels of all used constructs (ADAM17-TM-TEV-Z-tag, ADAM17-TM-Z-tag, ADAM17-TM-GB1, iRhom2-TM1-Z-tag, and iRhom2-TM1-GB1). Previous experiments performed by Dr. Elisabeth Häusler showed that it was not possible to produce sufficient amounts of the transmembrane domains of ADAM17 and iRhom2 without co-expression of Get3. Therefore, co-expression with Get3 appeared to be a useful method to allow expression of proteins with a single α-helical transmembrane domain of about twenty residues that otherwise cannot be produced in sufficient quantity using standard expression systems. However, separation of the transmembrane domains from Get3 appeared to be challenging and did not lead to success here (FIGURE 72, FIGURE 78), using the rather harsh detergent DPC to solubilize the transmembrane domain during purification via Ni-NTA affinity chromatography.

Subsequent SEC in the presence of DPC also did not allow separation of the transmembrane domains from Get3 and revealed the formation of larger assemblies of both proteins for non-cleaved and thrombin-cleaved ADAM17-TM-Z-tag (FIGURE 73, FIGURE 74). In addition, the use of constructs with GB1 instead of the Z-tag did not improve separation and resulted in comparable results (FIGURE 75, FIGURE 76, FIGURE 77).

Interestingly, the complex of the transmembrane domain with Get3 proved to be very stable, as even denaturing conditions such as GdmCl or urea did not completely unfold Get3 during Ni-NTA affinity chromatography, allowing the transmembrane domain to be released from Get3 and both proteins to be separated. The transition from the open to the closed state of the Get3 dimer occurs around the hinge point centered zinc ion, and these conformations are regulated by interactions with nucleotides such as ATP / ADP (FIGURE 11). However, the separation of the transmembrane domains from Get3 in the presence of ATP, ADP, $MgCl_2$, or $ZnCl_2$ during Ni-NTA affinity chromatography was not significantly improved. This could be because Get3 cannot release its target protein and incorporate it into a membrane in the presence of ATP / ADP and magnesium without its partner proteins [201, 202]. In addition, loss of the transmembrane domain was observed when Get3 was partially removed, suggesting that DPC was not a perfect membrane mimetic for the transmembrane domain of ADAM17 and the first transmembrane domain of iRhom2.

To produce all seven transmembrane domains of iRhom2, two different constructs were used, where iRhom2-7TM was either fused to a thrombin-cleavable α_5 integrin fragment or fused to a thrombin-cleavable Z-tag and was co-expressed with Get3-MBP. For both constructs, no protein production was observed (FIGURE 79). This could be due to several reasons, in particular, the linker sequences between the transmembrane domains could be too short in the case of co-expression with Get3, since only transmembrane domain can be embedded per dimeric Get3.

In summary, single helical transmembrane domains can be produced at a high expression level by co-expression of Get3, but subsequent separation of the two proteins remains challenging and further conditions can be tested to solve this separation problem. It is also conceivable that other transmembrane domains in the presence of a well-suited detergent may be more efficiently separated from Get3, as the detergent provides an appropriate membrane mimetic. In general, the application of Get3 appears to be limited to those approaches in which Get3 can be separated from the target protein without unfolding a cytosolic domain through harsh or denaturing conditions, as the use of mild conditions may not allow the separation from Get3.

5. REFERENCES

[1] Schuhmann, P., Engstler, C., Klöpfer, K., Gügel, I. L., Abbadi, A., Dreyer, F., Leckband, G., Bölter, B., Hagn, F., Soll, J. & Carrie, C. 2022. Article Two wrongs make a right: Heat stress reversion of a male-sterile *Brassica napus* line. *J Exp Bot*.

[2] Klöpfer, K. & Hagn, F. 2019. Beyond detergent micelles: The advantages and applications of non-micellar and lipid-based membrane mimetics for solution-state NMR. *Prog Nucl Magn Reson Spectrosc,* 114-115**,** 271-283.

[3] Vinothkumar, K. R. & Henderson, R. 2010. Structures of membrane proteins. *Q Rev Biophys,* 43**,** 65-158.

[4] Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K. & Walter, P. 2015. *Molecular Biology of the Cell (Sixth Edition)*. Garland Science, Taylor & Francis Group.

[5] Almeida, J. G., Preto, A. J., Koukos, P. I., Bonvin, A. & Moreira, I. S. 2017. Membrane proteins structures: A review on computational modeling tools. *Biochim Biophys Acta Biomembr,* 1859**,** 2021-2039.

[6] Macher, B. A. & Yen, T. Y. 2007. Proteins at membrane surfaces-a review of approaches. *Mol Biosyst,* 3**,** 705-13.

[7] Rask-Andersen, M., Almén, M. S. & Schiöth, H. B. 2011. Trends in the exploitation of novel drug targets. *Nat Rev Drug Discov,* 10**,** 579-90.

[8] Doerr, A. 2009. Membrane protein structures. *Nature Methods,* 6**,** 35-35.

[9] Seddon, A. M., Curnow, P. & Booth, P. J. 2004. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta,* 1666**,** 105-17.

[10] Drew, D., Fröderberg, L., Baars, L. & de Gier, J. W. 2003. Assembly and overexpression of membrane proteins in *Escherichia coli*. *Biochim Biophys Acta,* 1610**,** 3-10.

[11] Watson, H. 2015. Biological membranes. *Essays Biochem,* 59**,** 43-69.

[12] Singer, S. J. & Nicolson, G. L. 1972. The fluid mosaic model of the structure of cell membranes. *Science,* 175**,** 720-31.

[13] Linke, D. 2009. Detergents: an overview. *Methods Enzymol,* 463**,** 603-17.

[14] Malia, T. J. & Wagner, G. 2007. NMR structural investigation of the mitochondrial outer membrane protein VDAC and its interaction with antiapoptotic Bcl-x$_L$. *Biochemistry,* 46**,** 514-25.

[15] Losonczi, J. A., Olejniczak, E. T., Betz, S. F., Harlan, J. E., Mack, J. & Fesik, S. W. 2000. NMR studies of the anti-apoptotic protein Bcl-x$_L$ in micelles. *Biochemistry,* 39**,** 11024-33.

[16] Zhou, H. X. & Cross, T. A. 2013. Influences of membrane mimetic environments on membrane protein structures. *Annu Rev Biophys,* 42**,** 361-92.

[17] Pabst, G., Kučerka, N., Nieh, M. P. & Katsaras, J. 2014. *Liposomes, Lipid Bilayers and Model Membranes: From Basic Research to Application*. CRC Press, Taylor & Francis Group.

[18] Lee, A. G. 2003. Lipid-protein interactions in biological membranes: a structural perspective. *Biochim Biophys Acta,* 1612**,** 1-40.

[19] Sanders, C. R., 2nd & Landis, G. C. 1995. Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies. *Biochemistry,* 34**,** 4030-40.

[20] Tribet, C., Audebert, R. & Popot, J. L. 1996. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci U S A,* 93**,** 15047-50.

[21] Popot, J. L., Althoff, T., Bagnard, D., Banères, J. L., Bazzacco, P., Billon-Denis, E., Catoire, L. J., Champeil, P., Charvolin, D., Cocco, M. J., Crémel, G., Dahmane, T., de la Maza, L. M., Ebel, C., Gabel, F., Giusti, F., Gohon, Y., Goormaghtigh, E., Guittet, E., Kleinschmidt, J. H., Kühlbrandt, W., Le Bon, C., Martinez, K. L., Picard, M., Pucci, B., Sachs, J. N., Tribet, C., van Heijenoort, C., Wien, F., Zito, F. & Zoonens, M. 2011. Amphipols from A to Z. *Annu Rev Biophys,* 40**,** 379-408.

[22] Fang, Y., Gursky, O. & Atkinson, D. 2003. Lipid-binding studies of human apolipoprotein A-I and its terminally truncated mutants. *Biochemistry,* 42**,** 13260-8.

[23] Krieger, M., Brown, M. S., Faust, J. R. & Goldstein, J. L. 1978. Replacement of endogenous cholesteryl esters of low density lipoprotein with exogenous cholesteryl linoleate. Reconstitution of a biologically active lipoprotein particle. *J Biol Chem,* 253**,** 4093-101.

[24] Denisov, I. G., Grinkova, Y. V., Lazarides, A. A. & Sligar, S. G. 2004. Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *J Am Chem Soc,* 126**,** 3477-87.

[25] Bayburt, T. H., Grinkova, Y. V. & Sligar, S. G. 2002. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Letters,* 2**,** 853-856.

[26] Nasr, M. L., Baptista, D., Strauss, M., Sun, Z. J., Grigoriu, S., Huser, S., Plückthun, A., Hagn, F., Walz, T., Hogle, J. M. & Wagner, G. 2017. Covalently circularized nanodiscs for studying membrane proteins and viral entry. *Nat Methods,* 14**,** 49-52.

[27] Miehling, J., Goricanec, D. & Hagn, F. 2018. A Split-Intein-Based Method for the Efficient Production of Circularized Nanodiscs for Structural Studies of Membrane Proteins. *Chembiochem,* 19**,** 1927-1933.

[28] Flayhan, A., Mertens, H. D. T., Ural-Blimke, Y., Martinez Molledo, M., Svergun, D. I. & Löw, C. 2018. Saposin Lipid Nanoparticles: A Highly Versatile and Modular Tool for Membrane Protein Research. *Structure,* 26**,** 345-355 e5.

[29] Frauenfeld, J., Löving, R., Armache, J. P., Sonnen, A. F., Guettou, F., Moberg, P., Zhu, L., Jegerschöld, C., Flayhan, A., Briggs, J. A., Garoff, H., Löw, C., Cheng, Y. & Nordlund, P. 2016. A saposin-lipoprotein nanoparticle system for membrane proteins. *Nat Methods,* 13**,** 345-51.

[30] Popovic, K., Holyoake, J., Pomès, R. & Privé, G. G. 2012. Structure of saposin A lipoprotein discs. *Proceedings of the National Academy of Sciences,* 109**,** 2908-2912.

[31] Tonge, S. R. & Tighe, B. J. 2001. Responsive hydrophobically associating polymers: a review of structure and properties. *Adv Drug Deliv Rev,* 53**,** 109-22.

[32] Knowles, T. J., Finka, R., Smith, C., Lin, Y. P., Dafforn, T. & Overduin, M. 2009. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc,* 131**,** 7484-5.

[33] Ravula, T., Hardin, N. Z., Ramadugu, S. K., Cox, S. J. & Ramamoorthy, A. 2018. Formation of pH-Resistant Monodispersed Polymer-Lipid Nanodiscs. *Angew Chem Int Ed Engl,* 57**,** 1342-1345.

[34] Oluwole, A. O., Danielczak, B., Meister, A., Babalola, J. O., Vargas, C. & Keller, S. 2017. Solubilization of Membrane Proteins into Functional Lipid-Bilayer Nanodiscs Using a Diisobutylene/Maleic Acid Copolymer. *Angew Chem Int Ed Engl,* 56**,** 1919-1924.

[35] Long, A. R., O'Brien, C. C., Malhotra, K., Schwall, C. T., Albert, A. D., Watts, A. & Alder, N. N. 2013. A detergent-free strategy for the reconstitution of active enzyme complexes from native biological membranes into nanoscale discs. *BMC Biotechnol,* 13**,** 41.

[36] Anantharamaiah, G. M., Jones, J. L., Brouillette, C. G., Schmidt, C. F., Chung, B. H., Hughes, T. A., Bhown, A. S. & Segrest, J. P. 1985. Studies of synthetic peptide analogs of the amphipathic helix. Structure of complexes with dimyristoyl phosphatidylcholine. *J Biol Chem,* 260**,** 10248-55.

[37] Epand, R. M., Gawish, A., Iqbal, M., Gupta, K. B., Chen, C. H., Segrest, J. P. & Anantharamaiah, G. M. 1987. Studies of synthetic peptide analogs of the amphipathic helix. Effect of charge distribution, hydrophobicity, and secondary structure on lipid association and lecithin:cholesterol acyltransferase activation. *J Biol Chem,* 262**,** 9389-96.

[38] Midtgaard, S. R., Pedersen, M. C., Kirkensgaard, J. J., Sørensen, K. K., Mortensen, K., Jensen, K. J. & Arleth, L. 2014. Self-assembling peptides form nanodiscs that stabilize membrane proteins. *Soft Matter,* 10**,** 738-52.

[39] Kondo, H., Ikeda, K. & Nakano, M. 2016. Formation of size-controlled, denaturation-resistant lipid nanodiscs by an amphiphilic self-polymerizing peptide. *Colloids Surf B Biointerfaces,* 146**,** 423-30.

