Structural analysis of UBA5 and its binding partners

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften

am Fachbereich Chemie der Fakultät für Mathematik,

Informatik und Naturwissenschaften

Universität Hamburg

vorgelegt von

Sebastian Fuchs

Die vorliegende Arbeit wurde von Januar 2016 bis September 2019 in der Arbeitsgruppe von Prof. Dr. Christian Betzel im Laboratorium für Strukturbiologie von Infektionen und Entzündungen am Institut für Biochemie und Molekularbiologie am Fachbereich Chemie der Universität Hamburg durchgeführt.

1. Gutachter: Prof. Dr. Christian Betzel
2. Gutachter: Prof. Dr. Henning Tidow

Datum der Druckfreigabe: 07.12.2021

Datum der Disputation: 28.01.2022

I. Publikationsliste

Structure and dynamics of UBA5-UFM1 complex formation showing new insights in the UBA5 activation mechanism, Journal of Structural Biology, 213 (4), 2021 [1]

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAS *atomic absorption spectroscopy*
ASC1 *activating signal conintegrator 1*
Atg *autophagy related genes*
ATP *adenosine triphosphate*
AUC *analytical ultracentrifugation*
CD *circular dichroism*
DLS *dynamic light scattering*
D_{max} *maximal distance distribution*
DNA *desoxyribonucleic acid*
D_t *translational diffusion coefficient*
e.g. *exempli gratia, for example*
EDTA *ethylene diamine tetra acetic acid*
EMBL *european microbiology laboratory*
ER *endoplasmic reticulum*
ERα estrogen receptor α
et al. *et alii; and others*
FADH flavine- adenine dinucleotide
FAT10 *HLAF adjacent transcript 10*
FCCH *first catalytic cysteine half domain*
FRET *Förster resonance energy transfer*
GFP *green fluorescent protein*
GPCR *G-protein coupled receptors*
GST *glutathion S- transferase*
HECT *homologous to the E6AP carboxyl terminus*
HIV *human immunodeficiency virus*
HPLC *high performance liquid chromatography*
HRV *human rhino virus*
IPTG isopropyl β-D-1-thiogalctopyranoside
ISG15 *interferon stimulated gene 15*
k_B Boltzmann's constant
K_D *dissociation constant*, *dissociation constant*
LB *Luria Bertani*
MAD *multiple wavelength anomalous dispersion*
MIR *multiple isomorphous replacement*
MNSFβ *monoclonal nonspecific suppressor factor ß*
MRE. *mean residue ellipticity)*
MS *mass spectrometry*
MST *microscale thermophoresis*
MW molecular weight
NADH nicotinamide-adenine dinucleotide
NCAM140 *neuronal cell adhesion molecule 140*
NEDD8 *neural precursor cell-expresed, developmentally downregulated gene 8*
NLS *nuclear locating signal*
NMR *nuclear magnetic resonance*
NPE *nitrophenyl ester*
OD *optical density*
PBS *phosphate buffer saline*
PCI *proteasome COP9-initiation factor*
PCR *polymerase chain reaction*
pdb *protein data base*
pH *potentia Hydrogenii*
ppm *parts per million*
R_h hydrodynmaic radius
RING *really interesting new gene*
RNA *ribonucleic acid*
RP reversed phase
RT *room temperature*
SAD *single wavelength anomalous dispersion*
SAXS *small angle X-ray scattering*
SCCH *second catalytic cysteine half domain*
SDS-PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC *size exclusion chromatography*
SIR *single isomorphous replacement*
S_T *Soret coefficient*
SUMO *small ubiquitin-like modifier*
T temperature (K)
TBAB *tetrabutyle ammonium bromide*
TR *time resolved*
UBA5 *ubiquitin-like modifier activating enzyme 5*
UBL *ubiquitin like protein*
UFBP1 *ubiquitin fold modifier binding protein 1*
UFC1 *ubiquitin fold modifier conjugating enzyme 1*
UFL1 *ubiquitin fold modifier ligating enzyme 1*
UFM1 *ubiquitin fold modifier 1*
UfSP *UFM1 specific protease*
UIS *UFM1 interacting sequence*
URM1 *ubiquitin related modifier 1*
UV *ultra violet*
ΔG *free enthalpy*
ΔH enthalpy
ΔS entropy
η solvent viscosity

List of units

AU	absorption units
A	Ampere
Å	Angstrom
d	day
Da	Dalton
°	degree
°C	degree Celsius
g	gram
h	hour
K	Kelvin
L	liter
m	meter
M	molar, mol/L
min	minute
s	second
S	Svedberg
V	Volt

The nomenclature for the abbreviation of amino acids is based on the IUBAC-IUB agreement.

Zielsetzung

Im Fokus der vorliegenden Arbeit lag die funktionelle und strukturelle Analyse der Aktivierungskaskade, welche zur Aktivierung des ubiquitinähnlichen Proteins Ubiquitin Fold Modifier 1 (UFM1) führt. Die an dieser Kaskade beteiligten Enzyme sind das Ubiquitin Fold Modifier Activating Enzyme 5 (UBA5), das Ubiquitin Fold Modifier Conjugating Enzyme 1 (UFC1) und das Ubiquitin Fold Modifier Ligating Enzyme 1 (UFL1). Im Genaueren sollte hier der erste Aktivierungsschritt strukturell untersucht werden, der in der Komplexbildung zwischen UFM1 und UBA5 besteht. Ein weiterer Aspekt der Untersuchung sollte die Frage beantworten, ob UFC1 und UFM1 ohne vorherige Aktivierung durch UBA5 einen Komplex bilden können.

Ein weiterer Fokus lag bei der Evaluierung, ob das ATP-bindende Protein UBA5 für sogenannte „Pump-and-Probe-Experimente“ geeignet ist, um für zeitaufgelöste kristallographische Untersuchungen verwendet zu werden und um den Enzymmechanismus bei hoher Orts- und Zeitauflösung zu analysieren. Im Detail soll dabei ein stabiler Komplex hergestellt und kristallisiert werden, der aus UBA5 und einer caged Verbindung besteht. *Caged-Verbindungen* (engl. *cage* Käfig) sind chemische Verbindungen, die bei Bestrahlung mit Licht bestimmter Wellenlängen das aktivierte Substrat freisetzen. Als caged-Verbindung wurde hierzu ein modifiziertes ATP mit der photolabilen Nitrophenyl-Ethylester- (NPE) Gruppe eingesetzt.

Zusammenfassung

Nach Etablierung und Optimierung der Reinigungsprotokolle für die untersuchten Proteine (UBA5, UFM1 und UFC1) wurde zunächst untersucht, ob das nativ an UBA5 gebundene ATP durch das NPE-ATP Derivat ausgetauscht werden kann. Diese ist ein essenzieller Schritt für die nachfolgende Durchführung von Pump-Probe-Experimenten mit zeitaufgelöster Kristallographie. Dabei wurde aus verschiedenen methodischen Ansätzen ein Ansatz ausgewählt, der durch Entfernung aller gebundenen Kationen zum Ziel hatte, dadurch die Affinität des ATP zum Protein zu senken. Anschließend wurde versucht durch Zugabe eines Überschusses an NPE-ATP einen stabilen UBA/NPE-ATP Komplex zu erzeugen. Der Austausch wurde mittels HPLC Verfahren qualitativ und quantitativ analysiert. Im Zuge dieser Arbeit war es nicht möglich einen Komplex von UBA5 mit NPE-ATP zu kristallisieren. Mit einer systematischen Analyse durch eine Kombination aus verschiedenen biochemischen, -physikalischen und -informatischen Verfahren konnte allerdings nachgewiesen werden, dass insbesondere die Verwendung von EDTA zur Entfernung der Kationen für die Reduktion der ATP Affinität einen erheblichen negativen Einfluss auf die zeitliche Stabilität des Proteins hatte und so offensichtlich eine erfolgreiche Kristallisation verhinderte.

Im zweiten der Teil der Arbeit wurde die Komplexbildung der untersuchten Proteine analysiert, insbesondere zwischen UFM1 mit UBA5 und UFC1. Für eine erste Charakterisierung der Dispersität der Proteine und Komplexe wurden diese mittels DLS untersucht. Die Ergebnisse dieser Experimente erforderten zur besseren Quantifizierung der Zusammensetzung der Komplexe den Einsatz mehrerer komplementärer Methoden. Die Quantifizierung wurde durch Bestimmung der Dissoziationskonstanten mittels AUC zunächst eingehender untersucht. Die Analyse ergab, dass insbesondere die Mischung von UBA5 und UFM1 sich aus verschiedenen Oligomeren zusammensetzte. Diese Ergebnisse wurden ergänzt durch Untersuchungen der einzelnen Proteine und der Komplexe mittels Kleinwinkelröntgenbeugung (SAXS). Dabei konnten im Vergleich zu bekannten Kristallstrukturen Modelle entwickelt werden, die den Zustand der Proteine in Lösung zeigen. In weiteren Schritten wurden die Oligomerisierungszustände von UBA5 und dem UBA5-UFM1 Komplex bei verschiedenen Konzentrationen durch Anwendung bioinformatischer Modellierungen quantitativ bestimmt. Auch eine

zeitaufgelöste Messung der Komplexbildung und die damit einhergehende Bestimmung des ungefähren Zeitrahmens konnte durch Kleinwinkelröntgenbeugung erfolgreich durchgeführt werden. Weiterhin konnten durch Einsatz von Microscale Thermophorese und Fluoreszenzspektroskopie die Dissoziationskonstanten der Komplexe bestimmt werden. Zwischen UFM1 und UFC1 konnte keine Komplexbildung beobachtet werden. Bei der Bestimmung der Affinitäten des UBA5-UFM1 Komplexes wurde eine negative Kooperativität festgestellt. In Kombination der Ergebnisse aus DLS, AUC und SAXS wurde ein Modell entwickelt, das eine schrittweise Komplexbildung zeigt. Die Bindung von UFM1 an UBA5 läuft über einen „cross binding“ Mechanismus, bei dem UFM1 über eine Untereinheit UBA5 gebunden wird, die Adenylierungsreaktion jedoch über die zweite UBA5 Untereinheit durchgeführt wird. Das entwickelte Modell zeigt dabei eine stufenweise Aktivierung, bei dem zunächst ein UFM1 gebunden und adenyliert wird, bevor ein zweites UFM1 Molekül bindet und den bisher bekannten Tetramerkomplex bildet.

Aims

The focus of this work was the analysis of the activation cascade, which activates the ubiquitin-like protein Ubiquitin Fold Modifier 1 (UFM1). Proteins involved in this cascade are the Ubiquitin Fold Modifier Activating Enzyme 5 (UBA5), the Ubiquitin Fold Modifier Conjugating Enzyme 1 (UFC1) and the Ubiquitin Fold Modifier Ligating Enzyme 1 (UFL1). Explicitly, this work shall investigate the first activation step structurally, which consists of the complex formation between UFM1 and UBA5. Another aspect of the investigation engages the question if UFM1 and UFC1 can form a complex prior to the activation by UBA5.

Another focus is the evaluation, if the ATP-binding protein is suitable for so called „pump-probe-experiments“ to investigate it with time resolved crystallographic experiments for analysis of the enzymatic mechanism with high resolution of time and space. In detail, it is necessary to produce and crystallize a stable complex consisting of UBA5 and a caged compound. Caged compounds are chemical compounds, which release its activated substrate upon radiation with light of specific wave lengths. The used caged compound group is a modified ATP with the photolabile nitrophenyl ethyl ester (NPE) group

Abstract

After Establishment and optimization of the purification protocols for the proteins under investigation (UFM1, UBA5, UFC1) it was firstly investigated if the natively to UBA5 bound ATP could be exchanged with NPE-ATP derivate. This is an essential step for the following pump-probe experiments with time resolved crystallography. Therefore, one approach out of several methods was selected, which aimed to reduce the affinity of ATP by removing all bound cations from the proteins. After that, it was tried to produce a stabile UBA5/NPE-ATP complex by adding an excess of NPE-ATP. The exchange was monitored quantitatively and qualitatively with HPLC procedures. During this work, it was not possible to crystallize UBA5 together with NPE-ATP. With a systematic analysis by a combination of different biophysical, biochemical and bioinformatic procedures, it could be shown that especially the usage of EDTA in the reduction of ATP affinity had a considerable effect on the temporal stability of the protein and therefore inhibited the successful crystallization.

The complex formation of the proteins under research was investigated, especially between UFM1 with UBA5 and UFC1. For a first characterization, the dispersity of the proteins and complexes were investigated with DLS. The results of this experiments made it necessary to quantify the composition of the complexes by the usage of several complementary methods. Firstly, the sedimentation constants were determined by AUC. The analysis revealed that especially the mixture of UFM1 and UBA5 consists of several different oligomers. The results were expanded by investigation of the single proteins and complexes with small angle X-ray scattering. In addition to the crystallographic model, models were developed which showed the state of the proteins in solution. In additional steps, the composition of UBA5 and the UBA5-UFM1 complex at different concentrations were determined quantitatively by using bioinformatical modeling. Additionally, time resolved measurements of the complex formation and the determination of time scale of complex formation was successfully performed. Furthermore, the dissociation constant was determined by using microscale electrophoresis and fluorescence spectroscopy. No complex formation between UFC1 and UFM1 was observed. The determination of binding constant of the UBA5-UFM1 complex a negative cooperativity was revealed. In combination with the results from

DLS, AUC and SAXS it was possible to develop a model, which showed a stepwise complex formation. The binding if performed via a cross binding reaction, where one UFM1 molecule binds to one UBA5 subunit and is adenylated by a second UBA5 subunit. The developed model however shows a stepwise activation, where one UFM1 molecule is first bound and adenylated before the second UFM1 binds and the so far known tetrameric complex is formed.

1 Introduction

1.1 Ubiquitin and Ubiquitin like Proteins

1.1.1 Ubiquitin and Ubiquitin Regulation

Ubiquitin is a very common protein and is found in all eukaryotic species. It was discovered in 1975 [2] as a new form of posttranslational modification and has a wide range of functions necessary for homeostasis of cells. The homeostasis is regulated by the ubiquitination of target proteins through isopeptide linkages via the N-terminal amino group or one of the 7 lysine ε-amino groups of ubiquitin in specific patterns (mono-, multi-, or polyubiquitination) [3].

One of the first discovered functions was the role in degradation of proteins by the 26S proteasome, when such proteins are ubiquitinated [4][5]. The ubiquitination and following degradation can also be used for regulation of protein activities. One example is the regulation of transcription factors, i.e. NF-κB, responsible for responses to inflammatory, immune system, developmental or stress stimuli [6]. The transcription factor is expressed in an inactive form with attached inhibitory molecules IκBs. Under external stimuli, these molecules are degraded by ubiquitin and the proteasome, and the active factor translocates into the nucleus for transcriptional activation [7][8]. Additionally, ubiquitin influences the cell cycle by regulating the levels of E2F through degradation, which is an important transcriptional factor in cell transition from G_1- to S-phase [9][10]. Ubiquitin is also responsible for tumor suppression by degradation of p53 [11]. Levels of ubiquitination can also influence the degradation of endoplasmic reticulum (ER) related proteins [12] or apoptosis [13].

Polyubiquitin chains are attached in most cases with the K48 residue to the to be degraded proteins [14]. Activation of ubiquitin is conserved in all species as a three-step cascade. It begins with the activating enzyme, generally called E1, which performs an ATP dependent ligation reaction between the C-terminal glycine residue of ubiquitin and a catalytic cysteine of E1 [15]. This thioester linkage is further transferred to the E2 enzyme, which is a conjugating enzyme. The final step is the transfer of ubiquitin to the final target with the enzyme E3 as a ligating enzyme [16]. Furthermore, deubiquitinating enzymes are additional means to regulate ubiquitin activity, either to release ubiquitin,

when its pool is depleted, proofreading false attached ubiquitin molecules [17], or regulating non-degenerative functions of ubiquitin, like cell cycle control [18].

Beside ubiquitin, there are several ubiquitin-like proteins (UBL) known, that all share a similar structure and mechanism of activation, but have various, more specialized targets and functions and fewer activating enzymes [19]. One of the most studied UBL proteins is the small ubiquitin-like modifier (SUMO). Functions vary from translocation of proteins [20] to response to stress like heat shock,[21] oxidative stress [22] [23], starvation [24], or transcriptional activity in immune responses [25]. The neural precursor cell-expressed, developmentally downregulated protein 8 (NEDD8) is an important activator of cullin-RING ligases [26], which can promote the transfer of ubiquitin to growing polyubiquitin chains [27]. The product of interferon stimulated gene 15 (ISG15) is an UBL which is activated by type I interferons as an immune response, secreted by virus-infected cells [28]. Two other UBL might also be involved in immune response: the HLAF adjacent transcript 10 (FAT10), and the monoclonal nonspecific suppressor factor β (MNSFβ). FAT10 can stimulate apoptosis as a reaction to human immunodeficiency virus (HIV) infected cells [29]. MNSFβ is able to stimulate phagocytosis [30]. Another function of UBL can be the stimulation of autophagy during starvation conditions by Atg8 [31] and Atg12 [32] (autophagy related genes). Prokaryotes contain no ubiquitin-like proteins, but share proteins with a high structural similarity involved in sulfur transport and transfer [33]. A linkage between the prokaryotic enzymes and UBLs was found in URM1 (ubiquitin related modifier), which can thiolate tRNA [34]. One of the latest identified UBL is the ubiquitin fold modifier 1 (UFM1), which is discussed together with its related proteins in the next sections.

The first crystal structure of ubiquitin was solved in 1987 at a resolution of 1.8 Å [35]. All UBL show a similar structure. They are composed of two short helices and 5 β-strands, known as the β-grasp fold. Two of the β-strands are parallel; the rest are antiparallel. Usually, the C-terminal residues are two glycines, which are involved in the formation of isopeptide bonds with the substrate (s. **Fig. *1-1***).

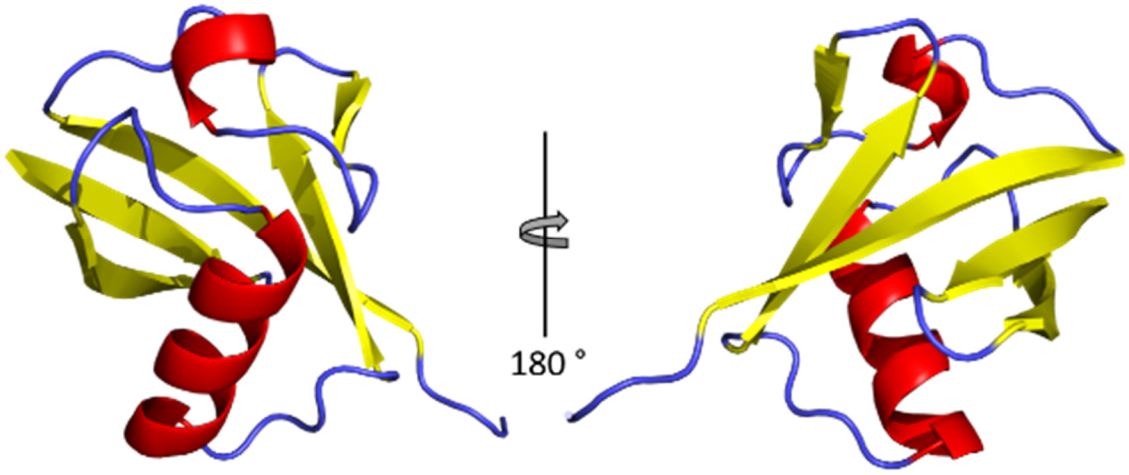


Fig. 1-1 Ribbon model of the crystal structure of ubiquitin showing the typical β-grasp fold consisting of five β-strands and two α-helices. Adapted from structural model by Vijay-Kumar et al[35] (pdb code 1UBQ).