[40] Stillwell, W. 2016. *An Introduction to Biological Membranes: Composition, Structure and Function (Second Edition)*. Elsevier.

[41] Yeagle, P. L. 2016. *The Membranes of Cells (Third Edition)*. Academic Press.

[42] Chou, K. C. & Elrod, D. W. 1999. Prediction of membrane protein types and subcellular locations. *Proteins,* 34**,** 137-53.

[43] Bölter, B., Soll, J., Hill, K., Hemmler, R. & Wagner, R. 1999. A rectifying ATP-regulated solute channel in the chloroplastic outer envelope from pea. *EMBO J,* 18**,** 5505-16.

[44] Chou, M. L., Fitzpatrick, L. M., Tu, S. L., Budziszewski, G., Potter-Lewis, S., Akita, M., Levin, J. Z., Keegstra, K. & Li, H. M. 2003. Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. *EMBO J,* 22**,** 2970-80.

[45] Düsterhöft, S., Babendreyer, A., Giese, A. A., Flasshove, C. & Ludwig, A. 2019. Status update on iRhom and ADAM17: It's still complicated. *Biochim Biophys Acta Mol Cell Res,* 1866**,** 1567-1583.

[46] Huxley-Jones, J., Clarke, T.-K., Beck, C., Toubaris, G., Robertson, D. L. & Boot-Handford, R. P. 2007. The evolution of the vertebrate metzincins; insights from *Ciona intestinalis* and *Danio rerio*. *BMC Evolutionary Biology,* 7**,** 63.

[47] Lemberg, M. K. & Freeman, M. 2007. Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases. *Genome Res,* 17**,** 1634-46.

[48] Kovács-Bogdán, E., Soll, J. & Bölter, B. 2010. Protein import into chloroplasts: the Tic complex and its regulation. *Biochim Biophys Acta,* 1803**,** 740-7.

[49] Soll, J. & Schleiff, E. 2004. Protein import into chloroplasts. *Nat Rev Mol Cell Biol,* 5**,** 198-208.

[50] Kao, Y. F., Lou, Y. C., Yeh, Y. H., Hsiao, C. D. & Chen, C. 2012. Solution structure of the C-terminal NP-repeat domain of Tic40, a co-chaperone during protein import into chloroplasts. *J Biochem,* 152**,** 443-51.

[51] Sveshnikova, N., Soll, J. & Schleiff, E. 2000. Toc34 is a preprotein receptor regulated by GTP and phosphorylation. *Proc Natl Acad Sci U S A,* 97**,** 4973-8.

[52] Young, M. E., Keegstra, K. & Froehlich, J. E. 1999. GTP promotes the formation of early-import intermediates but is not required during the translocation step of protein import into chloroplasts. *Plant Physiol,* 121**,** 237-44.

[53] Jackson-Constan, D., Akita, M. & Keegstra, K. 2001. Molecular chaperones involved in chloroplast protein import. *Biochim Biophys Acta,* 1541**,** 102-13.

[54] Stahl, T., Glockmann, C., Soll, J. & Heins, L. 1999. Tic40, a new "old" subunit of the chloroplast protein import translocon. *J Biol Chem,* 274**,** 37467-72.

[55] Schnell, D. J., Blobel, G., Keegstra, K., Kessler, F., Ko, K. & Soll, J. 1997. A consensus nomenclature for the protein-import components of the chloroplast envelope. *Trends Cell Biol,* 7**,** 303-4.

[56] Andrès, C., Agne, B. & Kessler, F. 2010. The TOC complex: preprotein gateway to the chloroplast. *Biochim Biophys Acta,* 1803**,** 715-23.

[57] Lee, D. W. & Hwang, I. 2019. Protein import into chloroplasts via the Tic40-dependent and -independent pathways depends on the amino acid composition of the transit peptide. *Biochem Biophys Res Commun,* 518**,** 66-71.

[58] Barth, M. A., Soll, J. & Akbaş, Ş. 2022. Prokaryotic and eukaryotic traits support the biological role of the chloroplast outer envelope. *Biochim Biophys Acta Mol Cell Res,* 1869**,** 119224.

[59] Breuers, F. K., Bräutigam, A. & Weber, A. P. 2011. The Plastid Outer Envelope - A Highly Dynamic Interface between Plastid and Cytoplasm. *Front Plant Sci,* 2**,** 97.

[60] Gross, L. E., Klinger, A., Spies, N., Ernst, T., Flinner, N., Simm, S., Ladig, R., Bodensohn, U. & Schleiff, E. 2021. Insertion of plastidic β-barrel proteins into the outer envelopes of plastids involves an intermembrane space intermediate formed with Toc75-V/OEP80. *Plant Cell,* 33**,** 1657-1681.

[61] McClain, A. M. & Sharkey, T. D. 2019. Triose phosphate utilization and beyond: from photosynthesis to end product synthesis. *J Exp Bot,* 70**,** 1755-1766.

[62] Heineke, D., Riens, B., Grosse, H., Hoferichter, P., Peter, U., Flügge, U.-I. & Heldt, H. W. 1991. Redox Transfer across the Inner Chloroplast Envelope Membrane 1. *Plant Physiology,* 95**,** 1131-1137.

[63] Flügge, U. I. 1998. Metabolite transporters in plastids. *Curr Opin Plant Biol,* 1**,** 201-6.

[64] Fischer, K. 2011. The import and export business in plastids: transport processes across the inner envelope membrane. *Plant Physiol,* 155**,** 1511-9.

[65] Trentmann, O., Mühlhaus, T., Zimmer, D., Sommer, F., Schroda, M., Haferkamp, I., Keller, I., Pommerrenig, B. & Neuhaus, H. E. 2020. Identification of Chloroplast Envelope Proteins with Critical Importance for Cold Acclimation. *Plant Physiol,* 182**,** 1239-1255.

[66] Flügge, U. I., Häusler, R. E., Ludewig, F. & Gierth, M. 2011. The role of transporters in supplying energy to plant plastids. *J Exp Bot,* 62**,** 2381-92.

[67] Lee, D. W., Lee, J. & Hwang, I. 2017. Sorting of nuclear-encoded chloroplast membrane proteins. *Curr Opin Plant Biol,* 40**,** 1-7.

[68] Flügge, U. I. & Benz, R. 1984. Pore-forming activity in the outer membrane of the chloroplast envelope. *FEBS Letters,* 169**,** 85-89.

[69] Flügge, U. I. 2000. Transport in and out of plastids: does the outer envelope membrane control the flow? *Trends Plant Sci,* 5**,** 135-7.

[70] Bölter, B. & Soll, J. 2001. Ion channels in the outer membranes of chloroplasts and mitochondria: open doors or regulated gates? *EMBO J,* 20**,** 935-40.

[71] Pottosin, I. & Shabala, S. 2016. Transport Across Chloroplast Membranes: Optimizing Photosynthesis for Adverse Environmental Conditions. *Mol Plant,* 9**,** 356-370.

[72] Pohlmeyer, K., Soll, J., Steinkamp, T., Hinnah, S. & Wagner, R. 1997. Isolation and characterization of an amino acid-selective channel protein present in the chloroplastic outer envelope membrane. *Proc Natl Acad Sci U S A,* 94**,** 9504-9.

[73] Götze, T. A., Patil, M., Jeshen, I., Bölter, B., Grahl, S. & Soll, J. 2015. Oep23 forms an ion channel in the chloroplast outer envelope. *BMC Plant Biol,* 15**,** 47.

[74] Pohlmeyer, K., Soll, J., Grimm, R., Hill, K. & Wagner, R. 1998. A high-conductance solute channel in the chloroplastic outer envelope from Pea. *Plant Cell,* 10**,** 1207-16.

[75] Götze, T. A., Philippar, K., Ilkavets, I., Soll, J. & Wagner, R. 2006. OEP37 is a new member of the chloroplast outer membrane ion channels. *J Biol Chem,* 281**,** 17989-98.

[76] Bräutigam, A., Hoffmann-Benning, S. & Weber, A. P. 2008. Comparative proteomics of chloroplast envelopes from C_3 and C_4 plants reveals specific adaptations of the plastid envelope to C_4 photosynthesis and candidate proteins required for maintaining C_4 metabolite fluxes. *Plant Physiol,* 148**,** 568-79.

[77] Facchinelli, F. & Weber, A. 2011. The Metabolite Transporters of the Plastid Envelope: An Update. *Frontiers in Plant Science,* 2.

[78] Hemmler, R., Becker, T., Schleiff, E., Bölter, B., Stahl, T., Soll, J., Götze, T. A., Braams, S. & Wagner, R. 2006. Molecular properties of Oep21, an ATP-regulated anion-selective solute channel from the outer chloroplast membrane. *J Biol Chem,* 281**,** 12020-9.

[79] Blatch, G. L. & Lässle, M. 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays,* 21**,** 932-9.

[80] Bédard, J., Kubis, S., Bimanadham, S. & Jarvis, P. 2007. Functional similarity between the chloroplast translocon component, Tic40, and the human co-chaperone, Hsp70-interacting protein (Hip). *J Biol Chem,* 282**,** 21404-14.

[81] Chou, M. L., Chu, C. C., Chen, L. J., Akita, M. & Li, H. M. 2006. Stimulation of transit-peptide release and ATP hydrolysis by a cochaperone during protein import into chloroplasts. *J Cell Biol,* 175**,** 893-900.

[82] Frydman, J. & Höhfeld, J. 1997. Chaperones get in touch: the Hip-Hop connection. *Trends Biochem Sci,* 22**,** 87-92.

[83] Chiu, C. C. & Li, H. M. 2008. Tic40 is important for reinsertion of proteins from the chloroplast stroma into the inner membrane. *Plant J,* 56**,** 793-801.

[84] Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C. F., Tolmie, D. E., Kent Wenger, R., Yao, H. & Markley, J. L. 2008. BioMagResBank. *Nucleic Acids Research,* 36**,** D402-D408.

[85] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. 2000. The Protein Data Bank. *Nucleic Acids Research,* 28**,** 235-242.

[86] Calligaris, M., Cuffaro, D., Bonelli, S., Spanò, D. P., Rossello, A., Nuti, E. & Scilabra, S. D. 2021. Strategies to Target ADAM17 in Disease: From its Discovery to the iRhom Revolution. *Molecules,* 26.

[87] Lichtenthaler, S. F., Lemberg, M. K. & Fluhrer, R. 2018. Proteolytic ectodomain shedding of membrane proteins in mammals-hardware, concepts, and recent developments. *EMBO J,* 37.

[88] Black, R. A., Rauch, C. T., Kozlosky, C. J., Peschon, J. J., Slack, J. L., Wolfson, M. F., Castner, B. J., Stocking, K. L., Reddy, P., Srinivasan, S., Nelson, N., Boiani, N., Schooley, K. A., Gerhart, M., Davis, R., Fitzner, J. N., Johnson, R. S., Paxton, R. J., March, C. J. & Cerretti, D. P. 1997. A metalloproteinase disintegrin that releases tumour-necrosis factor-alpha from cells. *Nature,* 385**,** 729-33.

[89] Hsia, H. E., Tüshaus, J., Brummer, T., Zheng, Y., Scilabra, S. D. & Lichtenthaler, S. F. 2019. Functions of 'A disintegrin and metalloproteases (ADAMs)' in the mammalian nervous system. *Cell Mol Life Sci,* 76**,** 3055-3081.

[90] Nagase, H., Visse, R. & Murphy, G. 2006. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovasc Res,* 69**,** 562-73.

[91] Srour, N., Lebel, A., McMahon, S., Fournier, I., Fugère, M., Day, R. & Dubois, C. M. 2003. TACE/ADAM-17 maturation and activation of sheddase activity require proprotein convertase activity. *FEBS Lett,* 554**,** 275-83.

[92] Zunke, F. & Rose-John, S. 2017. The shedding protease ADAM17: Physiology and pathophysiology. *Biochim Biophys Acta Mol Cell Res,* 1864**,** 2059-2070.

[93] Peschon, J. J., Slack, J. L., Reddy, P., Stocking, K. L., Sunnarborg, S. W., Lee, D. C., Russell, W. E., Castner, B. J., Johnson, R. S., Fitzner, J. N., Boyce, R. W., Nelson, N., Kozlosky, C. J., Wolfson, M. F., Rauch, C. T., Cerretti, D. P., Paxton, R. J., March, C. J. & Black, R. A. 1998. An essential role for ectodomain shedding in mammalian development. *Science,* 282**,** 1281-4.

[94] Hosur, V., Farley, M. L., Burzenski, L. M., Shultz, L. D. & Wiles, M. V. 2018. ADAM17 is essential for ectodomain shedding of the EGF-receptor ligand amphiregulin. *FEBS Open Bio,* 8**,** 702-710.

[95] Sunnarborg, S. W., Hinkle, C. L., Stevenson, M., Russell, W. E., Raska, C. S., Peschon, J. J., Castner, B. J., Gerhart, M. J., Paxton, R. J., Black, R. A. & Lee, D. C. 2002. Tumor necrosis factor-alpha converting enzyme (TACE) regulates epidermal growth factor receptor ligand availability. *J Biol Chem,* 277**,** 12838-45.

[96] Yan, I., Schwarz, J., Lücke, K., Schumacher, N., Schumacher, V., Schmidt, S., Rabe, B., Saftig, P., Donners, M., Rose-John, S., Mittrücker, H. W. & Chalaris, A. 2016. ADAM17 controls IL-6 signaling by cleavage of the murine IL-6Rα from the cell surface of leukocytes during inflammatory responses. *J Leukoc Biol,* 99**,** 749-60.

[97] Lee, M. Y., Nam, K. H. & Choi, K. C. 2016. iRhoms; Its Functions and Essential Roles. *Biomol Ther (Seoul),* 24**,** 109-14.

[98] Lisi, S., D'Amore, M. & Sisto, M. 2014. ADAM17 at the interface between inflammation and autoimmunity. *Immunol Lett,* 162**,** 159-69.

[99] Al-Salihi, M. A. & Lang, P. A. 2020. iRhom2: An Emerging Adaptor Regulating Immunity and Disease. *International Journal of Molecular Sciences,* 21**,** 6570.

[100] Düsterhöft, S., Künzel, U. & Freeman, M. 2017. Rhomboid proteases in human disease: Mechanisms and future prospects. *Biochim Biophys Acta Mol Cell Res,* 1864**,** 2200-2209.

[101] Mayer, U. & Nüsslein-Volhard, C. 1988. A group of genes required for pattern formation in the ventral ectoderm of the Drosophila embryo. *Genes Dev,* 2**,** 1496-511.

[102] Freeman, M. 2016. Rhomboids, signalling and cell biology. *Biochem Soc Trans,* 44**,** 945-50.

[103] Adrain, C., Zettl, M., Christova, Y., Taylor, N. & Freeman, M. 2012. Tumor necrosis factor signaling requires iRhom2 to promote trafficking and activation of TACE. *Science,* 335**,** 225-8.

[104] McIlwain, D. R., Lang, P. A., Maretzky, T., Hamada, K., Ohishi, K., Maney, S. K., Berger, T., Murthy, A., Duncan, G., Xu, H. C., Lang, K. S., Häussinger, D., Wakeham, A., Itie-Youten, A., Khokha, R., Ohashi, P. S., Blobel, C. P. & Mak, T. W. 2012. iRhom2 regulation of TACE controls TNF-mediated protection against *Listeria* and responses to LPS. *Science,* 335**,** 229-32.

[105] Cavadas, M., Oikonomidi, I., Gaspar, C. J., Burbridge, E., Badenes, M., Félix, I., Bolado, A., Hu, T., Bileck, A., Gerner, C., Domingos, P. M., von Kriegsheim, A. & Adrain, C. 2017. Phosphorylation of iRhom2 Controls Stimulated Proteolytic Shedding by the Metalloprotease ADAM17/TACE. *Cell Rep,* 21**,** 745-757.

[106] Grieve, A. G., Xu, H., Künzel, U., Bambrough, P., Sieber, B. & Freeman, M. 2017. Phosphorylation of iRhom2 at the plasma membrane controls mammalian TACE-dependent inflammatory and growth factor signalling. *Elife,* 6.

[107] Baker, R. P., Young, K., Feng, L., Shi, Y. & Urban, S. 2007. Enzymatic analysis of a rhomboid intramembrane protease implicates transmembrane helix 5 as the lateral substrate gate. *Proc Natl Acad Sci U S A,* 104**,** 8257-62.

[108] Wang, Y., Maegawa, S., Akiyama, Y. & Ha, Y. 2007. The role of L1 loop in the mechanism of rhomboid intramembrane protease GlpG. *J Mol Biol,* 374**,** 1104-13.

[109] Arutyunova, E., Panwar, P., Skiba, P. M., Gale, N., Mak, M. W. & Lemieux, M. J. 2014. Allosteric regulation of rhomboid intramembrane proteolysis. *EMBO J,* 33**,** 1869-81.

[110] Sampathkumar, P., Mak, M. W., Fischer-Witholt, S. J., Guigard, E., Kay, C. M. & Lemieux, M. J. 2012. Oligomeric state study of prokaryotic rhomboid proteases. *Biochim Biophys Acta,* 1818**,** 3090-7.

[111] Zoll, S., Stanchev, S., Began, J., Škerle, J., Lepšík, M., Peclinovská, L., Majer, P. & Strisovsky, K. 2014. Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures. *EMBO J,* 33**,** 2408-21.

[112] Borgese, N., Brambillasca, S. & Colombo, S. 2007. How tails guide tail-anchored proteins to their destinations. *Curr Opin Cell Biol,* 19**,** 368-75.

[113] Beilharz, T., Egan, B., Silver, P. A., Hofmann, K. & Lithgow, T. 2003. Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in *Saccharomyces cerevisiae*. *J Biol Chem,* 278**,** 8219-23.

[114] Mateja, A., Szlachcic, A., Downing, M. E., Dobosz, M., Mariappan, M., Hegde, R. S. & Keenan, R. J. 2009. The structural basis of tail-anchored membrane protein recognition by Get3. *Nature,* 461**,** 361-6.

[115] Favaloro, V., Spasic, M., Schwappach, B. & Dobberstein, B. 2008. Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *J Cell Sci,* 121**,** 1832-40.

[116] Stefanovic, S. & Hegde, R. S. 2007. Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell,* 128**,** 1147-59.

[117] Chang, Y. W., Lin, T. W., Li, Y. C., Huang, Y. S., Sun, Y. J. & Hsiao, C. D. 2012. Interaction surface and topology of Get3-Get4-Get5 protein complex, involved in targeting tail-anchored proteins to endoplasmic reticulum. *J Biol Chem,* 287**,** 4783-9.

[118] Mateja, A. & Keenan, R. J. 2018. A structural perspective on tail-anchored protein biogenesis by the GET pathway. *Curr Opin Struct Biol,* 51**,** 195-202.

[119] Mateja, A., Paduch, M., Chang, H. Y., Szydlowska, A., Kossiakoff, A. A., Hegde, R. S. & Keenan, R. J. 2015. Protein targeting. Structure of the Get3 targeting factor in complex with its membrane protein cargo. *Science,* 347**,** 1152-5.

[120] Bozkurt, G., Stjepanovic, G., Vilardi, F., Amlacher, S., Wild, K., Bange, G., Favaloro, V., Rippe, K., Hurt, E., Dobberstein, B. & Sinning, I. 2009. Structural insights into tail-anchored protein binding and membrane insertion by Get3. *Proc Natl Acad Sci U S A,* 106, 21131-6.

[121] Rule, G. S. & Hitchens, T. K. 2006. *Fundamentals of Protein NMR Spectroscopy*. Springer Dordrecht.

[122] Teng, Q. 2013. *Structural Biology: Practical NMR Applications (Second Edition)*. Springer Science & Business Media.

[123] Cavanagh, J., Fairbrother, W., Palmer III, A., Rance, M. & Skelton, N. 2007. *Principles and Practice: Protein NMR Spectroscopy (Second Edition)*. Elsevier Academic Press.

[124] Koharudin, L. M. I., Bonvin, A. M. J. J., Kaptein, R. & Boelens, R. 2003. Use of very long-distance NOEs in a fully deuterated protein: an approach for rapid protein fold determination. *Journal of Magnetic Resonance,* 163, 228-235.

[125] Griesinger, C., Sørensen, O. W. & Ernst, R. R. 1987. Novel three-dimensional NMR techniques for studies of peptides and biological macromolecules. *Journal of the American Chemical Society,* 109, 7227-7228.

[126] Zhang, O. & Forman-Kay, J. D. 1997. NMR Studies of Unfolded States of an SH3 Domain in Aqueous Solution and Denaturing Conditions. *Biochemistry,* 36, 3959-3970.

[127] Diercks, T., Coles, M. & Kessler, H. 1999. An efficient strategy for assignment of cross-peaks in 3D heteronuclear NOESY experiments. *Journal of Biomolecular NMR,* 15, 177-180.

[128] Fesik, S. W. & Zuiderweg, E. R. 1988. Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra. *Journal of Magnetic Resonance (1969),* 78, 588-593.

[129] Sattler, M., Schleucher, J. & Griesinger, C. 1999. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution. *Progress in nuclear magnetic resonance spectroscopy,* 34, 93-158.

[130] Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. 1990. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance (1969),* 89, 496-514.

[131] Grzesiek, S. & Bax, A. 1992. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *Journal of Magnetic Resonance (1969),* 96, 432-440.

[132] Muhandiram, D. R. & Kay, L. E. 1994. Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *Journal of Magnetic Resonance, Series B,* 103, 203-216.

[133] Clubb, R. T., Thanabal, V. & Wagner, G. 1992. A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $^{1}H^{N}$, ^{15}N, and $^{13}C'$ chemical shifts in ^{15}N-^{13}C-labelled proteins. *Journal of Magnetic Resonance (1969),* 97, 213-217.

[134] Kay, L. E., Xu, G. Y. & Yamazaki, T. 1994. Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation. *Journal of Magnetic Resonance, Series A,* 109, 129-133.

[135] Engelke, J. & Rüterjans, H. 1995. Sequential protein backbone resonance assignments using an improved 3D-HN (CA) CO pulse scheme. *Journal of Magnetic Resonance,* 109, 318-322.

[136] Ikura, M., Kay, L. E. & Bax, A. 1990. A novel approach for sequential assignment of ^{1}H, ^{13}C, and ^{15}N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry,* 29, 4659-67.

[137] Ikura, M., Marion, D., Kay, L. E., Shih, H., Krinks, M., Klee, C. B. & Bax, A. 1990. Heteronuclear 3D NMR and isotopic labeling of calmodulin. Towards the complete assignment of the ^{1}H NMR spectrum. *Biochem Pharmacol,* 40, 153-60.

[138] Farmer, B. T., Venters, R. A., Spicer, L. D., Wittekind, M. G. & Müller, L. 1992. A refocused and optimized HNCA: Increased sensitivity and resolution in large macromolecules. *Journal of Biomolecular NMR,* 2, 195-202.

[139] Kay, L. E., Ikura, M., Zhu, G. & Bax, A. 1991. Four-dimensional heteronuclear triple-resonance NMR of isotopically enriched proteins for sequential assignment of backbone atoms. *Journal of Magnetic Resonance (1969),* 91, 422-428.

[140] Bax, A. & Ikura, M. 1991. An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the α-carbon of the preceding residue in uniformly $^{15}N/^{13}C$ enriched proteins. *Journal of Biomolecular NMR,* 1, 99-104.

[141] Grzesiek, S. & Bax, A. 1992. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance (1969),* 99, 201-207.

[142] Grzesiek, S. & Bax, A. 1992. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *Journal of the American Chemical Society,* 114, 6291-6293.

[143] Berliner, L. 2015. *Protein NMR: Modern Techniques and Biomedical Applications*. Springer Science+Business Media, LLC.

[144] Grzesiek, S., Anglister, J. & Bax, A. 1993. Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in ^{13}C/^{15}N-Enriched Proteins by Isotropic Mixing of ^{13}C Magnetization. *Journal of Magnetic Resonance, Series B,* 101**,** 114-119.

[145] Montelione, G. T., Lyons, B. A., Emerson, S. D. & Tashiro, M. 1992. An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. *Journal of the American Chemical Society,* 114**,** 10974-10975.

[146] Vuister, G. W. & Bax, A. 1993. Quantitative *J* correlation: a new approach for measuring homonuclear three-bond $J(H^NH^\alpha)$ coupling constants in ^{15}N-enriched proteins. *Journal of the American Chemical Society,* 115**,** 7772-7777.

[147] Archer, S. J., Ikura, M., Torchia, D. A. & Bax, A. 1991. An alternative 3D NMR technique for correlating backbone ^{15}N with side chain Hβ resonances in larger proteins. *Journal of Magnetic Resonance (1969),* 95**,** 636-641.