1.1.2 UFM1

The ubiquitin fold modifier 1 (UFM1) was first described by Komatsu et al., together with its corresponding E1 and E2 enzymes: ubiquitin-like modifier activating enzyme 5 (UBA5) and ubiquitin fold modifier conjugating enzyme 1 (UFC1) [36]. The ubiquitin fold modifier ligating enzyme (UFL1) as the corresponding E3 enzyme was identified shortly afterwards [37]. UFM1 is expressed as a precursor, consisting of 85 amino acids with a molecular weight of 9.1 kDa, found in both the nucleus and cytoplasm [36]. The active form is processed by cleavage of the final two residues at the C-terminus, a serine, and a cysteine residue. The processed form has a glycine residue, typical for UBL, but with the difference that other UBL typically have a diglycine motive at the C-terminus [36] Another difference to ubiquitin is, that UFM1 is not found in yeasts. The cleavage is performed by two proteases: the UFM1- specific proteases UfSP1 and UfSP2, which can also deconjugate the targets [38]. The proteases are composed of 217 and 461 amino acid residues, respectively, and show no significant sequence homology to other known proteases but contain a cysteine as the active site. Besides their proteolytic function, these proteases can also play a role in an UFM1 independent mechanism of G-protein coupled receptor (GPCR) biogenesis [39].

It was found that UFM1 influences a variety of cellular functions and is involved in several pathogenic disorders. The first identified target of UFM1 was C20orf116, later called ubiquitin fold modifier binding protein 1 (UFBP1) [37]. UFBP1 consists of 314 amino

acid residues with a molecular weight of 38 kDa and forms an isopeptide bond with its K267 residue. Sequence homology studies revealed that the protein has a nuclear locating signal (NLS), a transmembrane helix and a C-terminal proteasome COP9-initiation factor (PCI) domain, which is able to form large complexes [37]. It is mainly found in the cytosolic site of the ER. The activating signal cointegrator 1 (ASC1) [40] was identified as the corresponding binding partner for UFBP1. ASC1 is a coactivator of the estrogen receptor α (ERα), which upon activation forms a dimer and translocate into the nucleus [41]. There it binds other transcriptional activators, which have an important influence in the development of breast cancer, if ASC1 is polyufmylated via the UFM1 K69 [42]. It was shown that UBA5 and UfSP2 are also involved in the development of breast cancer. Knock out of UBA5 inhibited tumor cell growth, while the knock-out of UfSP2 promoted it [43].

One of the main functions of UFM1 seems to be the regulation of ER stress response, as it plays an important role in the development of ischemic heart disease resulting from ER stress [44], ER development upon ER stress, and the activation of unfolded protein response [45]. Additionally, UFM1 prevents apoptosis in pancreatic Langerhans cells [46]; especially, it suppresses the ER stress response in diabetic organisms [47]. Furthermore, UFM1 is very important for erythroid development, since the knock-down of either one of the three activating enzymes, UBA5, UFC1, or UFL1, as well as UfBP1, leads to a severe anemia and death of the organism [48] [49]. Moreover, UFM1 regulates also fatty acid metabolism and cell proliferation [50]. UFM1 is transcriptionally downregulated in Mallory-Denk bodies, which can form in liver cells after chronic alcohol abuse[51], and it seems to be dysregulated in patients with schizophrenia [52].

The first structure of UFM1 was solved by NMR spectroscopy by Sasakawa et al. [53] in 2006. The model shows a very similar fold as ubiquitin, although the sequence similarity is only 16 %. Another difference is the uncharged surface compared to ubiquitin, which is assumed to be necessary to interact mainly with this surface with other proteins [53] (s. **Fig. *1*-*2***). The crystal structure was solved by Padala et al. in 2016[54].

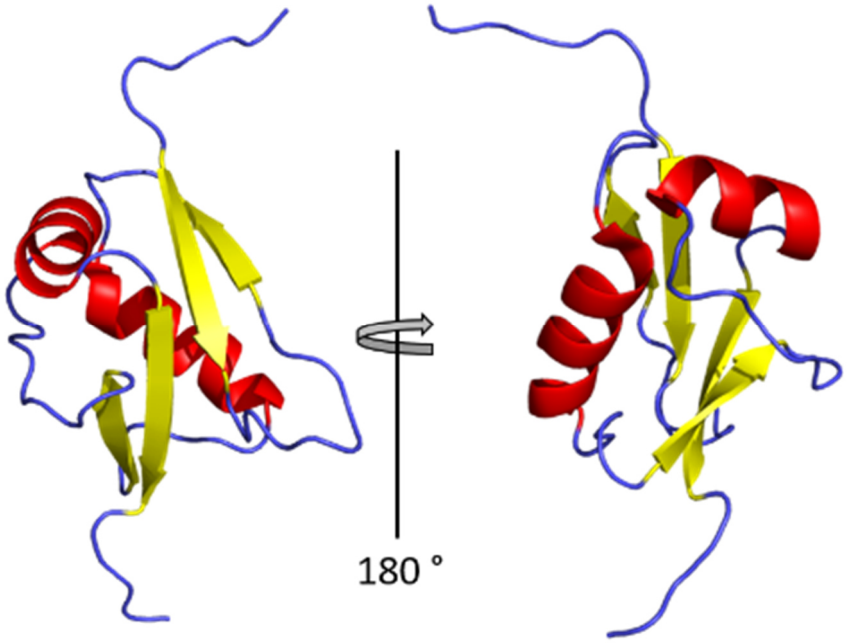


Fig. 1-2 Ribbon model of the NMR solution structure of UFM1 consisting of two α-helices and four antiparallel β-strands similar to the typical ubiquitin β-grasp fold. It has more flexible N- and C-terminal ends than the ubiquitin structure. Adapted from Sasakawa et al. (pdb id: 1WXS)

1.1.3 UBA5

The ubiquitin-like activating enzyme 5 (UBA5) is the only E1 enzyme for UFM1 discovered so far. It consists of 404 amino acid residues and has a molecular weight of 44.7 kDa [55]. It is mainly found in the cytoplasm and expressed in two isoforms: the full-length protein and a spliced variant, where the unconserved residues 1-56 are removed [56]. Usually, E1 enzymes are composed of three distinctive domains: an adenylation domain for binding of ATP, a domain for thioester formation, consisting of the first and second catalytic cysteine half domain (FCCH and SCCH, respectively), as well as the C-terminal ubiquitin fold domain, which is necessary for binding E2 [57]. Interestingly, the crystal structure of UBA5 shows, that it only has an adenylation domain composed of eight β-strands and seven α-helices with the catalytic cysteine C250 inside. On the contrary, UBA5 forms a dimer, at least in the crystal structure (**Fig. *1-3*** A) [56]. Furthermore, the activation process is a shorter two-step mechanism, instead of the typical E1 three-step mechanism [56]. Additionally, the thioesterfication of UFM1 works with a cross binding reaction. Two UFM1 molecules bind on each subunit of the UBA5 dimer, but the adenylation and thioester bond formation is performed by the other subunit (**Fig. *1-3*** B) [58] as it was shown with the crystal structure of the complex

by Oweis et al. in 2016. In addition, the binding of UFM1 also stabilizes the binding of the UBA5 dimer and the binding of ATP [59]. UFM1 is bound by the UFM1 interacting sequence (see Padala et al. 2016) (UIS) of UBA5 [54] and is necessary for the transfer of UFM1 to UFC1 [60]. UFC1 itself is bound by the C-terminal residues 381-404 [61]. Additionally, UBA5 has a zinc binding site, where the ion is coordinated by four cysteine residues [56].

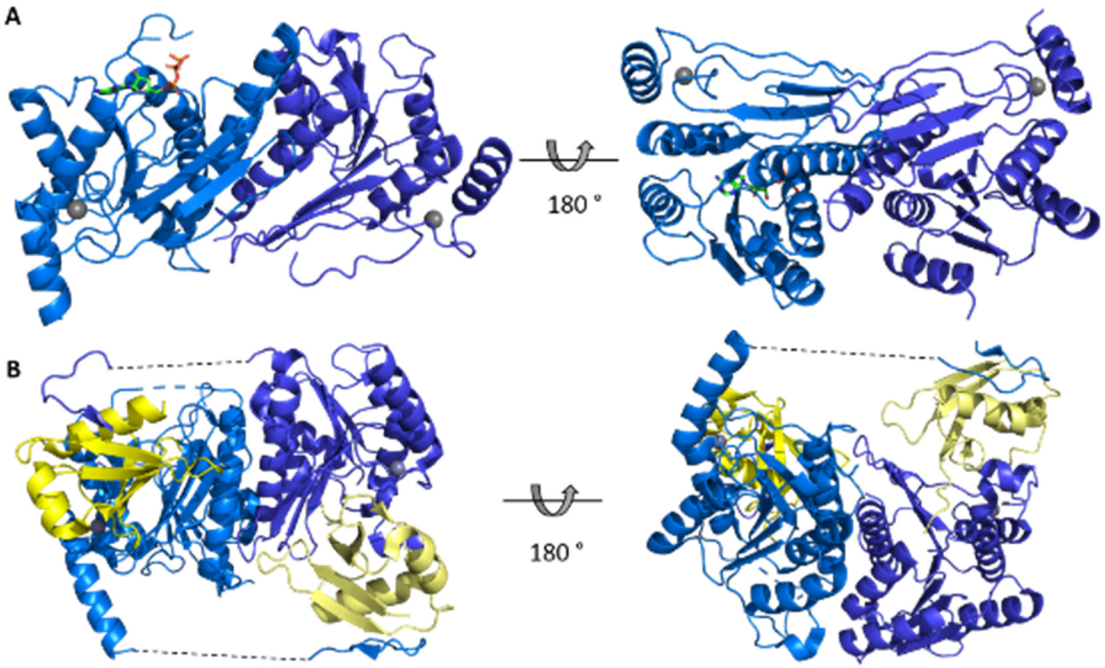


Fig. 1-3 Ribbon model of the crystal structure of UBA5 dimer (A) and in complex with UFM1 (B). A: Complex of two UBA5 monomers (light and dark blue). Visible are also the zinc (grey spheres) and bound ATP to one monomer (stick model). Model adapted from Bacik et al. (pdb id: 3H8V). B: Complex of two UBA5 monomers (blue) with UFM1 (yellow) bound to each subunit. The black dotted line shows a missing stretch of 15 residues. Model adapted from Oweis et al. (pdb id: 5IAA)

As already mentioned, knock-out of UBA5 leads to severe anemia in mice [49]. However, UBA5 is also known for other pathogenic symptoms caused by different single point mutations. Five different mutations were identified, which lead to severe disorders during early childhood development: movement disorders, microcephaly, deficits in the intellectual development and epilepsy. These mutations cause either loss of the thioester activity of UFM1 (V260M, M57V, K324N→fs14Stop, Q320Stop, G168E) or transthioester activity of UFM1 to UFC1 (A371T) [62]. Two other mutations were

identified related to autosomal recessive cerebellar ataxia, a neurodegenerative disease causing e.g., protein misfolding and degradation, defective DNA repair mechanisms and several metabolic defects. One of the identified mutations (R246X) leads to a loss of interaction between UBA5 and UFM1. The other mutation (K310E) still interacts with UFM1, but the exact pathogenesis remains unclear [63].

1.1.4 UFC1 and UFL1

The ubiquitin fold modifier 1 conjugating enzyme (UFC1) is 19.4 kDa protein consisting of 167 amino acid residues with the cysteine C116 as its catalytic residue [36,37]. Although the protein has relatively low sequence homology to other E2 enzymes, NMR and X-ray structures revealed a typical E2-like fold with an additional N-terminal helix, which is necessary for thermal stability (Mizushima et al., 2007)[64]. As mentioned in 1.1.3, UFM1 is transferred from UBA5 to UFC1 via a transthiolation reaction, which forms an isopeptide bond between the C-terminal glycine of UFM1 and the catalytic cysteine of UFC1. It was also shown that the neuronal cell adhesion molecule NCAM140 and together with UFM1 stimulates endocytosis. This might be a way to fine-tune the adhesion of proteins to the cell surface or surface of the ER [65].

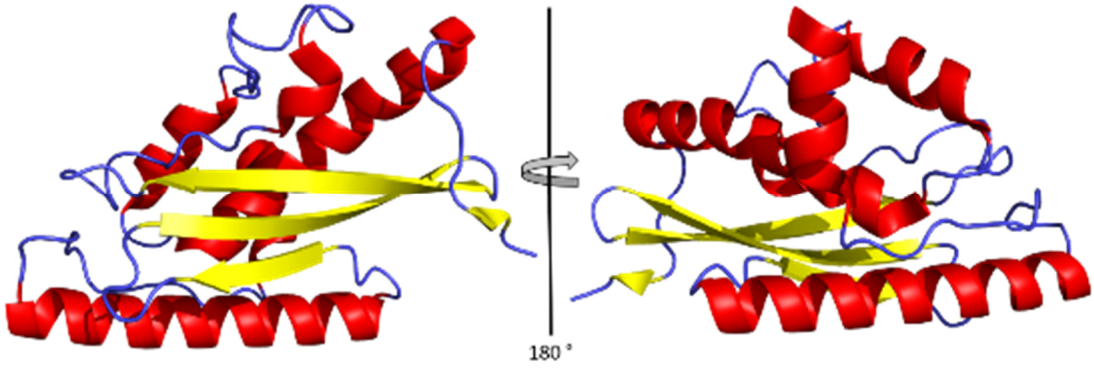


Fig. 1-4 Ribbon model of the crystal structure of UFC1. E2 enzymes have four β-strand propeller fold. Besides that, they have a high structural variation [66]. Model adapted from Mizushima et al. (PDB ID 2Z6O).

The ubiquitin fold modifier 1 ligating enzyme (UFL1) is an 89.5 kDa protein with 794 residues. It belongs to the type of non-canonical E3 enzymes as it contains neither the typical HECT domain (homologous to the E6AP carboxyl terminus) nor the RING domain (really interesting new gene) or the so-called Ub-Box [37]. Similar to UBA5, knock-out of UFC1 leads to severe anemia and death in mice [67]. Furthermore, it is involved in the

neurodegenerative disease, spinocerebellar ataxia type 1 by regulating the transfer from G1 to S phase during the cell proliferation in Berman glia cells [68].

The full process of UFM1 activation with all involved enzymes is depicted in **Fig. *1-5***. UBA5 acts as a homodimer and binds two UFM1 molecules by forming a thioester bond between the C-terminal glycine residue G83 of UFM1 and the catalytic cysteine C250 of UBA5 in a cross-binding mechanism, where the adenylation and thioester reaction is performed by the other UBA5 subunit. After binding of UFC1 by the terminal residues of UBA5, a heterohexamer is formed consisting of two molecules each of UBA5, UFM1 and UFC1. Then the thioester bond is transferred to the Cysteine residue C116 of UFC1. UFM1 molecule is consecutively transferred to the substrate molecule via a transfer of the thioester bond from UFC1 via UFL1 to the substrate.

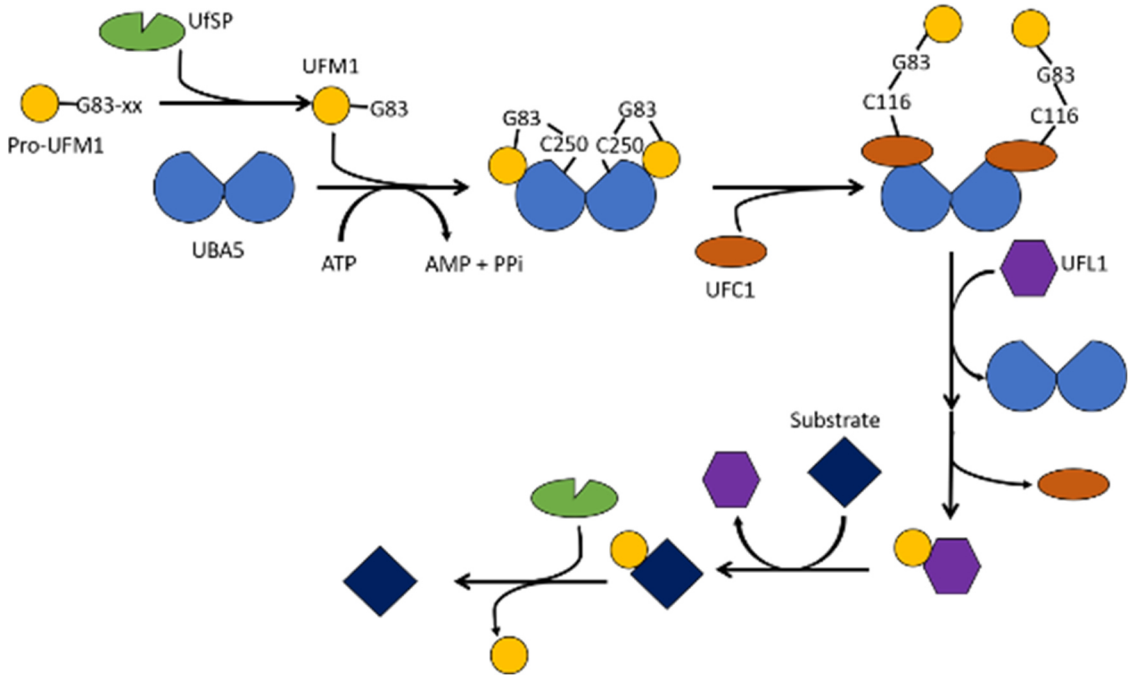


Fig. 1-5 Activation of UFM1 with a three-step mechanism. Pro-UFM1 (yellow) is cleaved by the UFM1 specific protease UfSP (green). UFM1 is bound to UBA5 (light blue) after adenylation of the C-terminal glycine G83 of UFM1 with a thioester bond with the cysteine residue C250 of UBA5. The thioester is transferred to the cysteine residue C116 of UFC1 (orange). The final activation step involves UFL1 (violet), which transfers UFM1 from UFC1 to the substrate (dark blue). This reaction can be reversed by cleavage of UFM1 by the protease UfSP.

1.2 Three-dimensional Structural Analysis of Proteins

1.2.1 Protein Crystallization

To obtain protein crystals, a highly pure, monodisperse solution in a supersaturated state is needed. Then, nucleation can occur by formation of weak intermolecular interactions, when the protein molecules self-assemble in a regular pattern [69].

In order to go from the soluble state into the crystal state, the free Gibson enthalpy ΔG (**Eq. *1***) has to reach a minimum [70].

$$\Delta G = \Delta H - T\Delta S$$

Eq. 1 Free Gibson enthalpy with ΔH: enthalpy; T: temperature in K and ΔS: entropy

Nevertheless, the transfer from the free translational and rotational transfusion of proteins to the ordered crystal state has a strong negative entropic term, which leads to a strong positive impact on the free enthalpy. Additionally, the enthalpic term due to forming crystal contacts is rather low. This raises the question of how the crystal formation can be thermodynamically favored if both terms would lead to a positive change in free enthalpy. The answer comes from the solvated water molecules, which are released when proteins form crystal contacts. This has a strong positive entropic term, which minimizes the free enthalpy of the crystallization process [69].

To reach the supersaturated state, the solubility of the protein is decreased, either by removing solvent, adding precipitants or adjusting the buffer composition (such as altering the pH) [71] [72]. Several crystallization methods were developed to achieve supersaturation, but the most common are vapor diffusion [73], batch crystallization [74] and counter diffusion [75]. Since crystallization is a form of phase separation, the nucleation process can be presented with a phase diagram.

Vapor diffusion is a widespread method, where small amounts of protein solution with high purity and concentration is mixed with small amounts of precipitant solution. The mixture is placed and sealed over a reservoir of the same precipitant solution, which usually contains, besides the precipitant, a buffer and additives like ligands [76] or detergents (if necessary) [77]. Since the concentration of the precipitants is initially higher in the reservoir than in the crystallization drop and the precipitants have hygroscopic properties, there is a concentration-dependent diffusion of water from the drop into the

reservoir until equilibrium is reached. This leads to a slow increase in protein concentration to a supersaturated solution, within which nucleation occurs.

Batch crystallization uses a different approach. A protein-precipitant mixture is placed under oil to avoid excessive evaporation. Commonly used oils for this purpose are usually paraffin oil, silicone oil or a mixture of both ("Al`s oil"). Here the effect of the precipitants to lower the protein solubility is used to achieve supersaturation in the solution. Nevertheless, water can also slowly diffuse into the oil (silicon oil), leading to an increase in protein concentration in the drop. When supersaturation is reached, nucleation and crystal growth can occur.

To crystallize proteins with the counter diffusion method, the protein solution is typically filled in a capillary and is placed in a precipitant solution or in a gel with precipitant. Due to the different concentrations of protein and precipitant on both ends of the capillary, these components diffuse to the other end. This leads to a concentration gradient along the capillary, where the protein can nucleate if supersaturation is reached on specific locations along the capillary. A schematic depiction of the common crystallization methods is depicted in **Fig. *1-6***.

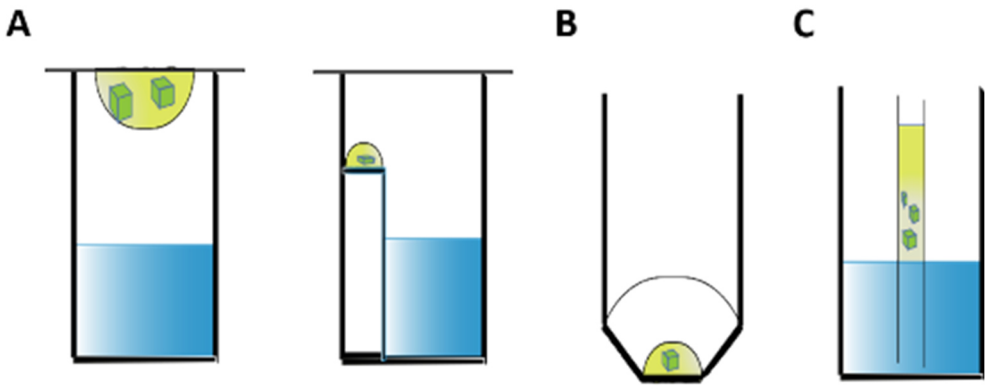


Fig. 1-6 Common crystallization methods. A: Vapor diffusion with the hanging drop (left) and sitting drop (right). The drop with the protein sample is placed over a reservoir on a cover slip or a small platform and sealed with grease. B: Batch crystallization under oil. C: Crystallization in a capillary with the counter diffusion method, where the capillary with protein solution is placed in a reservoir of precipitant solution

1.2.2 Protein crystallography

If protein crystals are obtained, they can be used for X-ray diffraction experiments with monochromatic, focused, polarized and coherent X-ray beams. In order to obtain high-resolution structural models of macromolecules like proteins, crystals of such macromolecules are exposed to X-ray radiation. The diffraction pattern, which occurs by the interaction of the X-rays with the electrons of the involved atoms, is recorded. From this pattern, the coordinates of the atom positions can be calculated. When the electromagnetic wave hits the atoms in the crystal, its electric field vector interacts with the electrons of that atom, leading to oscillation [69]. The oscillating electrons emit a photon with the same frequency (elastic scattering) in a discrete direction [78]. Inelastic scattering events (Compton scattering [79]) may occur. Nevertheless, they do not contribute to the diffraction, but to a diffuse background. Furthermore, 99 % of the X-ray photons pass the matter without any interaction. This leads to very weak diffraction from single molecules, explaining why crystals are needed to perform diffraction experiments, since the periodic arrangement of the molecules in the lattice leads to a strong amplification [69]. The emitted waves of the electrons have a phase difference depending on their relative position to each other and the incoming beam. This leads to constructive or destructive interference [80]. To observe diffraction, maximal constructive interference must take place. This happens, when Bragg's Law is fulfilled (Eq. *2*) [81].

$$n\lambda = 2d_{hkl}\sin\theta$$

Eq. 2 Bragg's Law. λ: wavelength; d_{hkl}: interplanar distance; θ: diffraction angle

This equation states that maximum interference occurs, when the path difference $d_{hkl}\sin\theta$ is equal to a multiple of λ, where θ is the diffraction angle and d the distance of the reciprocal lattice planes. The reciprocal lattice is theoretical construction to explain the diffraction and the corresponding diffraction pattern [82]. To construct a reciprocal lattice, the three vectors a^*, b^*, and c^* are calculated from their corresponding vectors a, b and c, which describe the real lattice. The vector a^* is perpendicular to b and c and has a magnitude of 1/a, b^* is perpendicular to a and c and has a magnitude of 1/b and c^* is perpendicular to a and b and has a magnitude of 1/c. A diffraction event can also

be viewed as a reflection of the X-ray on a plane. To further explain diffraction, the reciprocal lattice is intercepted by a set of reflection planes with a specific distance d. The number of times a plane intercepts the cell axis a^*, b^*, and c^* is described by their respective Miller indices h, k, and l [83]. From this construction a diffraction event can be explained with the Bragg equation: The interference from scattering electrons reaches a maximal constructive interference (and diffraction is measured) when all involved atoms are positioned on the constructed reflection planes which have a distance equal to a multiple of the wavelength of the incoming X-rays. All other atoms have a distance to each other, so that the Bragg equation is not fulfilled, and their scattered waves create destructive interference.

The amplitude of the scattered wave depends mostly on the type of atom itself, meaning the number of involved electrons (the electron density of the atom). This can be described by the atomic scattering factor (**Eq. *3***) [84].

$$f_s = \int_r^{V(atom)} \rho(r) \exp(2\pi i S r)\, dr$$

Eq. 3 The scattering f_s is the integration of the electron density ρ(r) and the phase of the emanating waves 2πiSr over the atom volume

The atomic scattering factors of all atoms participating in a diffraction event can be summarized with the structure factor, which sums up all scattering factors from all atoms in the unit cell of the lattice that contribute to the diffraction (**Eq. 4**) [85]. The intensity of the diffraction is proportional to the square of the modulus of the structure factor $|F_2|^2$ [69].

$$F_s = \sum_{j=1}^{atoms} f_{s,j}^0 * \exp\left(2\pi i S r_j\right)$$

Eq. 4 Structure factor equation for expressing the contribution of all atoms in the unit cell participating in a diffraction event

The typical setup for diffraction experiment (**Fig. *1-7***) consists of an X-ray source including monochromator and focusing elements, a goniometer, where the crystal is placed in the beam, and an array detector to collect the diffraction pattern (including a beam stop to protect the detector from radiation damage by the direct beam) [86]. In

order to collect a full dataset, the crystal must be turned with the goniometer. This brings new atoms in the Bragg conditions and more diffraction spots, which can be measured with the detector.

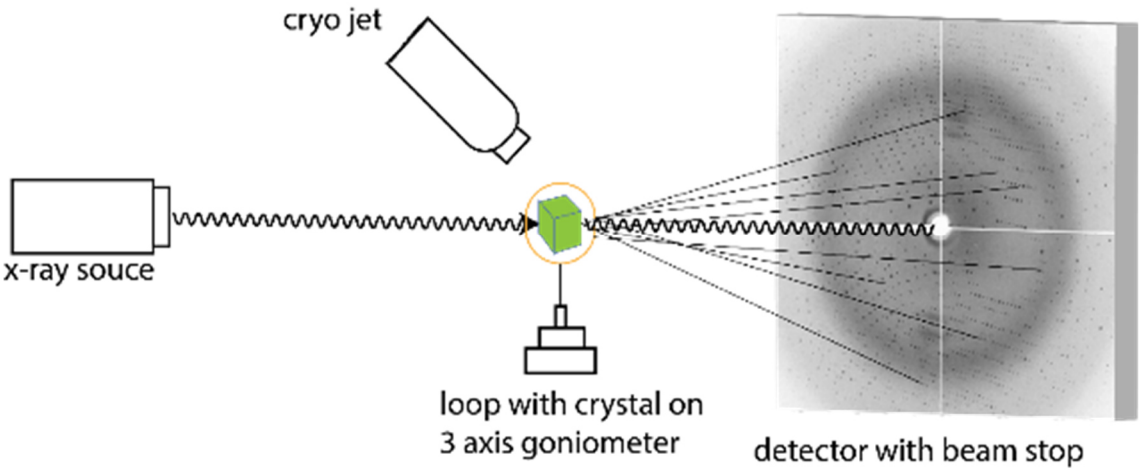


Fig. 1-7 Setup for crystal diffraction experiment. The crystal is collected from the crystallization solution with a loop and placed on a goniometer. The crystal is cooled to prevent radiation damage and reduce thermal vibration during illumination with the X-ray beam. The diffraction pattern is collected on a detector while turning the crystal with a goniometer. Diffraction pattern adapted from Chen et al[87].

The goal of a diffraction experiment is to calculate the electron density map from the diffraction pattern in order to find the coordinates of the respective atoms. Since the electron density is observed in real space and the diffraction pattern in reciprocal space, they can be calculated by each other's Fourier transform. The electron density ρ(r) can be calculated from the following equation **Eq. *5*** [88]:

$$\rho(r) = \frac{1}{V} \sum_{h=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} F(hkl) * \exp\left(-2\pi i(hx + ky + lz)\right)$$

Eq. 5 Calculation of the electron density from the phase and structure factors

The two important terms needed to calculate the electron density are the phases and the structure factors. The structure factors can be easily calculated from the square root of the measured intensity. This gives also information about the type of atoms involved in the diffraction. Unfortunately, the phase of the incoming beam cannot be extracted from the measured data. This is known as the phase problem in crystallography [89]. Luckily, several experimental and calculational methods have been developed to overcome this problem.

The most commonly used phasing methods are isomorphous replacement, anomalous dispersion and molecular replacement. Isomorphous replacement is an experimental phasing method, where the crystals are soaked with a small number of heavy atoms (e.g. Pb, Au, Ag), which should not affect the structure. Data from both the native and derivative crystal is collected and compared. The introduction of heavy atoms leads to strong intensity changes in some reflections, from which the position of these atoms can be deducted. This initial map can be used to calculate the rest of the atom positions. This method can be used by introduction of a single type of heavy atom (single isomorphous replacement, SIR) [90] or multiple heavy atoms (multiple isomorphous replacement, MIR) [91].

Anomalous Dispersion uses the effect of specific atoms to absorb X-rays at a specific energy [92]. This leads to a break of Friedel's law [93]. Usually to each of the structure factor F_{hkl}, there is a corresponding F_{-h-k-l} with the same magnitude. For anomalous dispersion, this is not the case, and leads to an intensity difference in specific reflections. These differences can be used to determine the position of atoms with the anomalous signal to get initial starting phases. From there, the other phases can be calculated. It is possible to measure only at one wavelength at the absorption edge of the atom (single-wavelength anomalous dispersion, SAD) [94]. To gain additional information, it is possible to measure at several wavelengths to get a data set with no anomalous signal, e.g. measuring at the high and low energy remote (multiple-wavelength anomalous dispersion, MAD) [95].

A known structure with high sequence similarity is used for molecular replacement, because high similarity often relates to high structural similarity. From the measured reflections and the model [96], a Patterson map[97] is calculated and compared. This can be used to determine known phases from the model and use them as initial phases for the calculation of electron density.

1.2.3 Small-Angle X-Ray Scattering

Small-angle X-ray (SAXS) scattering is a method to examine macromolecules in solution. It is possible to gather data about the molecular weight, size, structural information, as well as the behavior in solution. It is also possible to study structural changes upon addition of binding partners, ligands or buffer additives [98]. In addition to the advantage

of being able to examine the molecules of interest in a native state in solution, the method can gather data from small volumes and rather low concentrations. Data collection within the millisecond range is fast and sample preparation straight forward and established for standard experiments [99].

In a typical SAXS experiment, the sample is typically filled in a quartz capillary and illuminated by monochromatic, coherent X-rays. The scattering of the sample is recorded in a low degree range around the primary beam (0.3-5 °) by placing the detector usually several meters away from the sample [100]. Additionally, the pure solvent is measured in the same way and subtracted from the scattering of the sample to yield the scattering curve [100]. From analysis of the data, information like size, molecular weight or low-resolution structures like the overall shape can be extracted, as well as the behavior towards changes in the environment, like additives, denaturating agents or temperature [98]. It is also possible to analyze the behavior and quantify the composition and dynamics of polydisperse solutions through additional information like structural models from X-ray crystal or NMR structures [98].

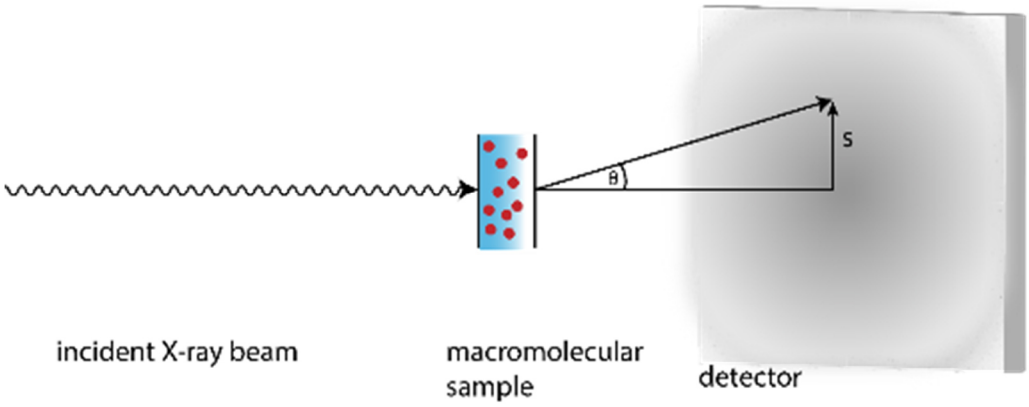


Fig. 1-8 Schematic representation of SAXS experimental setup. The macromolecular sample is illuminated with an X-ray beam with the wavelength λ. The radial scattering is collected with a detector in a small angle described by the scattering vector s.

The detector measures the scattering intensity of the solution. Since the macromolecules move freely, the scattering becomes isotropic and intensity is radially averaged and directly related to the momentum transfer s, expressed in equation **Eq. *6*** [100].

$$s = \frac{4\pi \sin\theta}{\lambda}$$

Eq. 6 Scattering intensity in relation to the momentum transfer with θ as the half-scattering angle and λ as the wavelength of the X-ray beam.

The measured intensity is the Fourier transform of the scattering amplitude A (**Eq. *7***) [99].

$$I(s) = \langle A(s)A^*(s)\rangle_\Omega$$

Eq. 7 Scattering Intensity as the Fourier transforms of the scattering amplitude. $\langle\rangle_\Omega$ is the average over all orientations

The amplitude depends on the electron density of the sample (**Eq. *8***) [99].

$$A(s) = \int \Delta\rho(r)\exp(isr)\,dr$$

Eq. 8 Scattering amplitude in relation to the electron density. Δρ(r) is the electron density difference between sample and solvent.

These calculations lead to a one-dimensional scattering curve, where the scattering intensity I is plotted over the scattering vector s. From the intensity, two important parameters can be extracted by the Guinier approximation (**Eq. *9***): the forward scattering I_0 and the radius of gyration R_g [101].

$$I(s) = I_0\exp\left(-\frac{1}{3}R_g^2 s^2\right)$$

Eq. 9 Guinier equation

Plotting the ln(I) over s^2 gives a curve that is linear for small angles, where R_g is the slope and I_0 is the intercept with ln(I). I_0 is a theoretical value for scattering intensity at angle, 0, which is proportional to the molecular weight [102]. By using references with known MW, the MW of the sample can directly be calculated from the I_0. The linearity of the Guinier approximation is also a quality factor for monodisperse systems, since a nonlinear plot can give evidence for polydispersity, imperfect background subtraction or interparticle interaction (repulsion or attraction) [103]. Assuming a uniform electron density distribution, the hydrated volume V_P can be estimated by the Porod equation (**Eq. *10***).

$$V_P = \frac{2\pi^2 I_0}{Q} \; with \; Q = \int_0^\infty s^2 I(s) ds$$

Eq. 10 Porod equation

The Porod equation can estimate the MW because the volume in nm³ is on average 1.66 times higher than the MW in kDa [104]. The interparticle distance is another parameter which can be extracted from the scattering profile by Fourier transform (**Eq. *11***) [105].

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty s^2 \quad I(s) \frac{\sin(sr)}{sr} ds$$

Eq. 11 Interparticle distance function

From the interparticle distance, the maximal distance D_{max} can be estimated; for spherical particles, the function would have bell-shaped structure with a maximum at $D_{max}/2$. Per definition, p is zero at r = 0 and r > D_{max} and non-negative between those values [99]. Additionally, R_g and I_0 can also be calculated from this function. Lastly, the folding state of the macromolecule can be analyzed from the scattering curve with the Kratky plot [106]. $s^2I(s)$ is plotted over s. Globular proteins show a peak at low angles with a decay of roughly $1/s^4$, while random coils increase at high angles and show no distinctive peak.

After analysis of these important parameters, additional tools can be used to model ab initio structures or analyze mixtures. The first approaches involved the calculation of scattering patterns from a simple bead model and compared them to the experimental data [107]. Improvements were made by using spherical harmonics for the calculations. A model composed of small beads is used, assuming a uniformly distributed electron density. These beads are modeled by an algorithm starting with an initial shape (usually a sphere) with a set diameter (usually D_{max}). Each bead is assigned either to solvent or protein. Each calculation step, one bead is changed, and a theoretical scattering curve is calculated from the shape. The goal is to minimize the difference between experimental and theoretical scattering curve [108].