[148] Hu, Q. 2021. NMR Approaches to Determine Protein Structure. *In:* BELL, E. (ed.) *Molecular Life Sciences: An Encyclopedic Reference.* New York, NY: Springer New York.

[149] Wishart, D. S. & Sykes, B. D. 1994. The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J Biomol NMR,* 4**,** 171-80.

[150] Wishart, D. S. & Sykes, B. D. 1994. Chemical shifts as a tool for structure determination. *Methods in Enzymology.* Academic Press.

[151] Cornilescu, G., Delaglio, F. & Bax, A. 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of Biomolecular NMR,* 13**,** 289-302.

[152] Karplus, M. 1963. Vicinal Proton Coupling in Nuclear Magnetic Resonance. *Journal of the American Chemical Society,* 85**,** 2870-2871.

[153] Wang, A. C. & Bax, A. 1996. Determination of the Backbone Dihedral Angles φ in Human Ubiquitin from Reparametrized Empirical Karplus Equations. *Journal of the American Chemical Society,* 118**,** 2483-2494.

[154] Wüthrich, K. 1986. *NMR of proteins and nucleic acids*. John Wiley & Sons.

[155] Wüthrich, K., Billeter, M. & Braun, W. 1995. Polypeptide Secondary Structure Determination by Nuclear Magnetic Resonance Observation of Short Proton?Proton Distances. *NMR in Structural Biology.* WORLD SCIENTIFIC.

[156] Clore, G. M. & Gronenborn, A. M. 1989. Determination of Three-Dimensional Structures of Proteins and Nucleic Acids in Solution by Nuclear Magnetic Resonance Spectroscop. *Critical Reviews in Biochemistry and Molecular Biology,* 24**,** 479-564.

[157] Lovell, S. C., Word, J. M., Richardson, J. S. & Richardson, D. C. 2000. The penultimate rotamer library. *Proteins: Structure, Function, and Bioinformatics,* 40**,** 389-408.

[158] Wüthrich, K., Billeter, M. & Braun, W. 1983. Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *Journal of Molecular Biology,* 169**,** 949-961.

[159] Harsch, T., Schneider, P., Kieninger, B., Donaubauer, H. & Kalbitzer, H. R. 2017. Stereospecific assignment of the asparagine and glutamine sidechain amide protons in proteins from chemical shift analysis. *Journal of Biomolecular NMR,* 67**,** 157-164.

[160] Vis, H., Boelens, R., Mariani, M., Stroop, R., Vorgias, C. E., Wilson, K. S. & Kaptein, R. 1994. ^{1}H, ^{13}C, and ^{15}N resonance assignments and secondary structure analysis of the HU protein from Bacillus stearothermophilus using two- and three-dimensional double- and triple-resonance heteronuclear magnetic resonance spectroscopy. *Biochemistry,* 33**,** 14858-14870.

[161] MacArthur, M. W. & Thornton, J. M. 1991. Influence of proline residues on protein conformation. *Journal of Molecular Biology,* 218**,** 397-412.

[162] Schubert, M., Labudde, D., Oschkinat, H. & Schmieder, P. 2002. A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ^{13}C chemical shift statistics. *Journal of Biomolecular NMR,* 24**,** 149-154.

[163] Güntert, P. 2003. Automated NMR protein structure calculation. *Progress in Nuclear Magnetic Resonance Spectroscopy,* 43**,** 105-125.

[164] Güntert, P. 1998. Structure calculation of biological macromolecules from NMR data. *Quarterly Reviews of Biophysics,* 31**,** 145-237.

[165] Grishaev, A. & Llinás, M. 2005. Protein structure elucidation from minimal NMR data: The CLOUDS approach. *Methods in enzymology.* Elsevier.

[166] Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. 1996. Discovering high-affinity ligands for proteins: SAR by NMR. *Science,* 274, 1531-1534.

[167] Dehner, A., Furrer, J., Richter, K., Schuster, I., Buchner, J. & Kessler, H. 2003. NMR Chemical Shift Perturbation Study of the N-Terminal Domain of Hsp90 upon Binding of ADP, AMP-PNP, Geldanamycin, and Radicicol. *ChemBioChem,* 4, 870-877.

[168] Fielding, L. 2003. NMR methods for the determination of protein-ligand dissociation constants. *Curr Top Med Chem,* 3, 39-53.

[169] Lee, W., Tonelli, M. & Markley, J. L. 2015. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics,* 31, 1325-7.

[170] Häusler, E. 2020. NMR structural and biochemical studies on the β-barrel membrane proteins OEP21 and hVDAC1. Doctoral Dissertation (Dr. rer. nat), *Technische Universität München*.

[171] Raltchev, K. P. 2020. NMR structural studies of full-length Bcl-xL in nanodiscs. Doctoral Dissertation (Dr. rer. nat.), *Technische Universität München*.

[172] Arcemisbéhère, L., Sen, T., Boudier, L., Balestre, M. N., Gaibelet, G., Detouillon, E., Orcel, H., Mendre, C., Rahmeh, R., Granier, S., Vivès, C., Fieschi, F., Damian, M., Durroux, T., Banères, J. L. & Mouillac, B. 2010. Leukotriene BLT2 receptor monomers activate the G$_{i2}$ GTP-binding protein more efficiently than dimers. *J Biol Chem,* 285, 6337-47.

[173] Kelly, S. M., Jess, T. J. & Price, N. C. 2005. How to study proteins by circular dichroism. *Biochim Biophys Acta,* 1751, 119-39.

[174] Wei, Y., Thyparambil, A. A. & Latour, R. A. 2014. Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230nm. *Biochim Biophys Acta,* 1844, 2331-7.

[175] Strop, P. & Brunger, A. T. 2005. Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci,* 14, 2207-11.

[176] Sreerama, N. & Woody, R. W. 2004. Computation and analysis of protein circular dichroism spectra. *Methods Enzymol,* 383, 318-51.

[177] Keegstra, K. & Yousif, A. E. 1986. Isolation and characterization of chloroplast envelope membranes. *Methods in Enzymology.* Academic Press.

[178] Waegemann, K., Eichacker, S. & Soll, J. 1992. Outer envelope membranes from chloroplasts are isolated as right-side-out vesicles. *Planta,* 187, 89-94.

[179] Morihara, K. & Tsuzuki, H. 1970. Thermolysin: Kinetic Study with Oligopeptides. *European Journal of Biochemistry,* 15, 374-380.

[180] Inagami, T. & Mitsuda, H. 1964. The Mechanism of the Specificity of Trypsin Catalysis: II. COMPARISON OF TRYPSIN AND α-CHYMOTRYPSIN IN THE NONSPECIFIC CATALYSES OF THE HYDROLYSIS OF ACETYLGLYCINE ETHYL ESTER. *Journal of Biological Chemistry,* 239, 1388-1394.

[181] Martin, O. A., Villegas, M. E., Vila, J. A. & Scheraga, H. A. 2010. Analysis of $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts of cysteine and cystine residues in proteins: a quantum chemical approach. *J Biomol NMR,* 46, 217-25.

[182] Voon, C. P., Guan, X., Sun, Y., Sahu, A., Chan, M. N., Gardeström, P., Wagner, S., Fuchs, P., Nietzel, T., Versaw, W. K., Schwarzländer, M. & Lim, B. L. 2018. ATP compartmentation in plastids and cytosol of *Arabidopsis thaliana* revealed by fluorescent protein sensing. *Proc Natl Acad Sci U S A,* 115, E10778-E10787.

[183] Hwang, T.-L., van Zijl, P. C. M. & Mori, S. 1998. Accurate Quantitation of Water-amide Proton Exchange Rates Using the Phase-Modulated CLEAN Chemical EXchange (CLEANEX-PM) Approach with a Fast-HSQC (FHSQC) Detection Scheme. *Journal of Biomolecular NMR,* 11, 221-226.

[184] Gupta, R. K., Gupta, P., Yushok, W. D. & Rose, Z. B. 1983. Measurement of the dissociation constant of MgATP at physiological nucleotide levels by a combination of ^{31}P NMR and optical absorbance spectroscopy. *Biochem Biophys Res Commun,* 117, 210-6.

[185] Portis, A. R., Jr. & Heldt, H. W. 1976. Light-dependent changes of the Mg^{2+} concentration in the stroma in relation to the Mg^{2+} dependency of CO$_2$ fixation in intact chloroplasts. *Biochim Biophys Acta,* 449, 434-6.

[186] Zhang, B., Zhang, C., Tang, R., Zheng, X., Zhao, F., Fu, A., Lan, W. & Luan, S. 2022. Two magnesium transporters in the chloroplast inner envelope essential for thylakoid biogenesis in Arabidopsis. *New Phytol*.

[187] Rausch, C. & Bucher, M. 2002. Molecular mechanisms of phosphate transport in plants. *Planta,* 216, 23-37.

[188] Nelson, J. W. & Kallenbach, N. R. 1986. Stabilization of the ribonuclease S-peptide α-helix by trifluoroethanol. *Proteins: Structure, Function, and Bioinformatics,* 1, 211-217.

[189] Shiraki, K., Nishikawa, K. & Goto, Y. 1995. Trifluoroethanol-induced Stabilization of the α-Helical Structure of β-Lactoglobulin: Implication for Non-hierarchical Protein Folding. *Journal of Molecular Biology,* 245**,** 180-194.

[190] Kaczka, P., Polkowska-Nowakowska, A., Bolewska, K., Zhukov, I., Poznański, J. & Wierzchowski, K. L. 2010. Backbone dynamics of TFE-induced native-like fold of region 4 of *Escherichia coli* RNA polymerase σ^{70} subunit. *Proteins,* 78**,** 754-68.

[191] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature,* 596**,** 583-589.

[192] Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S. A. A., Potapenko, A., Ballard, A. J., Romera-Paredes, B., Nikolov, S., Jain, R., Clancy, E., Reiman, D., Petersen, S., Senior, A. W., Kavukcuoglu, K., Birney, E., Kohli, P., Jumper, J. & Hassabis, D. 2021. Highly accurate protein structure prediction for the human proteome. *Nature,* 596**,** 590-596.

[193] Moreau, P., Bessoule, J. J., Mongrand, S., Testet, E., Vincent, P. & Cassagne, C. 1998. Lipid trafficking in plant cells. *Progress in Lipid Research,* 37**,** 371-391.

[194] Balint-Kurti, P. 2019. The plant hypersensitive response: concepts, control and consequences. *Molecular Plant Pathology,* 20**,** 1163-1178.

[195] Hatsugai, N., Perez Koldenkova, V., Imamura, H., Noji, H. & Nagai, T. 2012. Changes in Cytosolic ATP Levels and Intracellular Morphology during Bacteria-Induced Hypersensitive Cell Death as Revealed by Real-Time Fluorescence Microscopy Imaging. *Plant and Cell Physiology,* 53**,** 1768-1775.

[196] Zachariae, U., Schneider, R., Briones, R., Gattin, Z., Demers, J.-P., Giller, K., Maier, E., Zweckstetter, M., Griesinger, C., Becker, S., Benz, R., de Groot, Bert L. & Lange, A. 2012. β-Barrel Mobility Underlies Closure of the Voltage-Dependent Anion Channel. *Structure,* 20**,** 1540-1549.

[197] McClain, A. M. & Sharkey, T. D. 2019. Triose phosphate utilization and beyond: from photosynthesis to end product synthesis. *Journal of Experimental Botany,* 70**,** 1755-1766.

[198] Zahedi, R. P., Sickmann, A., Boehm, A. M., Winkler, C., Zufall, N., Schönfisch, B., Guiard, B., Pfanner, N. & Meisinger, C. 2006. Proteomic Analysis of the Yeast Mitochondrial Outer Membrane Reveals Accumulation of a Subclass of Preproteins. *Molecular Biology of the Cell,* 17**,** 1436-1450.

[199] Rostovtseva, T. & Colombini, M. 1996. ATP Flux Is Controlled by a Voltage-gated Channel from the Mitochondrial Outer Membrane. *Journal of Biological Chemistry,* 271**,** 28006-28008.

[200] Kreplak, J., Madoui, M.-A., Cápal, P., Novák, P., Labadie, K., Aubert, G., Bayer, P. E., Gali, K. K., Syme, R. A., Main, D., Klein, A., Bérard, A., Vrbová, I., Fournier, C., d'Agata, L., Belser, C., Berrabah, W., Toegelová, H., Milec, Z., Vrána, J., Lee, H., Kougbeadjo, A., Térézol, M., Huneau, C., Turo, C. J., Mohellibi, N., Neumann, P., Falque, M., Gallardo, K., McGee, R., Tar'an, B., Bendahmane, A., Aury, J.-M., Batley, J., Le Paslier, M.-C., Ellis, N., Warkentin, T. D., Coyne, C. J., Salse, J., Edwards, D., Lichtenzveig, J., Macas, J., Doležel, J., Wincker, P. & Burstin, J. 2019. A reference genome for pea provides insight into legume genome evolution. *Nature Genetics,* 51**,** 1411-1422.