The quality of the fit can be evaluated by the χ^2 value, which is also used to compare the fit quality of a calculated curve from an already existing model (X-ray-structure or NMR structure for example) and the experimental scattering curve. The lower χ2, the better

the fit [99]. This is especially a valuable tool for analysis of polydisperse systems by using different models to find the most probable composition, e.g. as used with the program CRYSOL [109].

$$\chi^2 = \frac{1}{N-1}\sum_{j=1}^{N}\left[\frac{I_{exp} - cI_{calc(s_j)}}{\sigma(s_j)}\right]^2$$

Eq. 12 Calculation of χ2 with N: number of data point; I_{exp}: experimental scattering intensity; c: correction factor; I_{calc}: theoretical scattering intensity, σ: experimental error

It is also possible to quantify the volume fractions of different components in polydisperse systems. This is based on the fact, that in an equilibrated, dilute system the scattering intensity is a linear combination of the scattering of individual components and their respective volume fractions (**Eq. *13***) [110] [111].

$$I(s) = \sum_k v_k I_k(s)$$

Eq. 13 Scattering intensity of a polydisperse system

Another way to analyze mixtures is the combination of SAXS with size exclusion chromatography (SEC-SAXS) [112]. In this approach, the sample is run through an SEC column directly prior to the illumination by X-rays. Assuming no formation of a new equilibrium, each component can be analyzed and modeled separately.

Furthermore, SAXS enables the possibility to analyze dynamics of structural changes, for example by addition of ligands, binding partners or environmental changes. With increasing brilliance and intensity of synchrotron beamlines, the time scale for data collection is reduced to microsecond range. This opens the possibility for time-resolved SAXS (TR-SAXS) experiments, when the time scale of the observed reaction is larger than the time collection of a scattering pattern and the system has a possibility to trigger the reaction. This can be temperature or pressure jumps, initiation of light-sensitive reactions by laser light, or, as most common fast mixing with stopped-flow mechanics. The drawback of TR-SAXS is the high sample consumption compared to conventional SAXS [113]. This adds to the drawbacks of conventional SAXS experiments, which include radiation damage [114], as well as the complicated analysis of polydisperse solutions and flexible systems [115]. Without additional information modelling has a high ambiguity due

to the fact that only the averaged scattering distribution of all involved compartments is recorded [116].

1.2.4 Comparison of both diffraction methods

Both X-ray diffraction and SAXS rely on the same measuring principle and share other similarities. However, they also differ in several details and therefore makes them predestined as complementary methods to answer different research questions.

Both methods require a highly, pure and homogenous sample. Data collection is also similar, as the data of the diffracted x-ray is recorded on a 2-dimensional detector. They also both suffer from issues, which are results of the radiation damage from the X-rays. While the crystals show a distinctive pattern of reflexes generated by the periodic arrangement of atom in wide angular range, SAXS only records the radially averaged, diffuse scattering at low angles.

Form the data, a structural model can be calculated. While x-ray diffraction of crystals results in molecular model with atomic resolution, SAXS is limited in its resolution yielding to more model showing the overall shape of molecule.

One big difference between them is the sample preparation. Crystallizing proteins can be a serious bottleneck, as without crystals, the structure can not be solved in atomic resolution. For SAXS a pure, homogenous solution of the molecule in an appropriate buffer is usually sufficient. One advantage of SAXS is also, that it can be used to investigate on the dynamics of molecules and complexes in solution, for example complex formation or structural changes from adding additives, temperature change or behavior under different conditions. X-ray structures from crystals usually provide are more detailed model of the structure and the enzymatic activity of proteins can still be preserved in the crystal. Nevertheless, the close crystal contacts and the arrangement of molecules in the crystal can lead to artifacts which do not represent their native state. Due to the flexibility of the proteins, it is often possible to crystalize only parts of the protein, while SAXS can still provide useful information of the whole protein. However, solving a structure with an atomic resolution from crystals is still one of the best methods to fully understand the function of protein on a molecular and mechanistic level. Structural models from protein crystals are also immensely useful and often

necessary for example to create promising pharmaceutical compounds, which can bind to the catalytic center of a target protein.

1.3 Biophysical Methods to Characterize Protein-Protein Interactions

1.3.1 Fluorescence Spectroscopy

Fluorescence is the emission of photons by molecules after their excitation by photons from the ground state. Due to energy losses to the surroundings, the emitted energy when going back to the ground state is lower than the excitation energy (Stokes shift) [117]. This leads to different absorption and emission spectra [118]. The transition between the different energy states can be visualized with the Jablonski diagram (**Fig. *1*-*9***) [119]. The timescale of excitation and emission is very short (10^{-9} s) [120]. Each fluorophore has a specific quantum yield, as described by the ratio of absorbed and emitted photons [121]. Since the quantum yield of a fluorescence phenomenon is highly dependent of the fluorophore and its surroundings [122], it can be used in fluorescence spectroscopy to study molecular interactions including binding kinetics of protein-protein interactions [123] or folding state and stability [124]. It is also applied in fluorescence microscopy, where fluorophores are attached to specific molecules to allow the localization in tissues [125], e.g. with green fluorescent protein (GFP) [126].

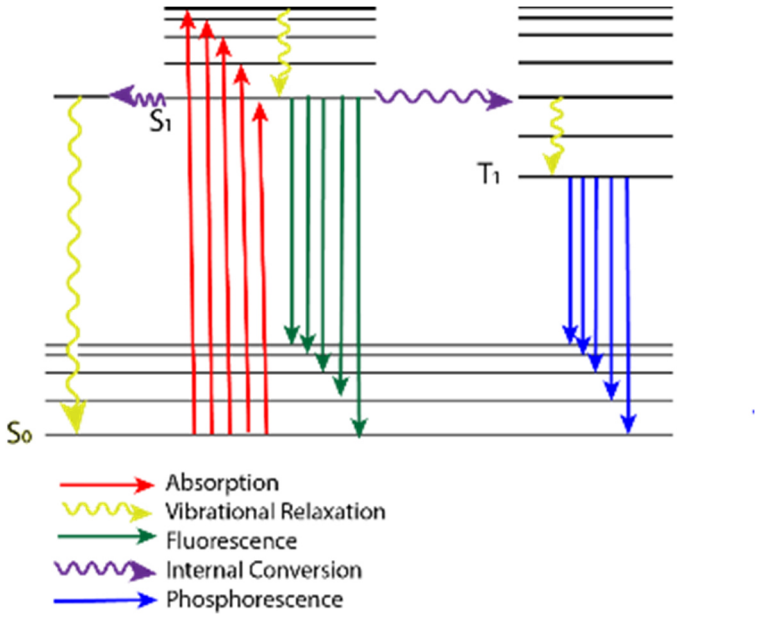


Fig. 1-9 Possible energy transitions visualized by the Jablonski diagram. After excitation of electrons from S_0 ground state to the S_1 excited state (red), electrons can return to the ground state by vibrational relaxation (yellow) or by emitting light via fluorescence (green). Another possibility is the emission of light due to phosphorescence (blue) after a slow internal conversion transition (purple).

To study such interactions, different types of fluorophores can be used: intrinsic, coenzymic, or extrinsic [120]. Intrinsic fluorophores involve amino acid residues with aromatic side chains: phenylalanine, tyrosine, and especially tryptophan. Coenzymic fluorophores are natural coenzymes like NADH (nicotinamide-adenine dinucleotide) or FADH (flavine- adenine dinucleotide). Extrinsic fluorophores must be introduced manually, if no of the other two types are abundant on the proper position in the protein. Types of extrinsic fluorophores range from organic compounds, peptides and proteins up to nanoparticles (quantum dots) [123].

Furthermore, it is possible to gain insight into structural properties with Förster resonance energy transfer (FRET). This phenomenon is observed when two different fluorophores are in close proximity to each other. The energy of an excited fluorophore can be transferred if there is a spectral overlap between the emission spectra of the donor and the excitation spectra of the acceptor. The quantum yield of the energy transfer is highly dependent from the distance of both fluorophores. Therefore, it can be used to study the distance of two different species, by labeling them on specific

positions and measure the fluorescence of the acceptor molecule, when exciting the donor molecule (s. **Eq. *14***) [127].

$$\kappa_T = \left(\frac{1}{\tau_D}\right) * (\frac{r_0}{r})^6$$

Eq. 14 FRET rate κ_T. τ_D is the donor fluorescent lifetime, r the distance between both fluorophore and r_0 the distance at which 50 % of the energy is transferred from donor to acceptor.

1.3.2 Analytical Ultracentrifugation

The analytical ultracentrifugation (AUC) is a method originally developed by Svedberg [128] to analyze polydispersity of solute macromolecular systems [129] [130], determine the molecular weight of particles over a wide range of size distribution ranging from sucrose to virus particles [131] [132], and study the interaction of molecules, e.g. self-association [133] or complex formation [134]. The sample particles are centrifuged at high velocity. In this produced gravitational field, three forces act on the particles: the gravitational force F_g, the buoyant force F_b and the frictional force F_f [135].

$$F_g = \frac{M}{N}\omega^2 r$$

Eq. 15 Gravitational force with M: molecular mass of particle; N: Avogadro's number; ω: angular velocity and r: distance from the rotation axis

The buoyant force and the frictional force act in the opposite direction of the gravitational force

$$F_b = -m_0\omega^2 r \qquad m_0 = m\bar{v}\rho$$

Eq. 16 Buoyant force with m_0: mass of fluid displaced by particle, $\bar{v}$ partial specific volume that each gram of the solute occupies in solution, ρ: density of the solvent

$$F_f = -fu$$

Eq. 17 Frictional force with f: frictional coefficient (shape factor) and u: particle velocity

After a short time, these forces balance and the equations can be summarized, where the term $u/\omega^2 r$ is called the sedimentation coefficient s:

$$\frac{M(1-\bar{v}\rho)}{Nf} = \frac{u}{\omega^2 r} \equiv s$$

Eq. 18 Calculation of the sedimentation coefficient

In order to calculate the sedimentation coefficient (**Eq. *18***) two additional parameters need to be known: the solvent density and the partial specific volume $\bar{v}$, which can both be calculated with high accuracy [136]. Additionally, sample viscosity and concentration effects must be considered [135]. Modern-day software uses the Lamm equation [137] to extract the Svedberg coefficient to further calculate the molecular weight of the sample.

$$\frac{\partial \chi(r,t)}{\partial t} = \frac{\partial}{r\partial r}\left[rD\frac{\partial \chi(r,t)}{\partial r\partial} - s\omega^2 r^2 \chi(r,t)\right]$$

Eq. 19 Lamm equation, describing the concentration distribution χ of the particle in relation to the radial distance and time

The Lamm equation describes the evolution of the concentration distribution in a sector-shaped centrifugation cell in relation to radial distance and time [138]. From this, the molecular weight distribution of mono- or even polydisperse systems can be modeled [139].

2 Results

2.1.1 Protein Purification and Characterization

2.1.1.1 Purification and Characterization of UBA5

Two different constructs were used for the experiments. The first construct including the residues 57-329 (sequence in **Fig. *2-3***) were used for crystallization and the experiments which involved NPE-ATP. It was cloned from the synthetic gene encoding the residues 37-404 of human UBA5. The second construct including the residues 57-346 (sequence displayed in **Fig. *2-3***) was mostly used for the complex formation investigations involving UFM1. The additional C-terminal residues are necessary for proper binding of UFM1. For UBA5, several expression methods have been tested, but the highest protein expression was found using the methods described in section 4.2.2.2. The purification was performed via affinity chromatography and size exclusion chromatography as described in sections 4.2.3.1 and 4.2.3.2. The two-step purification as described in the literature [56] is sufficient to yield pure UBA5 (s **Fig. *2-1*** and **Fig. *2-2***). Yield and purity were checked by SDS-PAGE.

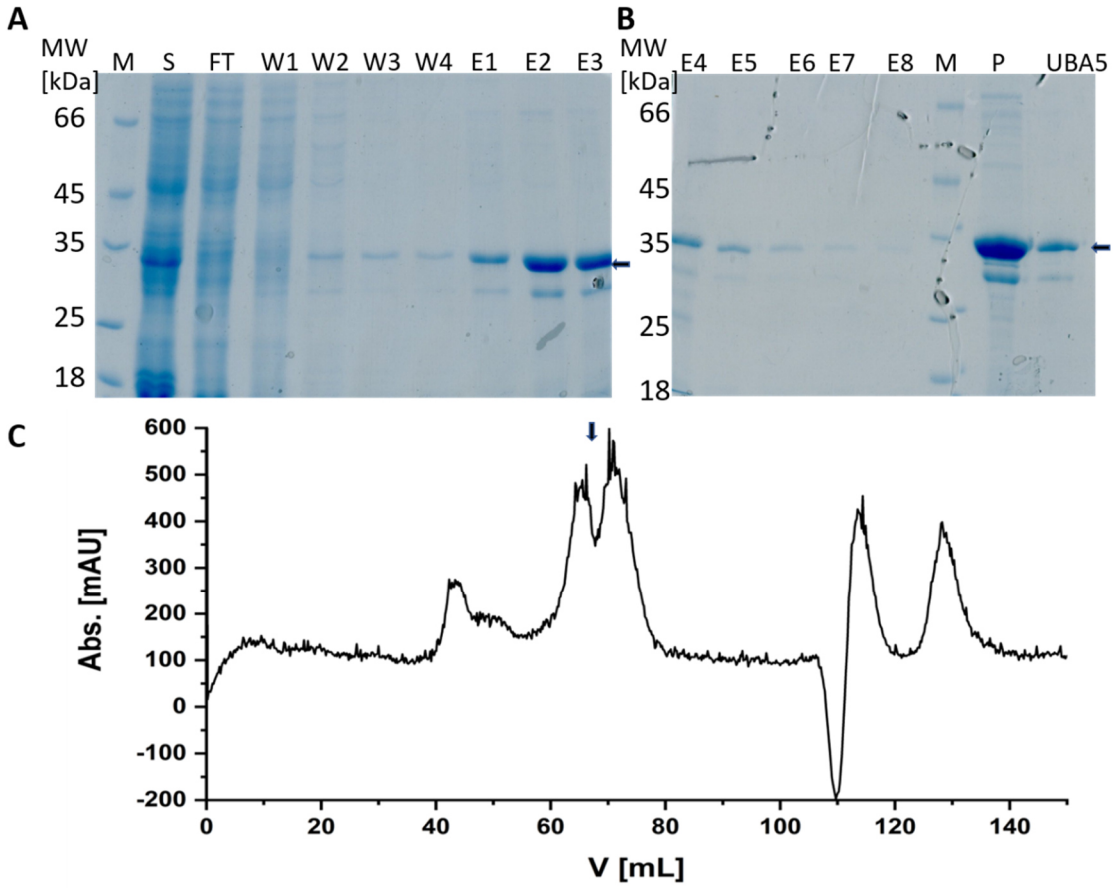


Fig. 2-1 Purification of UBA5 57-329. Top left: Gel shows the supernatant (S) of the centrifugation step (s. 4.2.3.1), the flow-through from the affinity chromatography (FT), the washing steps (W1-W4) and the first three elution steps (E1-E3). The UBA5 protein band is marked with a black arrow. Top right: Gel shows the last five elution steps (E4-E8), the pooled fractions applied to the SEC (P), and the final purified UBA5. The UBA5 protein band is marked with a black arrow. Molecular weights of the marker bands (M) are also displayed. Bottom: Chromatogram of the SEC purification by plotting the absorption in mAU over the elution volume. The UBA5 double peak is marked with the black arrow. The noisy line comes from the addition of 2 mM ATP to the SEC buffer.

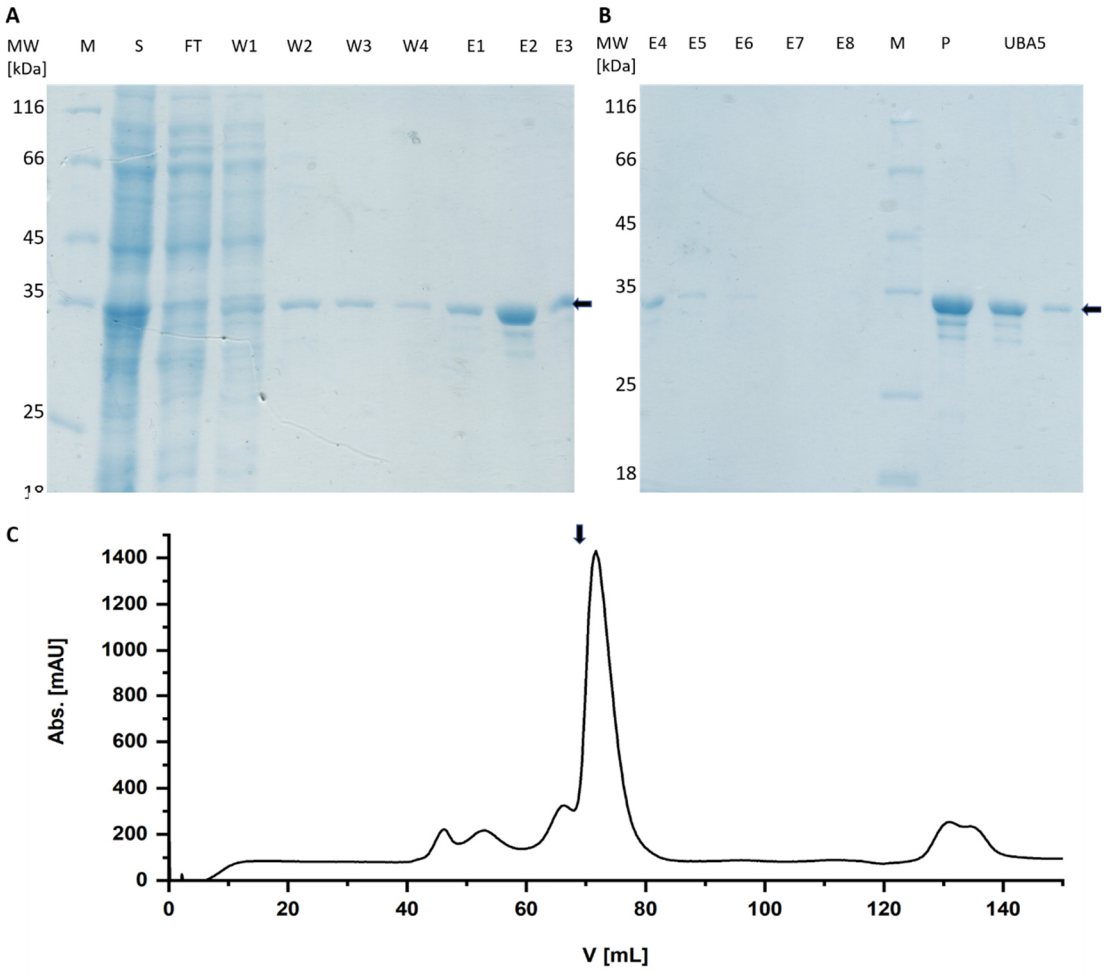


Fig. 2-2 Purification of UBA5 57-346. Top left: Gel shows the supernatant (S) of the centrifugation step (s. 4.2.3.1), the flow-through from the affinity chromatography (FT), the washing steps (W1-W4) and the first three elution steps (E1-E3). The UBA5 protein band is marked with a black arrow. Top right: Gel shows the last five elution steps (E4-E8), the pooled fractions applied to the SEC (P)s and the final purified UBA5 in the last two lanes from both elution peaks from the SEC. The UBA5 protein band is marked with a black arrow. Molecular weights of the marker bands (M) are also displayed. Bottom: Chromatogram of the SEC purification by plotting the absorption in mAU over the elution volume. The UBA5 double peak is marked with the black arrow.