[201] Gristick, H. B., Rao, M., Chartron, J. W., Rome, M. E., Shan, S.-o. & Clemons, W. M. 2014. Crystal structure of ATP-bound Get3–Get4–Get5 complex reveals regulation of Get3 by Get4. *Nature Structural & Molecular Biology,* 21**,** 437-442.

[202] Kubota, K., Yamagata, A., Sato, Y., Goto-Ito, S. & Fukai, S. 2012. Get1 Stabilizes an Open Dimer Conformation of Get3 ATPase by Binding Two Distinct Interfaces. *Journal of Molecular Biology,* 422**,** 366-375.

6. APPENDIX

TABLE 30: Amino acid sequences of the OEP21 constructs

The amino acid sequence, molecular weight, and extinction coefficient at 280 nm of the used OEP21 constructs are listed below. OEP21 (blue) was expressed with a His_{10}-tag (green). An additional arginine (red) was introduced into the sequence for loop five deletion. Based on OEP21 with a His_{10}-tag, several OEP21 variants were produced and their molecular weight as well as their extinction coefficient at 280 nm are listed. OEP21-ΔHis without a His_{10}-tag also lacked the two amino acids L and E (black) immediately before the His_{10}-tag.

OEP21-WT	
Number of amino acids:	189
Molecular weight:	22,058 $\text{g}\cdot\text{mol}^{-1}$
Extinction coefficient:	35,410 $\text{M}^{-1}\cdot\text{cm}^{-1}$

```
        10         20         30         40         50         60         70
METSLRYGGD SKALKIHAKE KLRIDTNTFF QVRGGLDTKT GQPSSGSALI RHFYPNFSAT LGVGVRYDKQ
        80         90        100        110        120        130        140
DSVGVRYAKN DKLRYTVLAK KTFPVTNDGL VNFKIKGGCD VDQDFKEWKS RGGAEFSWNV FNFQKDQDVR
       150        160        170        180        189
LRIGYEAFEQ VPYLQIRENN WTFNADYKGR WNVRYDLLEH HHHHHHHHH
```

OEP21-ΔLoop5	
Number of amino acids:	179
Molecular weight:	20,996 $\text{g}\cdot\text{mol}^{-1}$
Extinction coefficient:	33,920 $\text{M}^{-1}\cdot\text{cm}^{-1}$

```
        10         20         30         40         50         60         70
METSLRYGGD SKALKIHAKE KLRIDTNTFF QVRGGLDTKT GQPSSGSALI RHFYPNFSAT LGVGVRYDKR
        80         90        100        110        120        130        140
DKLRYTVLAK KTFPVTNDGL VNFKIKGGCD VDQDFKEWKS RGGAEFSWNV FNFQKDQDVR LRIGYEAFEQ
       150        160        170        179
VPYLQIRENN WTFNADYKGR WNVRYDLLEH HHHHHHHHH
```

OEP21 variant	**Molecular weight**	**Extinction coefficient**
K19A / K90A / K104A	22,001 $\text{g}\cdot\text{mol}^{-1}$	35,410 $\text{M}^{-1}\cdot\text{cm}^{-1}$
R33A / R51A / R66A / R84A / R140A / R142A / R157A / R174A	21,973 $\text{g}\cdot\text{mol}^{-1}$	35,410 $\text{M}^{-1}\cdot\text{cm}^{-1}$
R66A-N164R	22,015 $\text{g}\cdot\text{mol}^{-1}$	35,410 $\text{M}^{-1}\cdot\text{cm}^{-1}$
R66W	22,088 $\text{g}\cdot\text{mol}^{-1}$	40,910 $\text{M}^{-1}\cdot\text{cm}^{-1}$
C109A	22,026 $\text{g}\cdot\text{mol}^{-1}$	35,410 $\text{M}^{-1}\cdot\text{cm}^{-1}$
OEP21-ΔHis (without His_{10}-tag, L178, and E179: 177 amino acids)	20,444 $\text{g}\cdot\text{mol}^{-1}$	35,410 $\text{M}^{-1}\cdot\text{cm}^{-1}$

TABLE 31: Amino acid sequences of BnaA10 and BnaC9

The amino acid sequence, molecular weight, and extinction coefficient at 280 nm of BnaA10-Tic40 and BnaC9-Tic40 are listed below.

BnaA10-Tic40		**BnaC9-Tic40**	
Number of amino acids:	153	Number of amino acids:	151
Molecular weight:	17,640 g·mol^{-1}	Molecular weight:	17,522 g·mol^{-1}
Extinction coefficient:	15,470 M^{-1}·cm^{-1}	Extinction coefficient:	15,470 M^{-1}·cm^{-1}
305 315 325 335 GGAGLSVEAL EKMMEDPTVQ KMVYPHLPEE MRNPETFKWM		305 315 325 335 AGGSVEVLEK MIEDPTFQKM LYPHLPEEMR NPETFKWVRK	
345 355 365 375 LKNPQYRQQL QDMLNNMSGS GEWDKRMTET LKNFDLNSPE		345 355 365 375 NPQYRQQLQD MLNNMSESGE WDKRMTETLK NFDPNSPEVK	
385 395 405 415 VKQQFDQIGL TPEEVISKIM ENPDVAMAFQ NPRVQAALME		385 395 405 415 QGFDQLGLTP EEVISKIMEN PDVSMAFQNP RVEAALMDCS	
425 435 445 455 CSENPMNIMK YQNDKEVMDV FNKISQLFPG MTG		425 435 445 455 ENPMNIMKYQ NDKEVMDVFN KISQLFPGLT G	

TABLE 32: Amino acid sequence of Get3-MBP

The amino acid sequence, molecular weight, and extinction coefficient at 280 nm of the co-expressed Get3-MBP are listed below, with the sequence of Get3 is colored in blue.

Get3-MBP	
Number of amino acids:	731
Molecular weight:	80,698 g·mol^{-1}
Extinction coefficient:	86,290 M^{-1}·cm^{-1}
MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDE ALKDAQTNSS SGENQGSMDL TVEPNLHSLI TFTTHKWIFV GGKGGVGKTT SSCSIAIQMA LSQPNKQFLL ISTEPAHNLS DAFGEKFGKD ARKVTGMNNL SCMEIDPSAA LKDMNDMAVS RANNNGSDGQ GDDLGSLLQG GALADLTGSI PGIDEALSFM EVMKHIKRQE QGEGETFDTV IFDTAPTGHT LRFLQLPNTL SKLLEKFGEI TNKLGPMLNS FMGAGNVDIS GKLNELKANV ETIRQQFTDP DLTTFVCVCI SEFLSLYETE RLIQELISYD MDVNSIIVNQ LLFAENDQEH NCKRCQARWK MQKKYLDQID ELYEDFHVVK MPLCAGEIRG LNNLTKFSQF LNKEYNPITD GKVIYELEDK E	

TABLE 33: Amino acid sequences of the ADAM17-TM constructs

The amino acid sequence, molecular weight, and extinction coefficient at 280 nm of the used ADAM17-TM constructs are listed below. ADAM17-TM (blue) was combined with a His$_{10}$-tag (green), a Z-tag or GB1, and a sequence for TEV or thrombin cleavage (red; the cleavage site is in between the two underlined amino acids). In addition, the molecular weight and extinction coefficient for the TEV- or thrombin-cleaved products of ADAM17-TM are shown. All constructs were co-expressed with Get3-MBP.

ADAM17-TM-TEV-Z-tag				**(Get3-MBP co-expression)**
Number of amino acids:	184	/	44	TEV cleaved
Molecular weight:	21,022 g·mol^{-1}	/	4,960 g·mol^{-1}	
Extinction coefficient:	11,460 M^{-1}·cm^{-1}	/	6,990 M^{-1}·cm^{-1}	

MGHHHHHHHH HHGSNKFNKE QQNAFYEILH LPNLNEEQRN AFIQSLKDDP SQSANLLAEA KKLNDAQAPK
VAMNKFNKEQ QNAFYEILHL PNLNEEQRNA FIQSLKDDPS QSANLLAEAK KLNDAQAPKV AMSGENLYFQ
GNTFGKFLAD NIVGSVLVFS LIFWIPFSIL VHCVDKKLDK QYEG

ADAM17-TM-Z-tag				**(Get3-MBP co-expression)**
Number of amino acids:	183	/	45	Thrombin cleaved
Molecular weight:	20,779 g·mol^{-1}	/	5,047 g·mol^{-1}	
Extinction coefficient:	9,970 M^{-1}·cm^{-1}	/	6,990 M^{-1}·cm^{-1}	

MGHHHHHHHH HHGSNKFNKE QQNAFYEILH LPNLNEEQRN AFIQSLKDDP SQSANLLAEA KKLNDAQAPK
VAMNKFNKEQ QNAFYEILHL PNLNEEQRNA FIQSLKDDPS QSANLLAEAK KLNDAQAPKV AMSGLVPRGS
NTFGKFLADN IVGSVLVFSL IFWIPFSILV HCVDKKLDKQ YEG

ADAM17-TM-GB1				**(Get3-MBP co-expression)**
Number of amino acids:	133	/	49	Thrombin cleaved
Molecular weight:	14,856 g·mol^{-1}	/	5,335 g·mol^{-1}	
Extinction coefficient:	18,450 M^{-1}·cm^{-1}	/	6,990 M^{-1}·cm^{-1}	

MGHHHHHHHH HHDYDIPTTA MEYKLILNGK TLKGETTTEA VDAATAEKVF KQYANDNGVD GEWTYDDATK
TFTVTEIPTT LVPRGSGGSS NTFGKFLADN IVGSVLVFSL IFWIPFSILV HCVDKKLDKQ YEG

TABLE 34: Amino acid sequences of the iRhom2-TM1 constructs

The amino acid sequence, molecular weight, and extinction coefficient at 280 nm of the used iRhom2-TM1 constructs are listed below. iRhom2-TM1 (blue) was combined with a His$_{10}$-tag (green), a Z-tag or GB1, and a sequence for thrombin cleavage (red; the cleavage site is in between the two underlined amino acids). In addition, the molecular weight and extinction coefficient for the thrombin-cleaved products of iRhom2-TM1 are shown. All constructs were co-expressed with Get3-MBP.

iRhom2-TM1-Z-tag				**(Get3-MBP co-expression)**
Number of amino acids:	200	/	62	Thrombin cleaved
Molecular weight:	22,779 g·mol^{-1}	/	7,046 g·mol^{-1}	
Extinction coefficient:	14,440 M^{-1}·cm^{-1}	/	11,460 M^{-1}·cm^{-1}	

MGHHHHHHHH HHGSNKFNKE QQNAFYEILH LPNLNEEQRN AFIQSLKDDP SQSANLLAEA KKLNDAQAPK
VAMNKFNKEQ QNAFYEILHL PNLNEEQRNA FIQSLKDDPS QSANLLAEAK KLNDAQAPKV AMSGLVPRGS
GSQRQLESFD SHRPYFTYWL TFVHVIITLL VICTYGIAPV GFAQHVTTQL VLRNKGVYES

iRhom2-TM1-GB1				**(Get3-MBP co-expression)**
Number of amino acids:	148	/	64	Thrombin cleaved
Molecular weight:	16,711 g·mol^{-1}	/	7,190 g·mol^{-1}	
Extinction coefficient:	22,920 M^{-1}·cm^{-1}	/	11,460 M^{-1}·cm^{-1}	

MGHHHHHHHH HHDYDIPTTA MEYKLILNGK TLKGETTTEA VDAATAEKVF KQYANDNGVD GEWTYDDATK
TFTVTEIPTT LVPRGSGGSS QRQLESFDSH RPYFTYWLTF VHVIITLLVI CTYGIAPVGF AQHVTTQLVL
RNKGVYES

TABLE 35: Amino acid sequences of the iRhom2-7TM constructs

The amino acid sequence, molecular weight, and extinction coefficient at 280 nm of the used iRhom2-7TM constructs are listed below. iRhom2-7TM (blue) was combined with a His$_{10}$-tag (green), a sequence for thrombin cleavage (red; the cleavage site is in between the two underlined amino acids), and either with integrin (orange) or a Z-tag. In addition, the molecular weight and extinction coefficient for the thrombin-cleaved products of iRhom2-7TM are shown. The construct with the Z-tag was co-expressed with Get3-MBP.

iRhom2-7TM-Integrin

Number of amino acids:	548	/	247	Thrombin cleaved
Molecular weight:	60,036 g·mol^{-1}	/	28,342 g·mol^{-1}	
Extinction coefficient:	92,250 M^{-1}·cm^{-1}	/	62,910 M^{-1}·cm^{-1}	

```
MEGSGILSAT QEQIAESYYP EYLINLVQGQ LQTRQASSIY DDSYLGYSVA VGEFSGDDTE DFVAGVPKGN
LTYGYVTILN GSDIRSLYNF SGEQMASYFG YAVAATDVNG DGLDDLLVGA PLLMDRTPDG RPQEVGRVYV
YLQHPAGIEP TPTLTLTGHD EFGRFGSSLT PLGDLDQDGY NDVAIGAPFG GETQQGVVFV FPGGPGGLGS
KPSQVLQPLW AASHTPDFFG SALRGGRDLD GNGYPDLIVG SFGVDKAVVY RGRPIVSASA SLTIFPAMFN
PEERSCSDPS GSGSGSGLVP RGSSFDSHRP YFTYWLTFVH VIITLLVICT YGIAPVGFAQ HVTTQLVLRN
DQFYRLWLSL FLHAGVVHCL VSVVFQMTIL RDLEKLAGWH RIAIIFILSG ITGNLASAIF LPYRAEVGPA
GSQFGLLACL FVELFQSWPL LERPWKAFLN LSAIVLFLFI CGLLPWIDNI AHIFGFLSGL LLAFAFLPYI
TFGTSDKYRK RALILVSLLA FAGLFAALVL WLYIYPINWP WIEHLTLEHH HHHHHHHH
```

iRhom2-7TM-Z-tag **(Get3-MBP co-expression)**

Number of amino acids:	377	/	239	Thrombin cleaved
Molecular weight:	42,847 g·mol^{-1}	/	27,114 g·mol^{-1}	
Extinction coefficient:	65,890 M^{-1}·cm^{-1}	/	62,910 M^{-1}·cm^{-1}	

```
MGHHHHHHHH HHGSNKFNKE QQNAFYEILH LPNLNEEQRN AFIQSLKDDP SQSANLLAEA KKLNDAQAPK
VAMNKFNKEQ QNAFYEILHL PNLNEEQRNA FIQSLKDDPS QSANLLAEAK KLNDAQAPKV AMSGLVPRGS
GSSFDSHRPY FTYWLTFVHV IITLLVICTY GIAPVGFAQH VTTQLVLRND QFYRLWLSLF LHAGVVHCLV
SVVFQMTILR DLEKLAGWHR IAIIFILSGI TGNLASAIFL PYRAEVGPAG SQFGLLACLF VELFQSWPLL
ERPWKAFLNL SAIVLFLFIC GLLPWIDNIA HIFGFLSGLL LAFAFLPYIT FGTSDKYRKR ALILVSLLAF
AGLFAALVLW LYIYPINWPW IEHLTLE
```

TABLE 36: NMR chemical shift information of assigned resonances of BnaA10-Tic40

The NMR chemical shift information in ppm of all assigned resonances of BnaA10-Tic40 is shown below. The NMR chemical shifts were obtained with 1.1 mM ^{13}C,^{15}N-labelled protein in 20 mM NaPi pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM TCEP, 7 vol% D$_2$O, and 10 vol% TFE-d_3 at 308 K.

G306	
N	111.0
H	8.338
C	170.9
C_α	44.20
$H_{\alpha,1+2}$	4.020
A307	
N	123.7
H	8.309
C	175.2
C_α	51.53
H_α	4.377
C_β	18.51
M_β	1.399
G308	
N	107.6
H	8.240
C	171.0
C_α	44.08
$H_{\alpha,1+2}$	3.929
L309	
N	121.8
H	8.035
C	173.7
C_α	53.96
H_α	4.366
C_β	42.35
$H_{\beta,1+2}$	1.533
C_γ	26.03
S310	
N	117.4
H	7.979
C	172.5
C_α	56.07
H_α	4.675
C_β	64.20
$H_{\beta,1}$	3.998
$H_{\beta,2}$	4.283
V311	
N	121.5
H	8.572
C	174.4
C_α	65.20
H_α	3.689
C_β	30.69
H_β	2.027
$C_{\gamma1}$	20.88
$M_{\gamma1}$	0.953
$C_{\gamma2}$	21.25
$M_{\gamma2}$	1.027
E312	
N	119.5
H	8.581
C	176.1
C_α	58.98
H_α	4.099
C_β	27.95
$H_{\beta,1+2}$	2.010
C_γ	35.96
$H_{\gamma,1+2}$	2.315
A313	
N	122.2
H	7.853
C	177.3
C_α	53.79
H_α	4.128
C_β	17.59
M_β	1.503
L314	
N	119.4
H	8.028
C	175.8
C_α	57.07
H_α	4.060
C_β	40.98
$H_{\beta,1+2}$	1.748
E315	
N	117.7
H	8.427
C	176.0
C_α	59.20
H_α	3.789
C_β	28.33
$H_{\beta,1+2}$	2.112
C_γ	35.48
$H_{\gamma,1+2}$	2.413
K316	
N	117.0
H	7.499
C	177.1
C_α	57.49
H_α	4.117
C_β	31.00
$H_{\beta,1+2}$	1.951
M317	
N	119.5
H	8.138
C	174.9
C_α	57.79
H_α	4.176
C_β	32.27
$H_{\beta,1+2}$	2.181
C_γ	31.19
$H_{\gamma,1+2}$	2.640
M318	
N	113.5
H	7.745
C	174.0
C_α	54.90
H_α	4.271
C_β	30.16
$H_{\beta,1+2}$	2.086
C_γ	31.57
$H_{\gamma,1+2}$	2.544
E319	
N	116.5
H	7.458
C	173.7
C_α	55.77
H_α	4.159
C_β	29.05
$H_{\beta,1+2}$	2.128
C_γ	35.34
$H_{\gamma,1+2}$	2.570
D320	
N	122.1
H	7.701
C	172.8
C_α	50.60
H_α	4.913
C_β	41.81
$H_{\beta,1}$	2.685
$H_{\beta,2}$	3.179
P321	
C	176.1
C_α	64.05
H_α	4.153
C_β	31.34
$H_{\beta,1+2}$	1.977
C_γ	26.40
$H_{\gamma,1+2}$	2.381
C_δ	50.37
T322	
N	115.0
H	8.503
C	173.7
C_α	65.42
H_α	3.906
C_β	67.19
H_β	4.239
C_γ	21.09
M_γ	1.205
V323	
N	123.2
H	7.578
C	175.9
C_α	65.21
H_α	3.586
C_β	30.53
H_β	2.399
$C_{\gamma1}$	20.71
$M_{\gamma1}$	0.814
$C_{\gamma2}$	21.57
$M_{\gamma2}$	0.968
Q324	
N	117.9
H	7.959
C	175.0
C_α	58.77
H_α	3.346
C_β	26.77
$H_{\beta,1}$	1.935
$H_{\beta,2}$	1.540
C_γ	32.49
$H_{\gamma,1+2}$	2.380
C_δ	176.8
N_ε	108.8
$H_{\varepsilon,1}$	7.502
$H_{\varepsilon,2}$	6.625
K325	
N	114.9
H	7.456
C	175.2
C_α	58.01
H_α	3.858
C_β	31.52
$H_{\beta,1+2}$	1.890
M326	
N	115.7
H	7.467
C	174.8
C_α	56.42
H_α	4.326
C_β	32.25
$H_{\beta,1+2}$	2.198
C_γ	31.11
$H_{\gamma,1+2}$	2.611
V327	
N	117.1
H	7.920
C	174.2
C_α	64.10
H_α	3.984
C_β	32.33
H_β	1.934
$C_{\gamma1}$	20.42
$M_{\gamma1}$	0.929
$C_{\gamma2}$	21.33
$M_{\gamma2}$	0.961
Y328	
N	119.1
H	8.383
C	171.9
C_α	60.10
H_α	4.383
C_β	34.39
$H_{\beta,1}$	3.330
$H_{\beta,2}$	2.937
P329	
C	174.2
C_α	64.61
H_α	4.215
C_β	30.37
C_γ	31.65
C_δ	55.52
H330	
N	112.2
H	7.941
C	171.4
C_α	54.60
H_α	4.517
C_β	28.52
$H_{\beta,1}$	3.456
$H_{\beta,2}$	3.171
L331	
N	120.4
H	7.888
C	171.8
C_α	51.08
H_α	4.766
C_β	41.52
$H_{\beta,1}$	1.383
$H_{\beta,2}$	1.691
P332	
C	175.0
C_α	61.84
H_α	4.438
C_β	31.43
$H_{\beta,1+2}$	1.836
C_γ	26.91
$H_{\gamma,1+2}$	2.464
E333	
N	123.0
H	8.859
C	175.9
C_α	59.88
H_α	3.496
C_β	28.83
$H_{\beta,1+2}$	1.990
C_γ	35.26
$H_{\gamma,1+2}$	2.283
E334	
N	117.3
H	9.731
C	174.3
C_α	58.06
H_α	4.095
C_β	27.68
$H_{\beta,1+2}$	1.987
C_γ	35.33
$H_{\gamma,1+2}$	2.311
M335	
N	115.4
H	7.854
C	172.4
C_α	54.45
H_α	4.427
C_β	32.91
$H_{\beta,1}$	2.422
$H_{\beta,2}$	1.899
R336	
N	118.3
H	7.063
C	172.0
C_α	54.70
H_α	3.650
C_β	27.37
$H_{\beta,1}$	1.355
$H_{\beta,2}$	1.622
C_γ	25.86
$H_{\gamma,1+2}$	1.139
C_δ	43.31
$H_{\delta,1+2}$	2.923
N337	
N	119.0
H	7.893
C	170.8
C_α	50.38
H_α	4.961
C_β	38.80
$H_{\beta,1+2}$	2.862
C_γ	174.2
N_δ	114.7
$H_{\delta,1}$	7.587
$H_{\delta,2}$	6.884
P338	
C	176.0
C_α	64.53
H_α	4.428
C_β	31.33
$H_{\beta,1+2}$	2.075
C_γ	27.37
C_δ	49.98
E339	
N	117.0
H	8.380
C	176.5
C_α	58.55
H_α	4.140
C_β	28.04
$H_{\beta,1+2}$	2.096
C_γ	35.50
$H_{\gamma,1+2}$	2.365
T340	
N	116.6
H	7.973
C	173.5
C_α	64.89
H_α	4.048
C_β	67.41
H_β	4.173
C_γ	21.91
M_γ	1.256
F341	
N	121.9
H	8.132
C	174.1
C_α	60.81
H_α	4.231
C_β	38.18
$H_{\beta,1}$	3.140
$H_{\beta,2}$	3.428
K342	
N	116.1
H	7.943
C	176.0
C_α	59.06
C_β	31.31
$H_{\beta,1+2}$	1.931
W343	
N	118.5
H	7.598
C	176.6
C_α	59.43
H_α	4.332
C_β	27.80
$H_{\beta,1}$	3.528
$H_{\beta,2}$	3.391
M344	
N	120.3
H	8.366
C	175.0
C_α	58.29
H_α	3.745
C_β	31.55
$H_{\beta,1}$	2.046
$H_{\beta,2}$	1.703
L345	
N	114.6
H	7.539
C	174.9
C_α	55.63
H_α	3.828
C_β	40.63
$H_{\beta,u}$	1.406
$H_{\beta,d}$	1.721
C_γ	25.81
H_γ	1.562
$C_{\delta,u}$	22.54
$M_{\delta,u}$	0.674
$C_{\delta,d}$	24.37
$M_{\delta,d}$	0.700
K346	
N	115.4
H	7.241
C	173.4
C_α	55.74
H_α	4.141
C_β	31.62
$H_{\beta,1+2}$	1.935
N347	
N	120.6
H	7.611
C	171.3
C_α	49.57
H_α	4.982
C_β	39.27
$H_{\beta,1}$	3.061
$H_{\beta,2}$	2.515
C_γ	174.0
N_δ	113.0
$H_{\delta,1}$	7.125
$H_{\delta,2}$	6.883
P348	
C	175.8
C_α	64.57
H_α	4.140
C_β	31.29
$H_{\beta,1+2}$	2.034
C_γ	26.49
$H_{\gamma,1+2}$	2.397
C_δ	50.11
Q349	
N	116.9
H	8.109
C	175.7
C_α	58.02
H_α	4.114
C_β	27.21
$H_{\beta,1+2}$	1.938
C_γ	33.04
$H_{\gamma,1+2}$	2.249
C_δ	177.4
N_ε	111.2
$H_{\varepsilon,1}$	7.334
$H_{\varepsilon,2}$	6.740
Y350	
N	119.1
H	7.590
C	174.9
C_α	59.06
H_α	4.437
C_β	36.95
$H_{\beta,1+2}$	3.088
R351	
N	119.9
H	8.305
C	175.5
C_α	59.68
H_α	3.808
C_β	29.28
$H_{\beta,1+2}$	1.785
C_γ	27.23
$H_{\gamma,1+2}$	1.537
C_δ	42.55
$H_{\delta,1+2}$	3.242
Q352	
N	117.2
H	7.835
C	175.2
C_α	57.77
H_α	3.998
C_β	27.38
$H_{\beta,1+2}$	2.163
C_γ	32.81
$H_{\gamma,1+2}$	2.419
C_δ	177.6
N_ε	113.6
$H_{\varepsilon,1}$	7.652
$H_{\varepsilon,2}$	6.734
Q353	
N	118.6
H	7.707
C	176.5
C_α	58.20
H_α	4.145
C_β	27.41
$H_{\beta,1+2}$	2.179
C_γ	33.27
$H_{\gamma,1+2}$	2.495
C_δ	177.0
N_ε	109.5
$H_{\varepsilon,1}$	7.173
$H_{\varepsilon,2}$	6.657