SDS-PAGE shows clear distinctive bands of 32 kDa or 34 kDa, which are the expected size of monomeric UBA5 57-329 or 57-346, respectively. A second band appears in the elution fractions below the UBA5 main band (s SDS-PAGE of **Fig. *2-1*** and **Fig. *2-2***). It was identified as UBA5 by mass spectrometry (MS) analysis of a tryptic digest of both bands and showed no difference from the main band, except that for the lower band the his-tag is missing, probably due to proteolytic degradation (**Fig. *2-3***).

MGSSHHHHHHSSGLVPR||GS||MALKR||MGIVSDYEK||IR||TFAVAIVGVGGVGSVTAEMLTR||CGIGK||
LLLFDYDK||VELANMNR||LFFQPHQAGLSK||VQAAEHTLR||NINPDVLFEVHNYNITTVENFQHFMDR||I
SNGGLEEGKPVDLVLSCVDNFEAR||MTINTACNELGQTWMESGVSENAVSGHIQLIIPGESACFACAPPL
VVAANIDEK||TLKR||EGVCAASLPTTMGVVAGILVQNVLK||FLLNFGTVSFYLGYNAMQDFFPTMSMKP
NPQCDDR||NCRKQ||QEEYK||KK||VAALPK||QEVIQEEEEIIHEDNEWGIELV

MGSSHHHHHHSSGLVPRGS||MALKR||MGIVSDYEK||IR||TFAVAIVGVGGVGSVTAEMLTR||CGIGK||LL
LFDYDK||VELANMNR||LFFQPHQAGLSK||VQAAEHTLR||NINPDVLFEVHNYNITTVENFQHFMDR||ISN
GGLEEGKPVDLVLSCVDNFEAR||MTINTACNELGQTWMESGVSENAVSGHIQLIIPGESACFACAPPLVV
AANIDEK||TLKR||EGVCAASLPTTMGVVAGILVQNVLK||FLLNFGTVSFYLGYNAMQDFFPTMSMKPNP
QCDDR||NCRK||QQEEYK||KK||VAALPK||QEVIQEEEEIIHEDNEWGIELV

Fig. 2-3 Comparison of identified MS fragments of tryptic digest of both UBA5 57-346 bands. Red shows the residues which could be identified by MS, back double bars mark the fragments of the tryptic digest. Upper sequence shows the result of the upper band from the SDS-PAGE, lower sequence shows the result of the lower band. The sequence of UBA5 57-329 ends before the second last fragments with three lysine residues.

The chromatogram of the SEC run shows several peaks. One elutes at around 45 mL, which are mostly aggregates eluting at the void volume of the column. UBA5 is eluted in two peaks between 60 and 80 mL and their affiliated MW could be determined by calibration of the column as dimer and monomer peak (s. **Fig. S *7***). The last peaks are buffer components. Comparison of the SEC of both constructs shows several differences. The SEC of UBA5 57-329 shows a noisy signal. This is caused by the addition of ATP, which influences the measured absorption signal by its property to slightly absorb at 280 nm [140].

To check the oligomeric and folding state, the sample was analyzed with DLS and CD spectroscopy as described in sections 4.2.4.2 and 4.2.4.3. The folding state of UBA5 47-329 could be verified by CD spectroscopy showing a similar secondary structure of both constructs compared to the secondary structure estimation by the online single spectrum and fold recognition tool BeStSel (**Tab. *2-1***, **Fig. S *3***, **Fig. S *4***) and to the annotation in the pdb.

Tab. 2-1 Calculated secondary structure of both UBA5 constructs from CD spectra with the BeStSel server in comparison with the sequence annotation deposited in the pdb.

Protein	α-helix [%]	β-sheet [%]
UBA5 57-329	33.7	15.9
UBA5 (3H8V)	32.0	15.0
UBA5 57-346	28.4	6.3
UBA5 (5IAA)	33.0	16.0

The secondary structure annotation of UBA5 57-346 shows variation to the composition deposited in the respective pdb entry (pdb ID: 5IAA). Both, the α helical proportion and especially the β-fold proportion are lower than expected. This can be explained by the suboptimal fit in the 190-200 nm region (s. **Fig. S *4***), probably due to higher buffer absorption in this wavelength range. DLS of UBA5 57-329 with a concentration of 3.4 mg/mL gives a signal of 3.6 nm, which corresponds to a MW of 65 kDa, which is very close to the expected MW of 66 kDa. Nevertheless, the signal is broad, which means that the protein is not in the monodisperse state, but in an equilibrium of different oligomers (**Fig. *2-4***).

UBA5 57-346 with a concentration of 2.6 mg/mL gives a monodisperse signal at 3.4 nm, which corresponds to 56 kDa (**Fig. *2-5***). This is smaller than the expected MW of 68 kDa, which could also mean that there is an equilibrium between mono- and dimer influencing the signal.

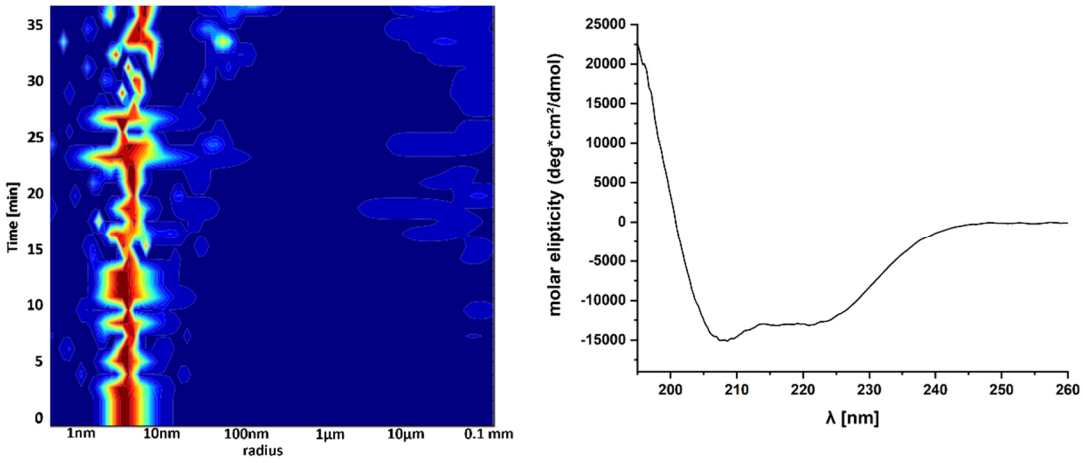


Fig. 2-4 DLS and CD spectra of purified UBA5 57-329. Left: DLS of purified UBA5 measured at 3.4 mg/mL shows a monodisperse peak with a radius of 3.6 nm, but with a relatively broad distribution. Right: CD spectra of UBA5 of molar ellipticity measured between 195 and 260 nm.

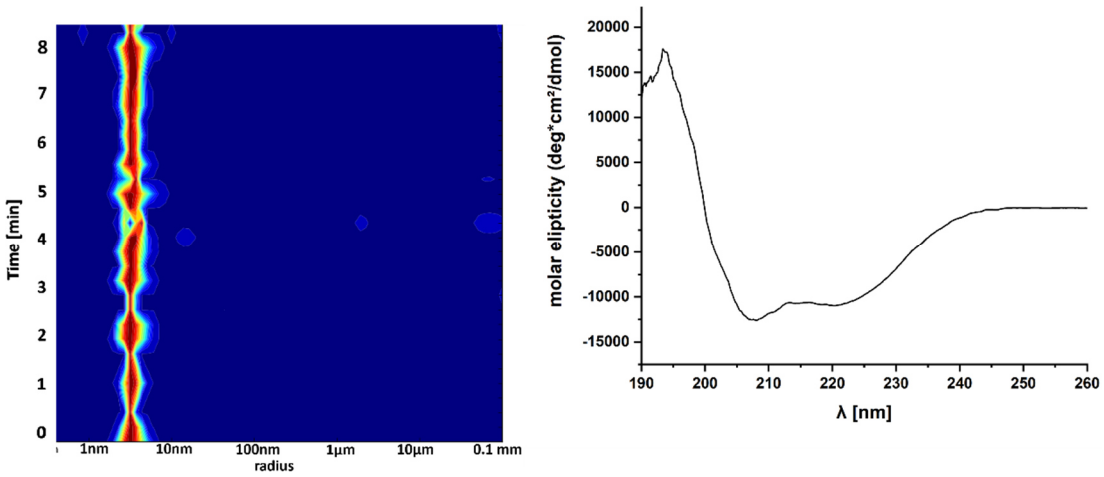


Fig. 2-5 DLS and CD spectra of purified UBA5 57-346. Left: DLS of purified UBA5 measured at 2.6 mg/mL shows a clear monodisperse peak with a radius of 3.4 nm. Right: CD spectra of UBA5 in molar ellipticity measured between 190 and 260 nm

2.1.1.2 Purification and Characterization of UFM1

The gene was cloned into the plasmid from a synthetic gene encoding the residues of human UFM1. After successful expression, purification of UFM1 was performed via affinity chromatography and size exclusion chromatography as described in sections 4.2.3.1 and 4.2.3.2. Yield and purity were checked by SDS PAGE (**Fig. *2-6* A+B).** The purified UFM1 band is clearly visible in the SEC fraction (**B**) while on (**A**) the uncleaved UFM1-GST slightly below 35 kDa marker band and cleaved GST in the middle between the 35 and 25kDa marker band are visible. The SEC chromatogram shows three peaks (**Fig. *2-6* C)**. The first elutes between 65 and 75 mL and was identified as a GST dimer by mass spectrometry. The second peak elutes around 90 mL and was identified as UFM1 by SDS-PAGE and the last peak contains buffer components. Although the UFM1 peak is quite small with a relative absorption of around 50 mAU, the yield is sufficient for further experiments. The peak size is caused by the low extinction coefficient as UFM1 is composed of only 83 residues and does not contain any high absorbing tryptophanes.

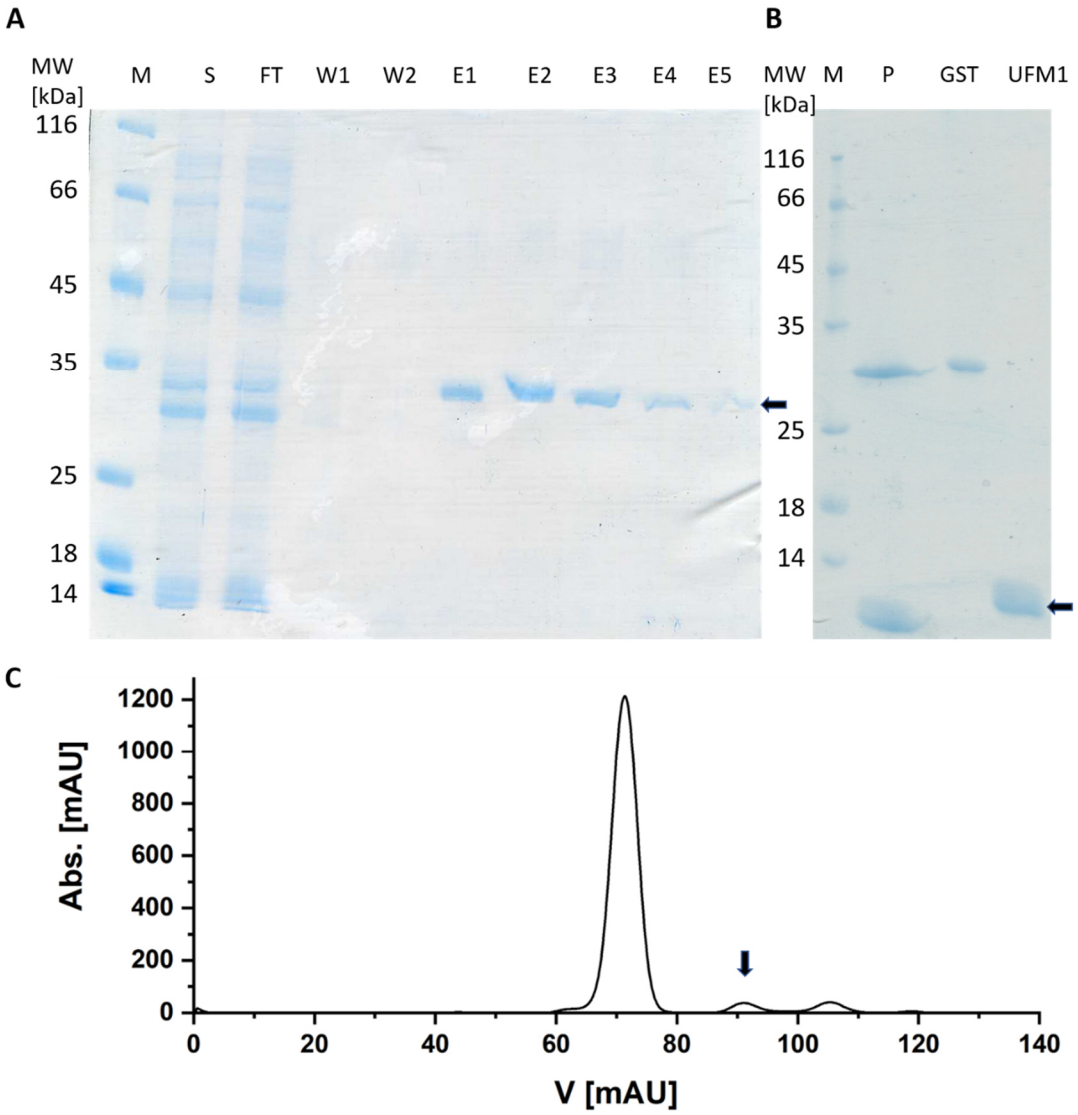


Fig. 2-6 Purification of UFM1. Top left: Gel shows the supernatant (S) of the centrifugation step (s. 4.2.3.1), the flow-through from the affinity chromatography (FT), the washing steps (W1-W2) and the first five elution steps (E1-E5). The UFM1-GST protein band is marked with a black arrow. Top right: Gel shows the pooled fractions applied to the SEC (P) and eluted GST, and the final purified UFM1. The UFM1 protein band is marked with a black arrow. Molecular weights of the marker bands (M) are also displayed. Bottom: Chromatogram of the SEC purification by plotting the absorption in mAU over the elution volume. The UFM1 peak is marked with a black arrow.

To check for oligomeric and folding state the sample was analyzed with DLS and CD spectroscopy as described in 4.2.4.2 and 4.2.4.3. The DLS signal shows with a concentration of 9 mg/mL a peak at 1.5 nm, which corresponds to a molecular weight of 8 kDa and is very close to the expected 8.9 kDa. However, the peak shows broad signals, which comes from the small size of the molecule. This makes it a weak scatterer

leading to a low signal-noise ratio. The CD measurement shows that the protein is properly folded and similar to the secondary structure assignment in the pdb deposition; however, the spectrum shows a smaller amount of β-fold.

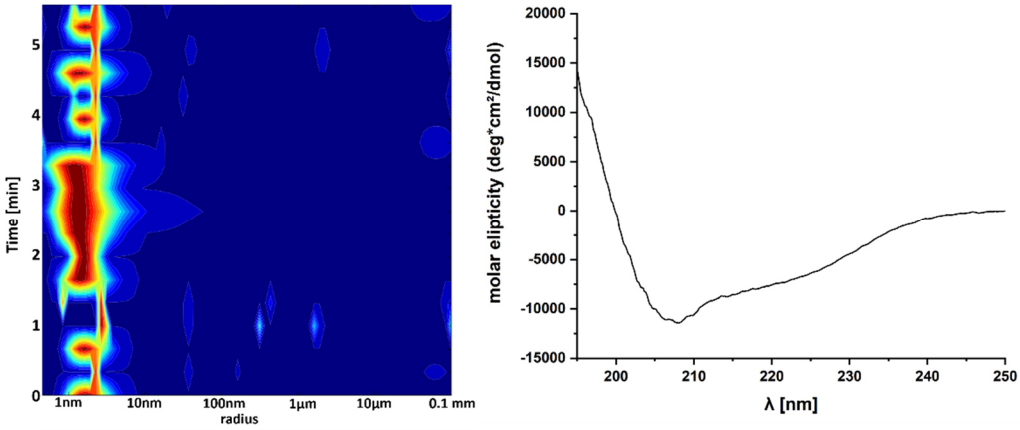


Fig. 2-7 DLS and CD spectra of purified UFM1. Left: DLS of purified UFM1 with a concentration of 9 mg/mL shows a monodisperse peak with a radius of 1.5 nm but with variation in size distribution due to the weak scattering from the small molecule. Right: CD spectra of UFM1 in molar ellipticity measured between 195 and 260 nm.

Tab. 2-2 Calculated secondary structure of UFM1 from the CD spectrum with the BeStSel server in comparison with the sequence annotation deposited in the pdb.

Protein	α-helix [%]	β-sheet [%]
UFM1	23.0	21.9
UFM1 (1WXS)	20.0	25.0

2.1.1.3 Purification and Characterization of UFC1

After successful expression, purification of UFC1 was performed via affinity chromatography and size exclusion chromatography as described in sections 4.2.3.1 and 4.2.3.2. The gene was cloned into the plasmid from a synthetic gene encoding the residues of human UFM1. Yield and purity were checked with SDS PAGE. Cleavage of the GST-Tag was performed as the optimized method used for UFM1-GST cleavage. The SDS gel (**Fig. *2-8* A+B**) shows that UFC1 can also be cleaved in the same way as UFM1, resulting in a single band after SEC. Nevertheless, the elution fractions of the affinity chromatography show a second band at 26 kDa in addition to the UFC1-GST band at 45 kDa. This could be identified as non-fused GST, which might reason by cleavage from proteases present in the supernatant. The SEC chromatogram shows four peaks (**Fig. *2-8***

C). The first elutes between 65 and 75 mL and contains GST dimer. UFC1 elutes around 90 mL, and the last two peaks contain buffer components.