L354	
N	120.7
H	8.307
C	175.7
C_α	56.94
H_α	4.136
C_β	40.77
$H_{\beta,1+2}$	1.663
Q355	
N	118.9
H	8.310
C	175.5
C_α	58.36
H_α	3.857
C_β	27.67
$H_{\beta,1+2}$	2.209
C_γ	33.55
$H_{\gamma,1+2}$	2.378
C_δ	177.0
N_ε	112.3
$H_{\varepsilon,1}$	7.251
$H_{\varepsilon,2}$	6.960
D356	
N	119.6
H	8.349
C	176.0
C_α	56.66
H_α	4.433
C_β	39.34
$H_{\beta,1}$	2.833
$H_{\beta,2}$	2.683
M357	
N	120.2
H	7.981
C	176.2
C_α	58.37
H_α	4.102
C_β	31.78
$H_{\beta,1+2}$	2.250
C_γ	30.85
$H_{\gamma,u}$	2.427
$H_{\gamma,d}$	2.658
L358	
N	119.6
H	8.383
C	177.0
C_α	56.83
H_α	4.043
C_β	40.24
$H_{\beta,1}$	1.465
$H_{\beta,2}$	1.867
N359	
N	118.0
H	8.386
C	174.7
C_α	54.45
H_α	4.505
C_β	37.42
$H_{\beta,1+2}$	2.916
C_γ	173.6
N_δ	111.5
$H_{\delta,1}$	7.736
$H_{\delta,2}$	6.892
N360	
N	118.1
H	7.987
C	173.9
C_α	53.98
H_α	4.679
C_β	37.98
$H_{\beta,1+2}$	2.862
C_γ	174.1
N_δ	112.7
$H_{\delta,1}$	7.489
$H_{\delta,2}$	6.875
M361	
N	119.3
H	8.016
C	174.3
C_α	56.37
H_α	4.313
C_β	32.12
$H_{\beta,1+2}$	2.064
C_γ	31.01
$H_{\gamma,1+2}$	2.556
S362	
N	114.7
H	8.034
C	172.8
C_α	58.54
H_α	4.344
C_β	62.83
$H_{\beta,1+2}$	3.915
G363	
N	110.1
H	8.185
C	172.0
C_α	44.78
$H_{\alpha,1+2}$	4.037
S364	
N	115.2
H	8.129
C	172.7
C_α	57.76
H_α	4.454
C_β	63.07
$H_{\beta,1+2}$	3.964
G365	
N	110.7
H	8.404
C	172.0
C_α	44.83
$H_{\alpha,1+2}$	3.992
E366	
N	120.2
H	8.230
C	174.5
C_α	56.50
H_α	4.239
C_β	28.89
$H_{\beta,1+2}$	1.992
C_γ	35.23
$H_{\gamma,1+2}$	2.233
W367	
N	120.8
H	8.072
C	174.0
C_α	57.85
H_α	4.444
C_β	28.39
$H_{\beta,1+2}$	3.276
D368	
N	119.9
H	8.116
C	175.0
C_α	54.74
H_α	4.368
C_β	39.86
$H_{\beta,1+2}$	2.611
K369	
N	121.6
H	7.987
C	175.4
C_β	31.38
R370	
N	119.2
H	8.115
C	176.1
C_α	57.14
H_α	4.162
C_β	28.97
$H_{\beta,1+2}$	1.836
C_γ	26.17
$H_{\gamma,1+2}$	1.673
C_δ	42.28
$H_{\delta,1+2}$	3.123
M371	
N	119.7
H	8.292
C	174.9
C_α	56.51
H_α	4.256
C_β	31.07
$H_{\beta,1+2}$	1.949
$H_{\gamma,1+2}$	2.328
T372	
N	113.8
H	7.973
C	173.3
C_α	64.39
H_α	4.011
C_β	68.06
H_β	4.288
C_γ	20.85
M_γ	1.244
E373	
N	120.8
H	8.044
C	175.3
C_α	57.54
H_α	4.126
C_β	28.69
$H_{\beta,1+2}$	2.087
C_γ	35.30
$H_{\gamma,1+2}$	2.334
T374	
N	114.8
H	7.896
C	173.0
C_α	64.19
H_α	4.089
C_β	68.24
H_β	4.256
C_γ	20.68
M_γ	1.213
L375	
N	121.4
H	7.976
C	175.5
C_α	55.76
H_α	4.199
C_β	40.84
$H_{\beta,u}$	1.580
$H_{\beta,d}$	1.759
K376	
N	118.9
H	7.842
C	174.4
C_α	56.99
H_α	4.157
C_β	31.55
$H_{\beta,1+2}$	1.835
N377	
N	116.7
H	7.864
C	172.3
C_α	52.92
H_α	4.650
C_β	38.11
$H_{\beta,1+2}$	2.700
C_γ	174.1
N_δ	112.4
$H_{\delta,1}$	7.444
$H_{\delta,2}$	6.861
F378	
N	120.0
H	7.876
C	172.2
C_α	57.19
H_α	4.549
C_β	38.57
$H_{\beta,1+2}$	3.089
D379	
N	121.6
H	7.982
C	173.7
C_α	52.92
H_α	4.617
C_β	40.25
$H_{\beta,1}$	2.818
$H_{\beta,2}$	2.627
L380	
N	122.9
H	8.104
C	173.9
C_α	55.14
H_α	4.135
C_β	40.92
$H_{\beta,1+2}$	1.637
C_γ	26.13
$C_{\delta u}$	22.47
$M_{\delta u}$	0.821
$C_{\delta d}$	24.02
$M_{\delta d}$	0.889
N381	
N	115.2
H	8.215
C	172.6
C_α	52.25
H_α	4.720
C_β	38.35
$H_{\beta,1+2}$	2.792
C_γ	174.7
N_δ	113.3
$H_{\delta,1}$	7.665
$H_{\delta,2}$	6.822
S382	
N	116.7
H	7.789
C	171.0
C_α	55.46
H_α	4.731
C_β	62.58
$H_{\beta,1}$	4.083
$H_{\beta,2}$	3.973
P383	
C	175.6
C_α	64.01
H_α	4.319
C_β	30.91
$H_{\beta,1+2}$	1.970
C_γ	26.52
$H_{\gamma,1+2}$	2.359
C_δ	49.82
E384	
N	118.0
H	8.721
C	175.4
C_α	58.08
H_α	4.115
C_β	28.21
$H_{\beta,1+2}$	1.999
C_γ	35.58
$H_{\gamma,1+2}$	2.297
V385	
N	120.6
H	7.533
C	174.6
C_α	63.87
H_α	3.807
C_β	31.01
H_β	2.105
$C_{\gamma u}$	20.65
$M_{\gamma u}$	0.860
$C_{\gamma d}$	21.14
$M_{\gamma d}$	0.962
K386	
N	121.1
H	8.052
C	175.5
C_α	58.13
C_β	31.48
Q387	
N	117.7
H	8.054
C	174.9
C_α	57.01
H_α	4.186
C_β	27.79
$H_{\beta,1+2}$	2.100
C_γ	32.93
$H_{\gamma,1+2}$	2.418
C_δ	177.5
N_ε	110.9
$H_{\varepsilon,1}$	7.403
$H_{\varepsilon,2}$	6.763
Q388	
N	119.0
H	7.770
C	175.0
C_α	56.90
H_α	4.135
C_β	27.66
$H_{\beta,1+2}$	2.027
C_γ	32.79
$H_{\gamma,1+2}$	2.284
C_δ	177.0
N_ε	110.4
$H_{\varepsilon,1}$	7.191
$H_{\varepsilon,2}$	6.670
F389	
N	118.7
H	8.050
C	174.8
C_α	58.95
H_α	4.438
C_β	37.38
$H_{\beta,1}$	3.059
$H_{\beta,2}$	3.319
D390	
N	120.4
H	8.411
C	175.4
C_α	55.48
H_α	4.528
C_β	39.79
$H_{\beta,1+2}$	2.761
Q391	
N	117.8
H	7.907
C	174.3
C_α	56.45
H_α	4.213
C_β	28.02
$H_{\beta,1+2}$	2.167
C_γ	33.01
$H_{\gamma,1+2}$	2.453
C_δ	177.5
N_ε	111.3
$H_{\varepsilon,1}$	7.426
$H_{\varepsilon,2}$	6.761
I392	
N	115.8
H	7.665
C	173.8
C_α	60.88
H_α	4.267
C_β	37.69
H_β	2.032
$C_{\gamma 1}$	26.40
$H_{\gamma 1,u}$	1.290
$H_{\gamma 1,d}$	1.518
$C_{\gamma 2}$	16.55
$M_{\gamma 2}$	0.913
C_δ	12.54
M_δ	0.780
G393	
N	109.5
H	8.042
C	171.3
C_α	44.95
$H_{\alpha,1+2}$	3.947
L394	
N	120.0
H	7.557
C	173.8
C_α	53.02
H_α	4.710
C_β	44.09
$H_{\beta,1+2}$	1.454
T395	
N	112.7
H	8.232
C	171.0
C_α	58.42
H_α	4.729
C_β	68.20
H_β	4.619
C_γ	21.08
M_γ	1.338
P396	
C	175.2
C_α	64.76
H_α	3.973
C_β	30.53
$H_{\beta,1+2}$	2.137
C_γ	27.23
E397	
N	115.0
H	8.811
C	176.6
C_α	59.55
H_α	3.845
C_β	27.81
$H_{\beta,1+2}$	1.954
C_γ	35.82
$H_{\gamma,1+2}$	2.220
E398	
N	120.1
H	7.723
C	176.2
C_α	58.10
H_α	4.033
C_β	29.44
$H_{\beta,u}$	2.024
$H_{\beta,d}$	2.323
C_γ	36.08
$H_{\gamma,1+2}$	2.317
V399	
N	120.4
H	7.813
C	175.0
C_α	66.00
H_α	3.519
C_β	30.61
H_β	2.079
$C_{\gamma 1}$	19.73
$M_{\gamma 1}$	0.698
$C_{\gamma 2}$	0.736
$M_{\gamma 2}$	22.13
I400	
N	118.4
H	8.305
C	175.1
C_α	63.50
H_α	3.547
C_β	36.23
H_β	1.829
$C_{\gamma 1}$	28.04
$H_{\gamma 1,1}$	1.154
$H_{\gamma 1,2}$	1.559
$C_{\gamma 2}$	16.55
$M_{\gamma 2}$	0.805
C_δ	10.88
M_δ	0.737
S401	
N	113.7
H	7.806
C	173.8
C_α	61.43
H_α	4.040
C_β	61.87
$H_{\beta,1+2}$	3.955
K402	
N	120.0
H	7.600
C	177.0
C_α	57.46
H_α	4.135
C_β	30.92
$H_{\beta,1+2}$	1.857
I403	
N	120.0
H	8.137
C	175.0
C_α	64.34
H_α	3.634
C_β	36.72
H_β	1.906
$C_{\gamma 1}$	27.70
$H_{\gamma 1,1}$	1.017
$H_{\gamma 1,2}$	1.673
$C_{\gamma 2}$	16.12
$M_{\gamma 2}$	0.692
C_δ	12.41
M_δ	0.580
M404	
N	114.8
H	7.942
C	174.4
C_α	55.77
H_α	4.134
C_β	30.60
$H_{\beta,1+2}$	2.068
C_γ	31.80
$H_{\gamma,1+2}$	2.527
E405	
N	116.2
H	7.384
C	173.7
C_α	56.23
H_α	4.156
C_β	29.06
$H_{\beta,1+2}$	2.128
C_γ	35.30
$H_{\gamma,1}$	2.333
$H_{\gamma,2}$	2.562
N406	
N	119.1
H	7.722
C	170.5
C_α	50.00
H_α	5.138
C_β	39.21
$H_{\beta,1+2}$	2.853
C_γ	174.7
N_δ	114.2
$H_{\delta,1}$	7.739
$H_{\delta,2}$	7.288
P407	
C	175.2
C_α	64.31
H_α	4.316
C_β	31.13
$H_{\beta,1+2}$	2.375
C_γ	26.43
$H_{\gamma,1+2}$	2.053

C_δ	49.96
$H_{\delta,1+2}$	3.916
D408	
N	116.4
H	8.296
C	175.9
C_α	56.05
H_α	4.430
C_β	38.96
$H_{\beta,1+2}$	2.683
V409	
N	120.7
H	7.761
C	174.4
C_α	64.32
H_α	3.753
C_β	30.84
H_β	2.262
$C_{\gamma1}$	21.29
$M_{\gamma1}$	0.861
$C_{\gamma2}$	21.40
$M_{\gamma2}$	0.952
A410	
N	121.7
H	8.175
C	177.8
C_α	54.77
H_α	3.948
C_β	17.30
M_β	1.489
M411	
N	114.4
H	7.975
C	176.4
C_α	57.10
H_α	4.229
C_β	31.23
$H_{\beta,1+2}$	2.169
C_γ	31.25
$H_{\gamma,1+2}$	2.673
A412	
N	123.6
H	7.695
C	176.5
C_α	53.94
H_α	4.223
C_β	17.52
M_β	1.502
F413	
N	114.6
H	8.044
C	172.9
C_α	58.74
H_α	4.139
C_β	38.11
$H_{\beta,1+2}$	3.133
Q414	
N	114.1
H	7.448
C	173.1
C_α	55.15
H_α	4.140
C_β	27.86
$H_{\beta,1}$	2.083
$H_{\beta,2}$	2.315
C_γ	33.11
$H_{\gamma,1+2}$	2.538
C_δ	177.8
N_ε	111.2
$H_{\varepsilon,1}$	7.311
$H_{\varepsilon,2}$	6.736
N415	
N	120.3
H	7.574
C	172.1
C_α	49.81
H_α	4.919
C_β	38.37
$H_{\beta,1}$	2.959
$H_{\beta,2}$	3.205
C_γ	174.4
N_δ	113.1
$H_{\delta,1}$	7.887
$H_{\delta,2}$	7.046
P416	
C	176.0
C_α	64.30
H_α	4.208
C_β	31.19
$H_{\beta,1+2}$	2.437
C_γ	26.57
$H_{\gamma,1+2}$	2.063
C_δ	50.11
$H_{\delta,1+2}$	3.902
R417	
N	118.0
H	8.094
C	176.3
C_α	58.10
H_α	4.120
C_β	29.55
$H_{\beta,1+2}$	1.826
C_γ	27.07
$H_{\gamma,1+2}$	1.572
C_δ	42.02
$H_{\delta,u}$	3.238
$H_{\delta,d}$	3.368
V418	
N	119.9
H	7.216
C	174.7
C_α	64.84
H_α	3.401
C_β	30.60
H_β	2.355
$C_{\gamma1}$	20.72
$M_{\gamma1}$	0.719
$C_{\gamma2}$	21.64
$M_{\gamma2}$	0.868
Q419	
N	118.0
H	7.872
C	175.3
C_α	58.80
H_α	3.648
C_β	27.16
$H_{\beta,1+2}$	2.022
C_γ	32.58
$H_{\gamma,1+2}$	2.338
C_δ	176.6
N_ε	109.6
$H_{\varepsilon,1}$	7.334
$H_{\varepsilon,2}$	6.346
A420	
N	119.3
H	7.852
C	177.2
C_α	54.26
H_α	4.064
C_β	17.07
M_β	1.508
A421	
N	121.0
H	7.519
C	177.4
C_α	54.41
H_α	3.957
C_β	17.29
M_β	0.760
L422	
N	118.6
H	8.388
C	177.4
C_α	57.25
H_α	3.920
C_β	40.11
$H_{\beta,1}$	1.502
$H_{\beta,2}$	1.836
C_γ	25.98
M423	
N	118.9
H	8.145
C	176.9
C_α	57.91
H_α	4.156
C_β	31.27
E424	
N	119.9
H	8.361
C	178.0
C_α	58.90
H_α	4.235
C_β	28.78
$H_{\beta,1}$	2.223
$H_{\beta,2}$	2.418
C_γ	35.67
$H_{\gamma,1+2}$	2.626
C425	
N	118.6
H	8.502
C	172.7
C_α	61.88
H_α	4.464
C_β	25.87
$H_{\beta,1}$	2.706
$H_{\beta,2}$	3.009
S426	
N	114.5
H	7.855
C	172.7
C_α	60.19
H_α	4.204
C_β	61.99
$H_{\beta,1+2}$	3.947
E427	
N	120.0
H	7.311
C	174.4
C_α	57.23
H_α	4.201
C_β	29.02
$H_{\beta,1+2}$	2.134
C_γ	35.06
$H_{\gamma,1+2}$	2.425
N428	
N	112.9
H	7.393
C	170.9
C_α	50.02
H_α	4.968
C_β	38.61
$H_{\beta,1+2}$	2.955
C_γ	175.0
N_δ	111.1
$H_{\delta,1}$	7.601
$H_{\delta,2}$	6.803
P429	
C	175.6
C_α	64.07
H_α	4.591
C_β	31.21
$H_{\beta,1+2}$	2.434
C_γ	26.18
$H_{\gamma,1+2}$	2.039
C_δ	50.08
$H_{\delta,1+2}$	3.705
M430	
N	115.3
H	7.979
C	176.0
C_α	56.27
H_α	4.324
C_β	30.63
$H_{\beta,1+2}$	2.044
C_γ	31.69
$H_{\gamma,1+2}$	2.590
N431	
N	117.3
H	8.512
C	173.9
C_α	54.33
H_α	4.590
C_β	37.65
$H_{\beta,1+2}$	2.926
C_γ	173.0
N_δ	114.4
$H_{\delta,1}$	7.534
$H_{\delta,2}$	7.428
I432	
N	121.8
H	7.893
C	174.8
C_α	61.95
H_α	4.133
C_β	37.11
H_β	1.819
$C_{\gamma1}$	28.28
$H_{\gamma1,u}$	0.825
$H_{\gamma1,d}$	1.319
$C_{\gamma2}$	16.74
$M_{\gamma2}$	0.818
C_δ	12.70
M_δ	0.568
M433	
N	118.1
H	7.515
C	175.9
C_α	56.75
H_α	4.188
C_β	30.98
$H_{\beta,1+2}$	2.107
$H_{\gamma,1+2}$	2.654
K434	
N	117.4
H	7.616
C	174.9
C_α	3.939
H_α	57.21
C_β	30.98
$H_{\beta,1+2}$	1.463
Y435	
N	116.1
H	7.657
C	173.0
C_α	57.49
H_α	4.559
C_β	38.62
$H_{\beta,1}$	2.596
$H_{\beta,2}$	3.392
Q436	
N	116.8
H	7.038
C	173.2
C_α	57.75
H_α	4.352
C_β	28.12
$H_{\beta,1+2}$	2.104
C_γ	32.92
$H_{\gamma,1+2}$	2.407
C_δ	176.9
N_ε	110.7
$H_{\varepsilon,1}$	7.245
$H_{\varepsilon,2}$	6.768
N437	
N	114.6
H	8.318
C	171.1
C_α	51.68
H_α	4.898
C_β	37.84
$H_{\beta,1}$	2.671
$H_{\beta,2}$	2.954
C_γ	175.0
N_δ	112.6
$H_{\delta,1}$	7.496
$H_{\delta,2}$	6.853
D438	
N	123.1
H	7.984
C	172.1
C_α	51.83
H_α	4.767
C_β	40.53
$H_{\beta,1}$	2.472
$H_{\beta,2}$	3.520
K439	
N	126.1
H	8.404
C	175.1
C_α	58.99
H_α	3.871
C_β	31.56
$H_{\beta,1+2}$	1.853
E440	
N	116.7
H	8.156
C	176.0
C_α	58.60
H_α	4.103
C_β	28.89
$H_{\beta,1+2}$	1.966
C_γ	35.94
$H_{\gamma,1+2}$	2.272
V441	
N	117.5
H	7.462
C	175.2
C_α	65.54
H_α	3.441
C_β	31.10
H_β	1.956
$C_{\gamma1}$	20.41
$M_{\gamma1}$	0.395
$C_{\gamma2}$	22.09
$M_{\gamma2}$	0.769
M442	
N	117.2
H	8.579
C	176.7
C_α	56.00
H_α	4.498
C_β	29.52
$H_{\beta,1+2}$	2.003
C_γ	31.46
D443	
N	119.8
H	8.639
C	176.5
C_α	56.67
H_α	4.513
C_β	39.46
$H_{\beta,1}$	2.589
$H_{\beta,2}$	2.891
V444	
N	119.1
H	7.359
C	175.1
C_α	65.62
H_α	3.667
C_β	30.31
H_β	2.173
$C_{\gamma1}$	20.81
$M_{\gamma1}$	0.814
$C_{\gamma2}$	22.41
$M_{\gamma2}$	1.047
F445	
N	120.4
H	8.489
C	175.9
C_α	60.66
H_α	4.153
C_β	37.42
$H_{\beta,1+2}$	3.150
N446	
N	117.9
H	8.898
C	175.2
C_α	55.46
H_α	4.459
C_β	37.55
$H_{\beta,1}$	2.848
$H_{\beta,2}$	3.030
C_γ	173.5
N_δ	111.5
$H_{\delta,1}$	7.640
$H_{\delta,2}$	6.823
K447	
N	120.4
H	7.817
C	176.3
C_α	58.16
H_α	4.110
C_β	30.98
I448	
N	118.3
H	8.441
C	175.1
C_α	64.25
H_α	3.767
C_β	36.96
H_β	1.993
$C_{\gamma1}$	27.96
$H_{\gamma1,u}$	1.157
$H_{\gamma1,d}$	1.751
$C_{\gamma2}$	16.33
$M_{\gamma2}$	0.947
C_δ	12.50
M_δ	0.780
S449	
N	114.2
H	8.055
C	173.0
C_α	60.52
H_α	4.074
C_β	62.04
$H_{\beta,1+2}$	3.839
Q450	
N	118.2
H	7.430
C	174.5
C_α	56.31
H_α	4.153
C_β	27.85
$H_{\beta,1+2}$	2.136
C_γ	33.04
$H_{\gamma,1+2}$	2.462
C_δ	177.6
N_ε	110.8
$H_{\varepsilon,1}$	7.240
$H_{\varepsilon,2}$	6.728
L451	
N	118.7
H	7.755
C	173.9
C_α	55.13
H_α	4.060
C_β	41.63
$H_{\beta,1}$	0.985
$H_{\beta,2}$	1.612
F452	
N	116.8
H	7.615
C	170.6
C_α	54.22
H_α	4.913
C_β	38.19
$H_{\beta,1+2}$	3.058
P453	
C	175.2
C_α	63.09
H_α	4.483
C_β	30.79
$H_{\beta,1+2}$	2.303
C_γ	26.42
$H_{\gamma,1+2}$	1.979
C_δ	49.39
$H_{\delta,1+2}$	3.595
G454	
N	108.1
H	8.419
C	172.2
C_α	44.54
$H_{\alpha,1+2}$	3.992
M455	
N	119.5
H	8.121
C	173.8
C_α	55.52
H_α	4.479
C_β	32.00
$H_{\beta,1+2}$	2.143
C_γ	29.76
$H_{\gamma,1+2}$	2.538
T456	
N	111.4
H	7.935
C	171.4
C_α	60.48
H_α	4.421
C_β	69.04
H_β	4.357
C_γ	20.39
M_γ	1.180
G457	
N	116.5
H	7.823
C	176.2
C_α	45.26
$H_{\alpha,1+2}$	3.809