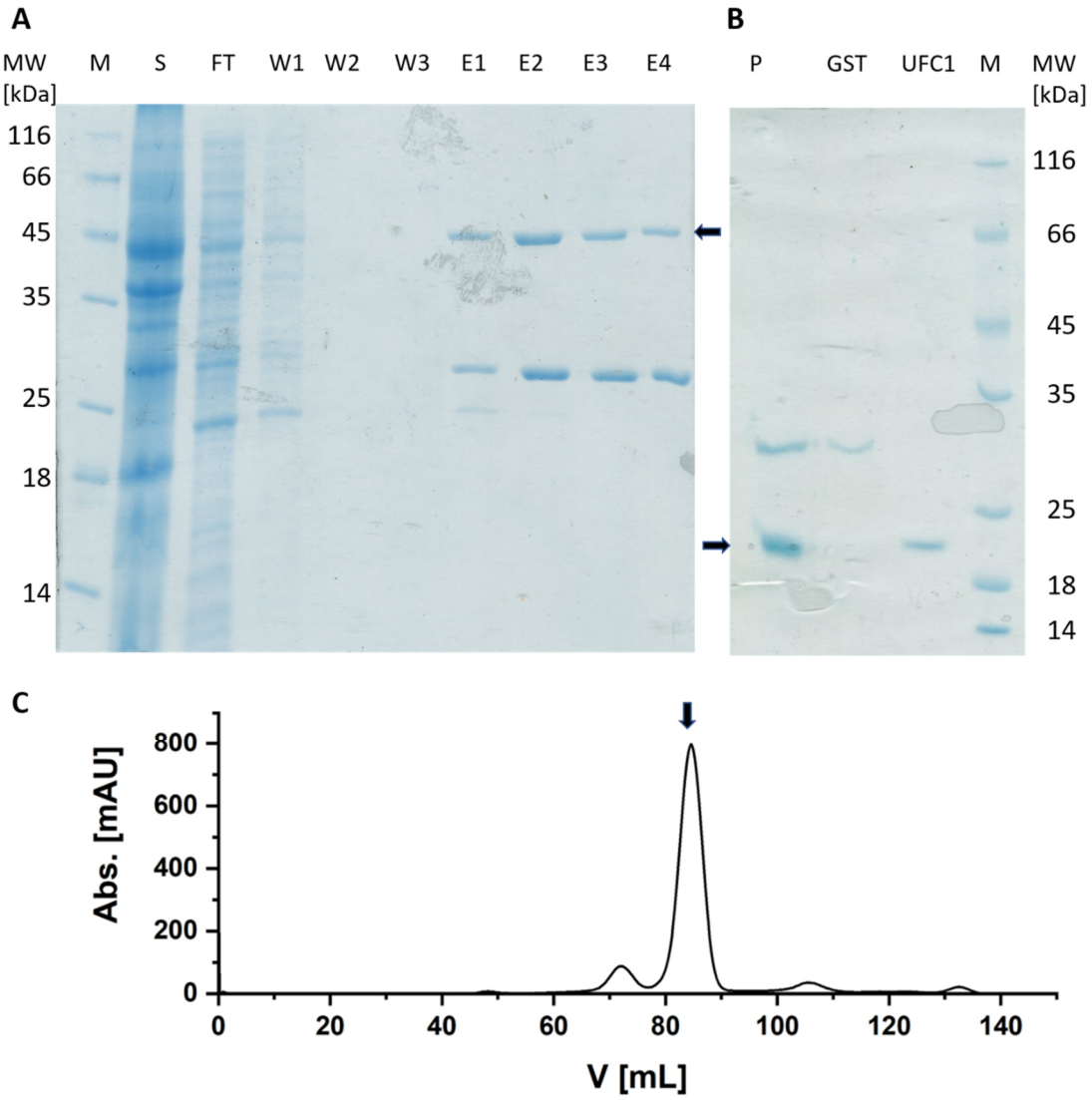


Fig. 2-8 Purification of UFC1. Top left: Gel shows the supernatant (S) of the centrifugation step (s. 4.2.3.1), the flow-through from the affinity chromatography (FT), the washing steps (W1-W3) and the first four elution steps (E1-E4). The UFC1-GST protein band is marked with a black arrow. The additional band at 26 kDa comes from free GST. Top right: Gel shows the pooled fractions applied to the SEC (P) and eluted GST, and the final purified UFC1. The UFC1 protein band is marked with a black arrow. Molecular weights of the marker bands (M) are also displayed. Bottom: Chromatogram of the SEC purification by plotting the absorption in mAU over the elution volume. The UFM1 peak is marked with a black arrow.

To check for oligomeric and folding the state the sample was then analyzed with DLS and CD spectroscopy (see sections 4.2.4.2 and 4.2.4.3; **Fig. *2-9***). DLS shows at a

concentration of 2.5mg/mL a clear monodisperse peak at 2.3 nm, which corresponds to a molecular weight of 22 kDa. The secondary structure estimation differs from the sequence annotation of the secondary structure in the pdb, but the CD spectra shows a properly folded protein (**Tab. *2-3*** and **Fig. S *5***). The ratio of α-helix is only half of what is found in the crystal structure and the amount of β-sheet is slightly higher. Due to high buffer absorption in the low UV regime, the secondary structure could only be annotated to the spectra between 200 and 260 nm. Around 210 nm, there is also variation in the fit and the measured spectra (s. **Fig. S *5***). This probably leads to the variation of the annotated values from the PDB and the expected values.

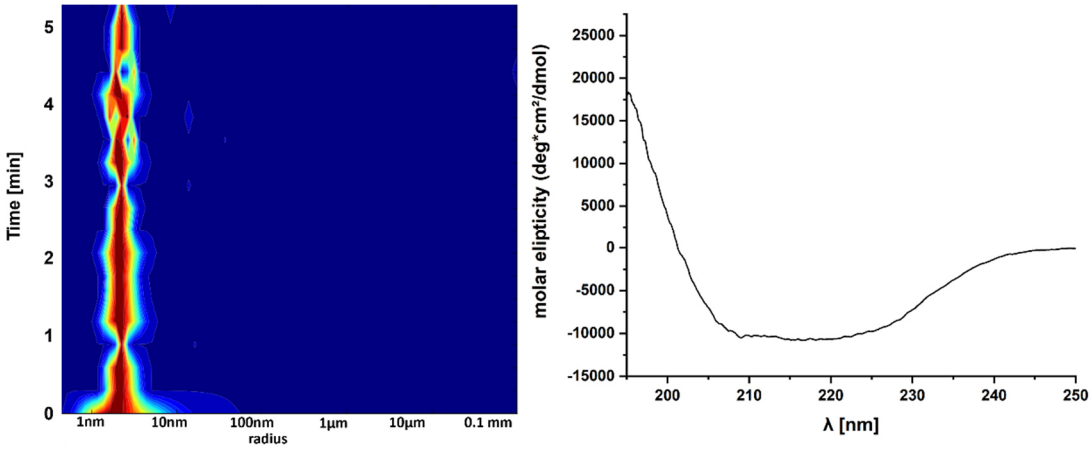


Fig. 2-9 DLS and CD spectra of purified UFC1. Left: DLS of purified UFC1 measured at 2.5 mg/mL shows a clear monodisperse peak with a radius of 2.3 nm but with variation in size distribution. Right: CD spectra of UFM1 in molar ellipticity measured between 195 and 260 nm.

Tab. 2-3 Calculated secondary structure of UFC1 from the CD spectrum with the BeStSel server in comparison with the sequence annotation deposited in the pdb.

Protein	**α-helix [%]**	**β-sheet [%]**
UFC1	18.7	26.6
UFC1 (2Z6O)	37.0	18.0

2.2 Biophysical Characterization of UBA5 with bound NPE-ATP

2.2.1 Exchange of ATP with NPE-ATP

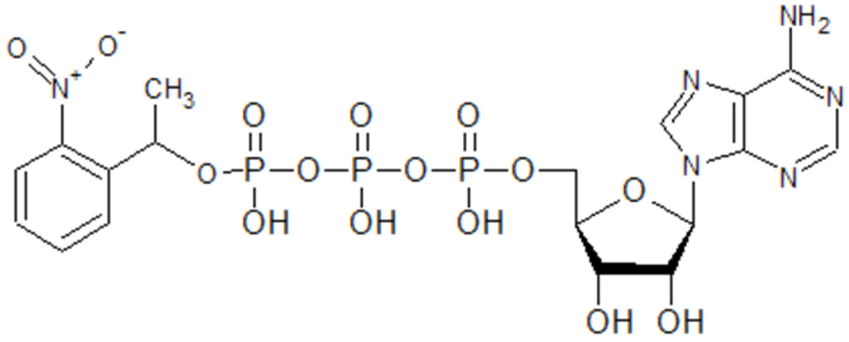


Fig. 2-10 Structure of caged ATP with a nitrophenyl ester as the caging group

Three different methods were tested to exchange ATP with NPE-ATP in UBA5 which are described in the literature [141–143]. Two of these methods rely on the magnesium-dependent affinity of nucleotides to proteins. By removing magnesium by addition of EDTA the affinity of ATP can be lowered by addition of an excess of NPE-ATP. Method A additionally includes $(NH4)_2SO_4$ to improve the affinity lowering with a slight salting-out effect. The third method is based on the affinity variations of different nucleotides (AMP, ADP, ATP) to the protein. Here, the holoenzyme is incubated with a phosphatase to hydrolyze ATP to AMP or ADP. These have a lower affinity, which can be used to exchange them with addition of an excess of NPE-ATP. The exchange rate was determined by applying the samples on a RP-HPLC column as described in section 4.2.5.3 (**Fig. *2-13***). Furthermore, different ATP and NPE-ATP concentrations were measured for better determination of the nucleotide concentration in the samples. This was used to determine the loading ratio of both nucleotides in the samples (**Tab. *2-5***).

First of all, the loading status of UBA5 with ATP was checked with RP-HPLC. One sample was purified with addition of 2 mM ATP and 5 mM $MgCl_2$ in the SEC buffer to saturate UBA5 with ATP. Subsequently two samples were prepared from purification of UBA5 without addition of ATP and $MgCl_2$. Hereby, the two different SEC elution peaks (s. **Fig. *2-2***) were analyzed separately. It was found that without ATP addition during purification, a small amount of ATP is still bound to the protein (**Fig. *2-12*** and **Tab. *2-4***)

With calibration of pure ATP and NPE-ATP the amount of bound nucleotide could be estimated (**Fig. *2-11***). ATP elutes around 9 minutes and NPE-ATP elutes around 12 minutes.

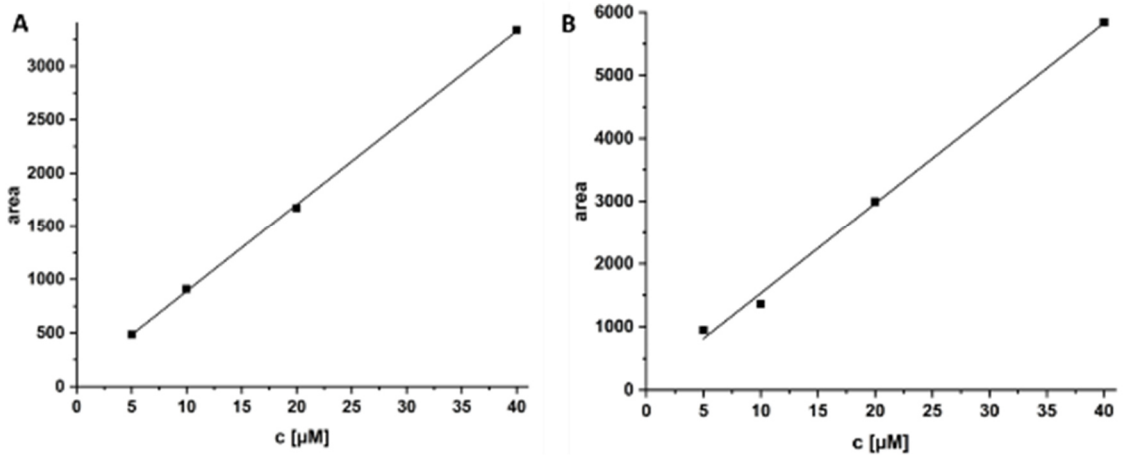


Fig. 2-11 Calibration curves of A: ATP and B: NPE-ATP.

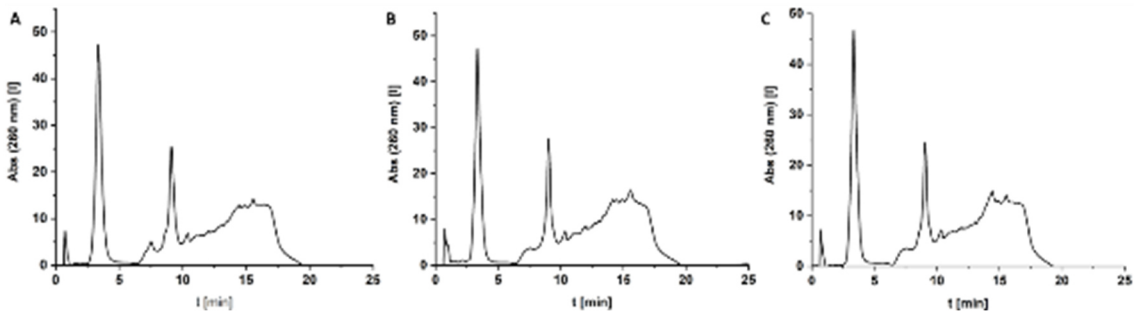


Fig. 2-12 Loading status of UBA5 with ATP. A: UBA5 after purification with ATP additive; B: UBA5 after purification without ATP, first SEC elution peak; C: UBA5 after purification without ATP, second SEC elution peak.

Tab. 2-4 Calculated loading of UBA5 with ATP by forming the ratio of ATP and protein concentration. UBA has as a dimer two binding sites. This is implemented in the calculation by doubling the protein concentration.

Sample	**protein concentration [µM]**	**ATP concentration [µM]**	**loading [%]**
UBA5 + ATP	3.7	6.5	87.8
UBA5 – ATP first peak	17.7	5.9	16.7
UBA5 – ATP second peak	12.9	5.4	20.9

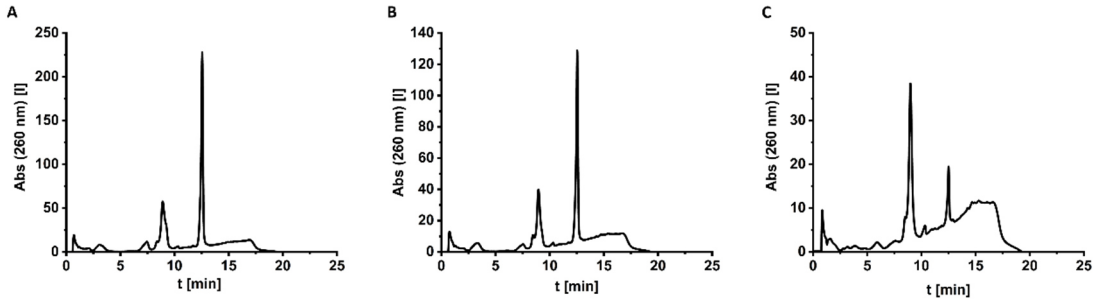


Fig. 2-13 Comparison of different elution profiles of ATP exchange methods.

Tab. 2-5 Measured nucleotide concentrations in the different nucleotide exchange methods and their ratios.

Method	ATP [µM]	NPE-ATP [µM]	ATP [%]	NPE-ATP [%]
A	22.7	19.5	53.8	46.2
B	12.7	11.7	52.0	48.0
C	10.5	2.3	82.0	18.0

Incubation with Method C shows that only small amounts of ATP were exchanged, while methods A and B show similar rates, but with a slightly higher amount of NPE-ATP for method B. The inability to successfully cleave the phosphate groups by the phosphatase comes probably from missing interactions between enzyme and nucleotide due to steric reasons. Therefore, this method was not selected as promising method to optimize the exchange of ATP. Hence Method B shows the highest exchange rates, it was used for further optimization by changing the parameters incubation time, excess of NPE-ATP, and dialysis during incubation (***Fig. 2-14***).

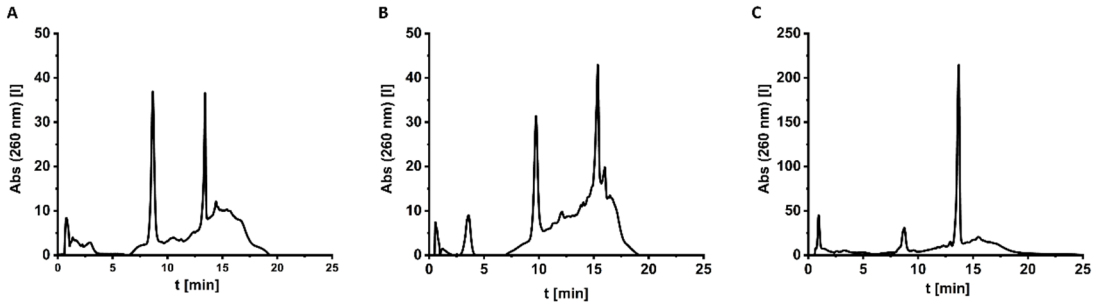


Fig. 2-14 Optimization of ATP exchange with NPE-ATP.The biggest influence in optimization is the incubation of the sample during dialysis against buffer to remove the ATP in the sample. Altering the incubation time or excess of NPE-ATP did not improve the exchange rate significantly. From the area and its respective concentration of both peaks, the ratio of both nucleotides was calculated.

Tab. 2-6 Measured nucleotide concentrations in the different optimization trials and their ratios.

Method	ATP [µM]	NPE-ATP [µM]	ATP [%]	ATP [%]
2 h	13.1	10.4	55.7	44.3
50x excess	7.6	3.6	67.8	32.2
dialysis	6.4	70.4	8.3	91.7

After optimizing the exchange reaction by using Method B and dialyze during EDTA incubation, DLS and CD spectroscopy experiments were performed to observe if the additives influence the dispersity and folding state of the protein. First, the dispersity after the exchange reaction was compared to UBA5 before the exchange (**Fig. *2-15*** A+B). The radius increased after the first dialysis and did not change after the second. Additionally, the peak showed a change from mono- to polydispersity. In order to analyze at what step this change occurs. DLS was then measured (**Fig. *2-15*** C+D). These experiments show that the protein changed after the first dialysis step. Therefore, each of the additives (ATP, EDTA, $MgCl_2$, $ZnCl_2$) were tested separately to examine which of these influence UBA5 (**Fig. *2-15*** E-H). These tests revealed that after 1 h none of these changed the polydispersity. Nevertheless, it was expected that EDTA might have a time-dependent influence by removing magnesium and zinc ions. DLS experiments showed that EDTA in fact influences the polydispersity of UBA5 after several hours (**Fig. *2-16*** A). Comparison of the DLS results after EDTA incubation and after the first dialysis revealed a similar behavior.

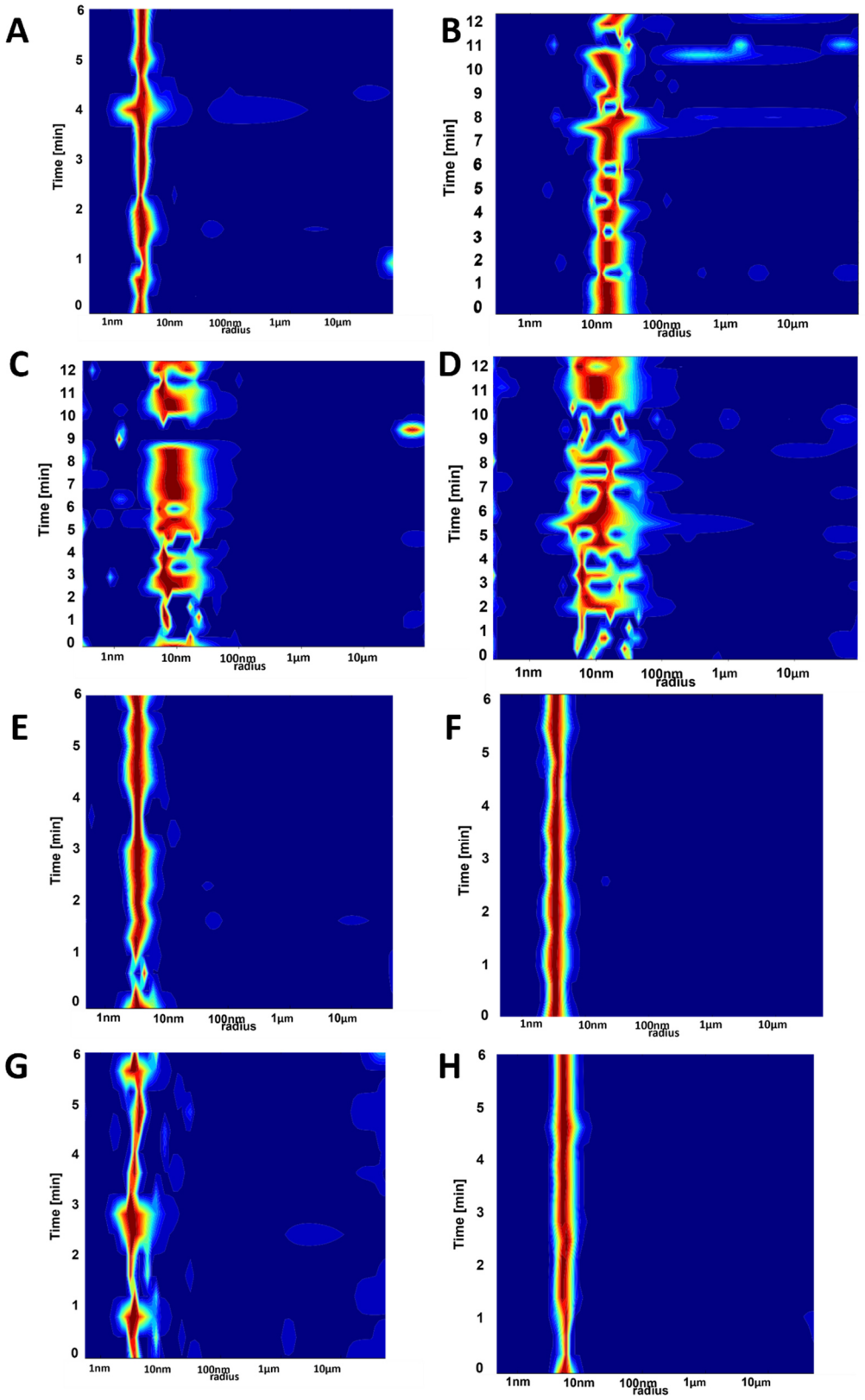


Fig. 2-15 Summary of DLS results from ATP exchange experiments. A: UBA5 before exchange; B: UBA5 after exchange; C. UBA5 after first dialysis; D: UBA5 second dialysis; E: UBA5 with ATP incubation; F: UBA5 after EDTA incubation; G: UBA5 after $MgCl_2$ incubation; H: UBA5 after $ZnCl_2$ incubation

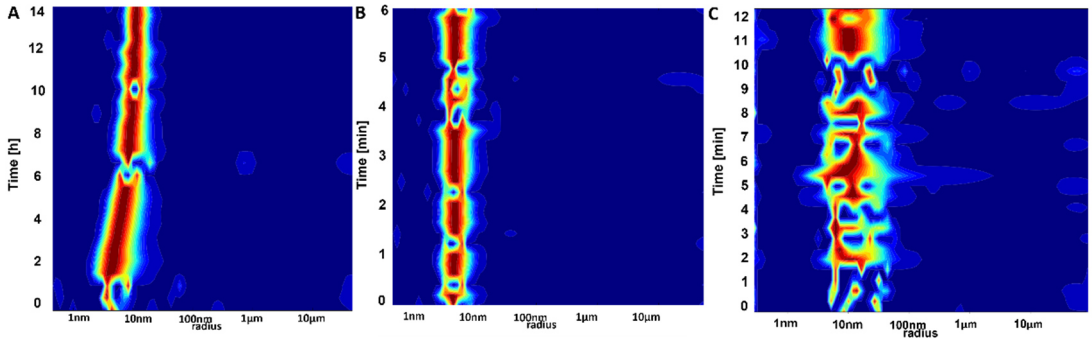


Fig. 2-16 Summary of DLS results from experiments with EDTA. A: UBA5 with EDTA measured over 14 h; B: UBA5 after incubation with EDTA; C: UBA5 after first dialysis.

Further investigations involved analysis of the samples with CD spectroscopy. Here, the additives did not affect the folding state except addition of $ZnCl_2$. With AAS, the effect of EDTA on the magnesium and zinc ions was studied. It was found that EDTA has a small effect on zinc ions by reducing their concentration bound to UBA5.

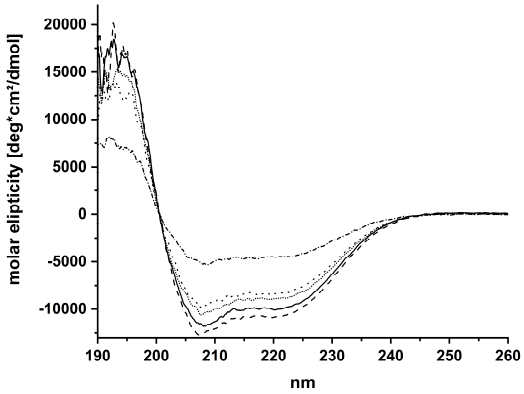


Fig. 2-17 CD spectra of UBA5 with different additives. Black line: UBA5; long-dashed line: UBA5 + EDTA; short-dashed line: UBA5 + ATP; dotted line: UBA5 + $MgCl_2$; dashed-dotted line: UBA5 + $ZnCl_2$.

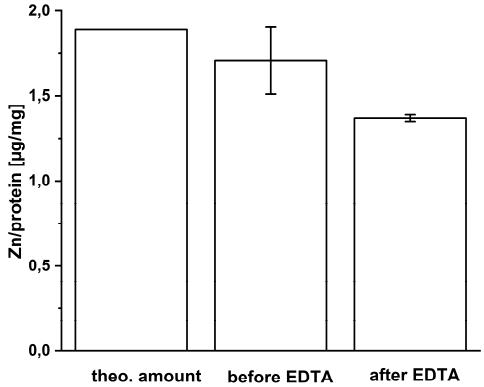


Fig. 2-18 Measured AAS Zn concentrations in µg per mg protein measured before and after EDTA treatment and compared to the theoretical expected value.

These experiments proved to be successful in order to optimize the exchange rate and a loading rate from ATP to NPE-ATP to over 90 %. The folding status was also not influenced by the different additives as shown by CD spectroscopy. Also, the addition of EDTA showed signs of influencing the zinc binding site in addition to the binding of the magnesium ions. However, the influence is only small, since the concentration of zinc was only lowered by around 15 % after incubation with EDTA (s. **Fig. *2-18***). DLS shows that EDTA has an influence on the dispersity of UBA5. However, if the incubation time is kept low but long enough to exchange the nucleotides (~ 4 h), the protein keeps its monodisperse behavior and can be used for crystallization trials.

2.2.2 Crystallization trials of UBA5 with NPE-ATP

Different methods with the two different UBA5 constructs were applied in order to find crystallization conditions of UBA5 with exchanged NPE-ATP. The first trials involved UBA5 57-329 with the optimized crystallization conditions shown (see section 2.2.4): Protein concentration 8 – 10 mg/mL, 0.8 – 1.2 M $LiSO_4$, 0.3 M $(NH_4)_2SO_4$, 0.1 M Na-citrate pH 5.9- 6.2. The crystallization was not successful, the crystallization drops developed only precipitate. Additionally, for an optimal screening, commercially available sparse matrix crystallization screens were applied over a concentration range of 4 to 16 mg/mL and at different temperatures (4 °C, 16 °C and 20 °C). Furthermore, the screening involved the setup of crystallization trials with the vapor diffusion, microbatch under oil method and counter diffusion method.

Complex studies with CD spectroscopy showed that UBA5 is stabilized by its binding partner UFM1 (**Fig. *2-19***). Therefore UBA5 57-346 in complex with UFM1 was also used for crystallization trials. The different approaches included screening around the crystallization conditions described by Oweis et al [58] and the screens mentioned above with vapor diffusion and microbatch under oil setups and variation of parameters protein concentration and temperature (similar to the condition variation mentioned above). Nevertheless, crystallization trials were unsuccessful.

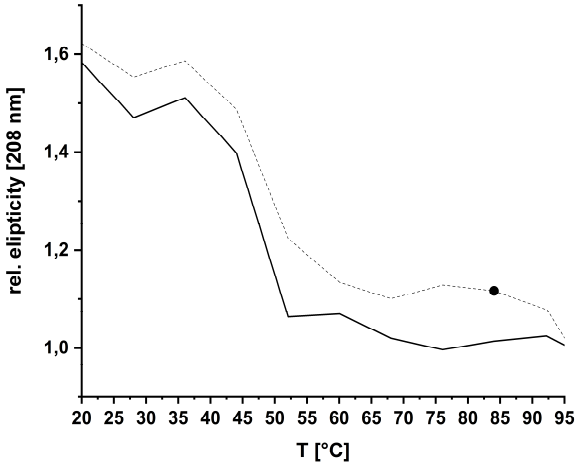


Fig. 2-19 Melting curves of UBA5 (black line) and UBA5-UFM1 measured with CD spectroscopy at 208 nm.

2.2.3 Docking studies of caged NPE-ATP to UBA5

In order to test if NPE-ATP with additional space requirements of nitrophenyl group can fit into the binding pocket of ATP, two different docking software packages were applied, as it was also necessary to find a proper conformation of the NPE-ATP molecule and to compare the consistence of the results. SwissDock, which is an online docking tool, requires the pdb file of the protein model and the mol2 file of the ligand. It generates different models with different possible ligand binding positions. Among the different models, one shows NPE-ATP in the ATP binding site with enough space for the nitrophenyl group (**Fig. *2-20***). Nevertheless, NPE-ATP adopts a different conformation with the third phenyl group with the additional nitrophenyl group turned 180 ° compared to ATP. In another, complementary approach the SYBYL software was used to get a deeper insight on the fit of NPE-ATP to the binding site (**Fig. *2-21***). The results are similar. NPE-ATP fits with its adenine part in the ATP pocket with a similar conformation. The ribose, phosphates and protecting group extend to the protein surface with slightly different conformation than ATP.

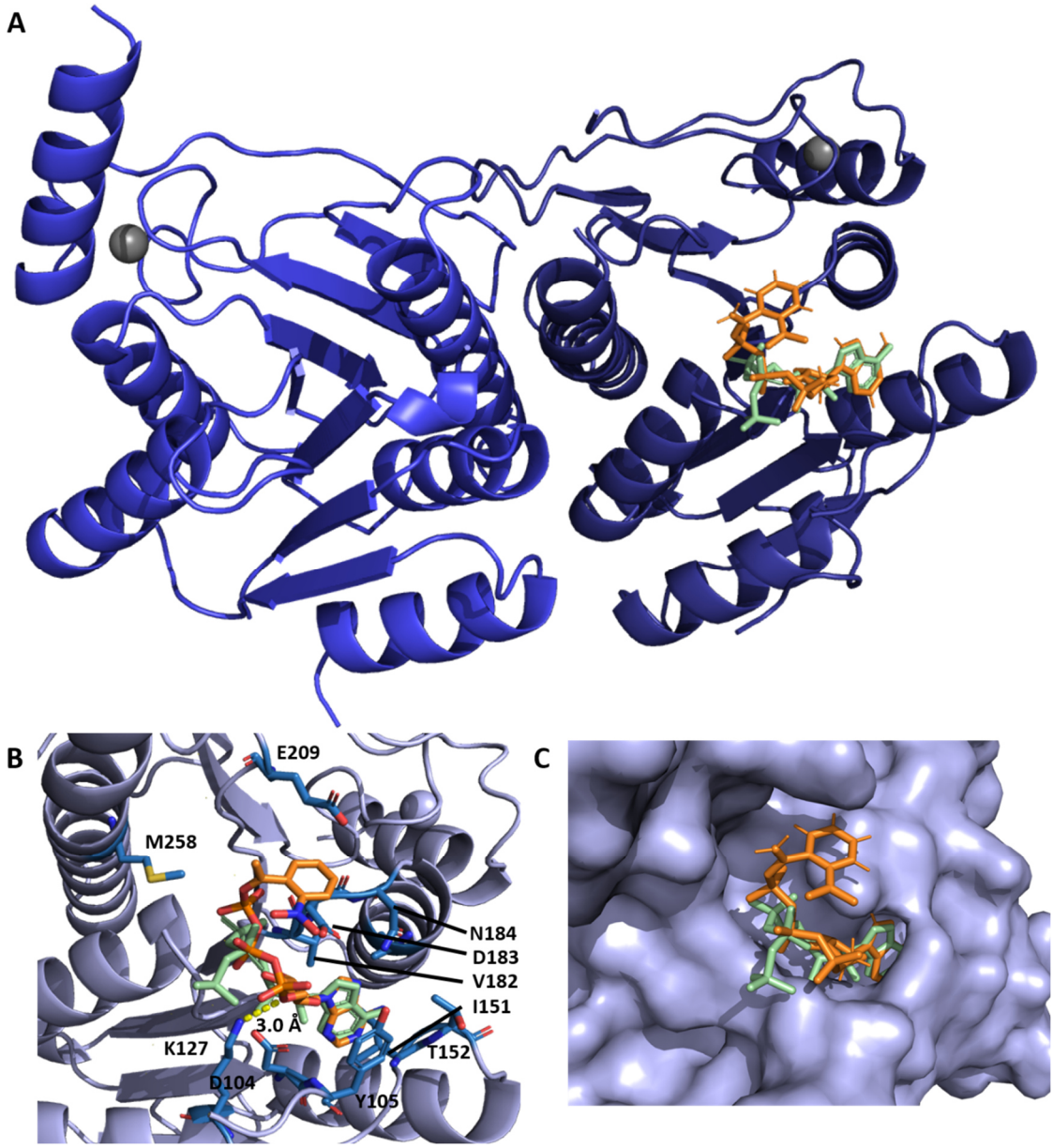


Fig. 2-20 Depiction of SwissDock model of UBA5 57-329 with NPE-ATP in ATP binding pocket. ATP is colored in pale green and NPE-ATP in orange. A: UBA5 dimer with ATP and NPE-ATP shown as stick models. B: Detailed view of ATP binding pocket with ATP and NPE-ATP as stick models, as well UBA5 residues with hydrophobic and H-bond interactions as stick models. C: Detailed view of ATP binding pocket displayed as surface structure with ATP and NPE-ATP as stick models.

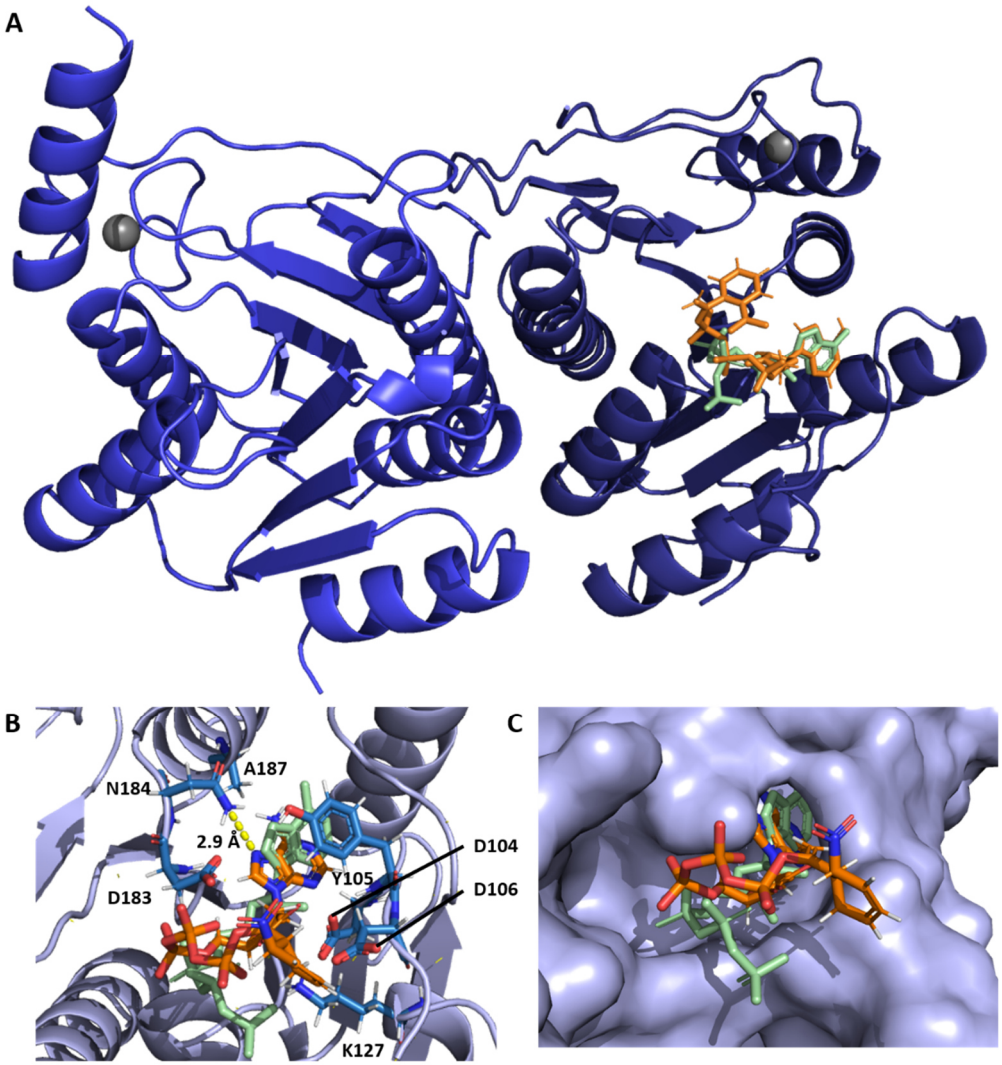


Fig. 2-21 Depiction of SYBYL model of UBA5 57-329 with NPE-ATP in ATP binding pocket. ATP is colored in pale green and NPE-ATP in orange. A: UBA5 dimer with ATP and NPE-ATP shown as stick models. B: Detailed view of ATP binding pocket with ATP and NPE-ATP as stick models as well UBA5 residues with hydrophobic and H-bond interactions as stick models. C: Detailed view of ATP binding pocket displayed as surface structure with ATP and NPE-ATP as stick models.

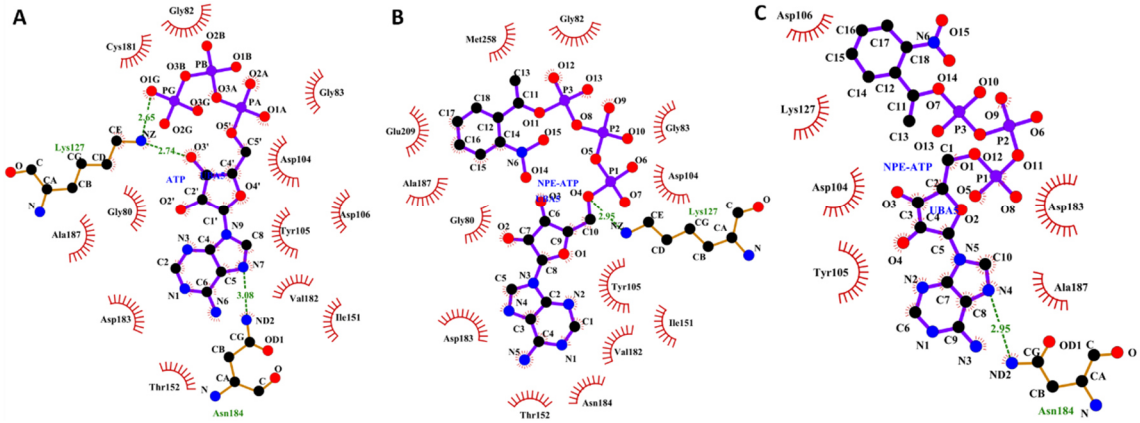


Fig. 2-22 Comparison of the nucleotide binding pocket of UBA5 with interacting residues generated by LigPlot.Hydrophobic interactions with amino acid residues are depicted with a red, spiked circle, H-bond interaction are depicted with a green dotted line and the distance in Å between the atoms of the nucleotide and the residue. A: Binding pocket of UBA5 with ATP; B: binding pocket of UBA5 with NPE-ATP found by SwissDock; C: binding pocket of UBA5 with NPE found by Sybl.

Tab. 2-7 Summary of interactions (H-bonds and hydrophobic) between nucleotides found by LigPlot between UBA5 and ATP and NPE-ATP from docking results. (A) and (D) mark acceptor and donor residues in H-bond interactions. Atoms in bracket mark the hydrophobic interaction partner of the nucleotide.

Type of interaction	**UBA5 with ATP**	**SwissDock**	**SYBYL**
H-bond	K127 (A)		
	N184 (D)		N184 (D)
Hydrophobic	D104 (C5')		
	Y105 (adenosine ring)		
	D183 (C5')		
	A187 (N6)		
	G80 (ribose)		
	G82 (phosphate groups)		
	G83 (phosphate groups)		
	I151 (N6)		
	V182 (adenosine ring)		
	D106 (phosphate groups)		D106 (nitrophenyl group)
	T152 (N6)		
	C181 (C5')		
		N184 (adenosine ring)	
		E209 (nitrophenyl group)	
		M258 (phosphate group of PG	
			K127 (ribose group)

If the results of the docking of NPE-ATP into the ATP binding pocket is visualized with the plotting software LigPlot and compared with the native interactions of ATP and UBA5 (**Fig. *2-22***, **Tab. *2-7***), these similarities and differences are found: 2 different H-bonds are identified (K127, N184), both are present in the native UBA5 model with ATP but only one appears either in the SwissDock (K127) or the SYBYL model (N184). Additionally, the H-bonds change. In the native structure the H-Bond of K127 consists between the amine group as acceptor and the oxygen atoms of O3' and OG1 as donors.

The SwissDock model shows interactions between the amine group (acceptor) and O5'. The other residues interact hydrophobically with the nucleotide: N184 in the SwissDock model with adenosine ring of the nucleotide and K127 with the ribose ring in the SYBYL model. For all models, LigPlot finds hydrophobic interactions between the nucleotides and the residues D104 (interacting with C5'), Y105 (adenosine ring), D183 (C5') and A187 (N6). In the SwissDock model, LigPlot identifies additional hydrophobic interactions between NPE-ATP and the glycine residues G80 interacting with the ribose ring, G82 and G83 interacting with the phosphate groups, as well as between NPE-ATP and I151 and V182, which are also present in the native model. The aspartate residue D106 interacts with nitrophenyl group the SYBYL model while interactions in the native model can be found between D106 and the phosphate groups. The T152 and C181 residue does neither interact with NPE-ATP in the SwissDock nor the SYBYL model. Furthermore, in the SwissDock model, two additional residues are identified making hydrophobic interactions with NPE-ATP, which are E209 and M258 interacting with the nitrophenyl group and the gamma phosphate group, respectively. These come probably from the different conformation of the third phosphate group and the nitrophenyl group.

Both programs calculated a ΔG score (s. **Tab. *2-8***) to evaluate the quality of the fit. SwissDock calculated ΔG with – 9 kcal/mol, which is in the range of what is typical found by this program [144]. SYBYL calculated a much lower value of – 704 kcal/mol, but this calculations is performed with a higher variety of parameters, e.g. rotatable bonds, evaluation of different binding modes or distance to important residues [145] Furthermore, it also finds less possible interactions than SwissDock. However, both results are below zero, meaning that the binding at the binding site is favored. However due to less interactions, meaning fewer H-Bonds and numbers of hydrophobic interactions, the binding of caged ATP is weaker than that of the native ATP. However, accurate binding energies of ATP to UBA5 have not been evaluated accurately in the literature so far.

Tab. 2-8 ΔG score calculated by the docking programs in kcal/mol and kJ/mol.

program	**ΔG [kcal/mol]**	**ΔG [kJ/mol]**
SwissDock	- 8,96	- 37,49
SYBYL	- 703.78	- 2944.62

To calculate the association constant from the free enthalpy, this formula is used:

$$\Delta G = RT * \ln \frac{K_d}{c^\theta}$$

Eq. 20 Calculation of ΔG from the dissociation constant with R as the ideal gas constant (8.314 $J*K^{-1}*mol^{-1}$), T as the temperature in K and c^θ as the standard reference concentration (1 mol/L).

The formula is rearranged to calculate K_d from ΔG:

$$K_d = c^\theta * e^{\frac{\Delta G}{RT}}$$

program	**K_d [µM]**
SwissDock	0.20
SYBYL	$1.07*10^{-515}$

The unrealistic low value calculated by SYBYL is further explained in section 2.2.5

2.2.4 Comparative crystallization studies of UBA5 in complex with ATP and NPE-ATP

In order to check the crystallization reproducibility of the native protein the native form with ATP was crystallized. UBA5 57-329 could be crystallized by varying the method described by Bacik et al (1 M $LiSO_4$, 0.3 M $(NH_4)_2SO_4$, 0.1 Na-citrate pH 6.2). However, in order to crystallize the construct some optimizations had to be done: the concentration was lowered from 20 mg/mL to 8 mg/ml. Additionally, seeding with crystalline precipitate from previous crystallization trials was applied. Despite these alterations in crystallization conditions, the biggest influence in crystallization success was the change of cell disruption method from sonication to freeze-thaw procedure and addition of fresh DTT in every purification step. This suggests that UBA5 is a sensitive protein, where the cell disruption method can influence the conformation and the aggregation probability [146]. Blimp shaped crystals appeared after 4 d and grew to a size of 200 x 50 x 50 µm.

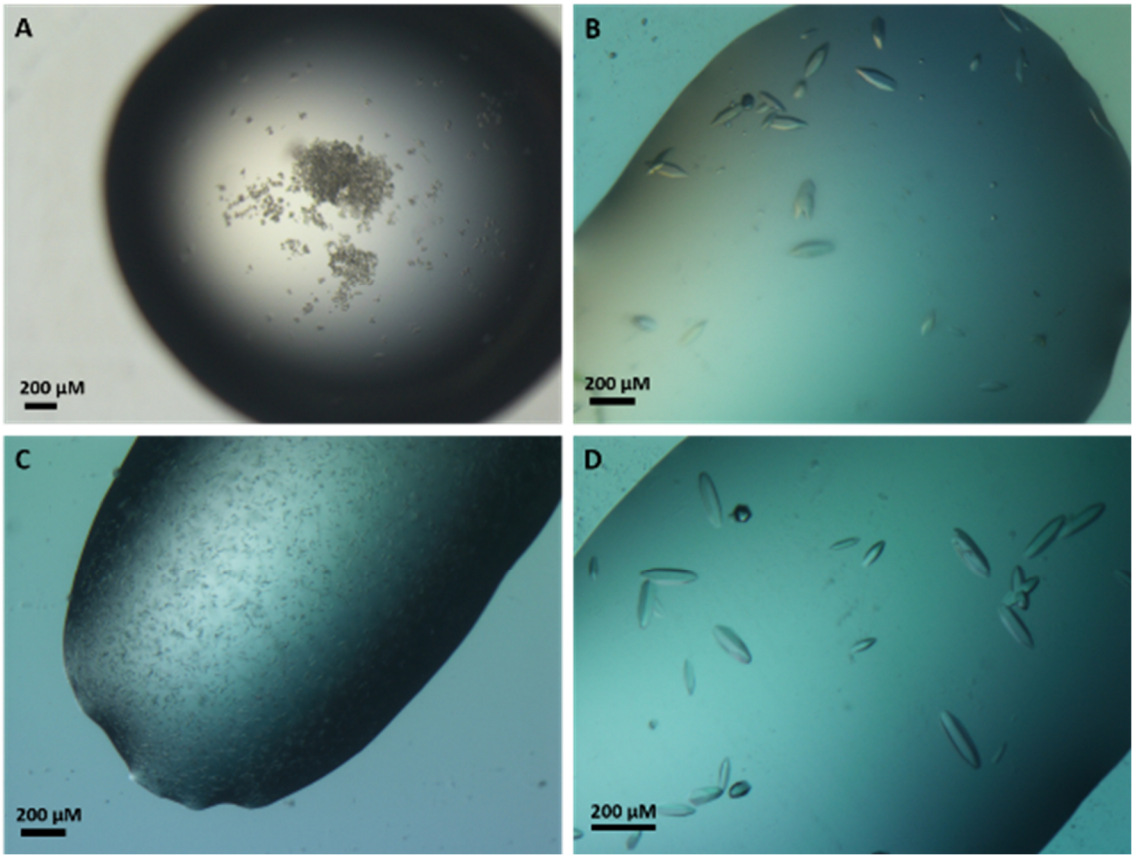


Fig. 2-23 Crystals of UBA5 57-329 (8.0 mg/mL) crystallized under different conditions with scale bar. A: 0.9 M $LiSO_4$; 0.3 M $(NH_4)_2SO_4$, 0.1 M Na-citrate pH 5.9; B: 0.8 M $LiSO_4$, 0.3 M $(NH_4)_2SO4$, 0.1 M Na-citrate pH 5.9; C: 1.2 M $LiSO_4$, 0.3 M $(NH_4)_2SO4$, 0.1 M Na-citrate pH 6.2 D: 1.0 M $LiSO_4$, 0.3 M $(NH_4)_2SO4$, 0.1 M Na-citrate pH 5.9. Precipitant in A was used to prepare the seed stock. B and C were crystallized by the addition of 1 µL seed stock; D by seed streeking.

Crystals used for diffraction experiments diffracted up to 2.2 Å. From the data, the structure could be solved with XDS and the CCP4 software package. Data statistics and comparisons to the published structure are depicted in **Tab. *2-9***. The comparison shows lower quality of the parameters mostly due to the lower resolution which was achieved. Especially the R factors of the data collection are higher. However, the refinement statistics are similar, as well as the space group and the cell parameters. The missing number atoms in the protein result from a valine residue (V108) which could not be modeled into the structure due to missing electron density.

Tab. 2-9 Data collection and refinement statistics of UBA5 57-329 crystal used for model building compared with statistics from crystal structure by Bacik et al. [56] (pdb code: 3H8V). Numbers in brackets show values of the highest resolution shell

Data collection		
model	own	Bacik et al.
X-ray source	PETRAIII, Hamburg, EMBL P14	APS, 24-ID-C
detector	EIGER 16M	ADSC QUANTUM 315
space group	$P3_221$	$P3_221$
cell dimensions		
a, b, c [Å]	77.5, 77.5, 205.7	78.0, 78.0, 207.0
α, β, γ [°]	90, 90, 120	90, 90, 120
wavelength [Å]	0.9763	0.97944
resolution [Å]	19.96 – 2.26 (2.46 – 2.26)	50.00-2.00 (2.00 – 2.03)
total reflections	741040 (120626)	177115 (n.a.)
total unique reflections	37310 (5908)	45414 (n.a.)
redundancy	19.9 (20.4)	3.9 (3.8)
Wilson B-factor [Å²]	19.0	31.1
R_{syms}	0.102 (1.493)	0.080 (0.859)
R_{meas}	0.105 (1.531)	n.a.
R_{pim}	0.108 (1.936)	0.045 (0.475)
$CC_{1/2}$ [%]	100.0 (81.3)	n.a.
completeness [%]	100.0 (99.9)	90.1 (93.3)
I/σI	19.10 (2.10)	16.75 (1.65)
Refinement		
resolution	19.1 – 2.2	25.5 – 2.0
reflections used	35351	42909
reflections used for R_{free}	1768	2372
R_{work}	0.1930	0.1910
R_{free}	0.2420	0.2240
number of atoms	3562	3615
protein	3415	3425
ligand/ions	33	33
water	114	157
rms deviations		
bond length [Å]	0.011	0.011
bond angle [°]	1.951	1.251
mean B value [Å²]	58.05	68.49
Ramachandran		
favored [%]	94.8	99.1
allowed [%]	4.3	0.9
outliers [%]	0.9	0

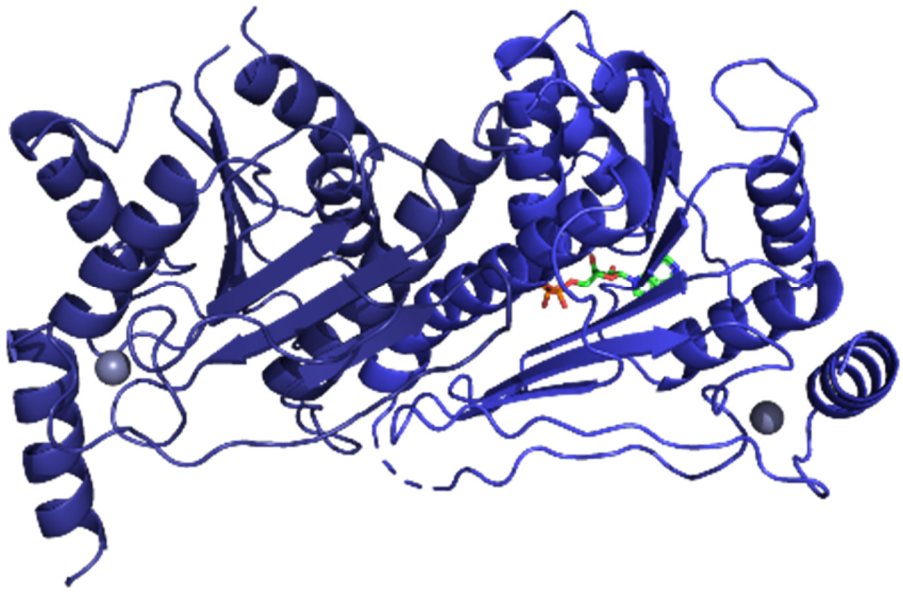


Fig. 2-24 Ribbon model of the crystal structure model of UBA5 57-329 with a resolution of 2.2 Å. Dotted line represents missing residues between N236 and R246.

Comparison with the published UBA5 structure (pdb id: 3H8V) shows only marginal differences and slightly lower quality of statistics (**Tab. *2-9***). The rsmd of the Cα positions between both structures is 0.198 Å. The biggest differences are found in on loop between residues D106 and L125 (Cα = 3.389 Å). In chain A, more residues could be modeled into the electron density of the published UBA5 structure in comparison to our model (**Fig. *2-25***). In chain B the beginning and the end of the loop has a slightly different conformation. Some residues adopt similar conformations, e.g. Y105 at ATP binding site.

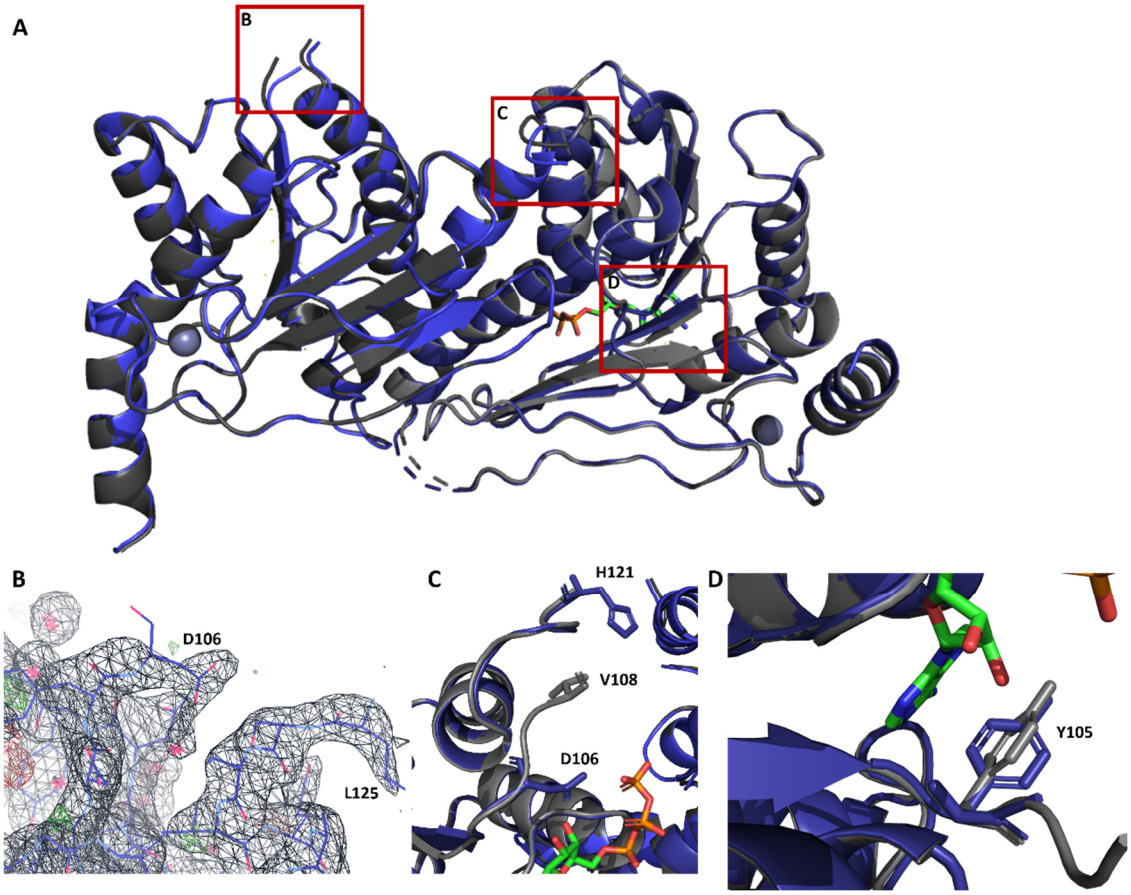


Fig. 2-25 Comparison of UBA5 57-329 models. Own model created from diffraction data is in blue; model by Bacik et al. is in dark grey. Red squares show the zooms B-D. A: Dimer models of UBA5 57-329 show a very similar fold. Also depicted are two zinc ions displayed as spheres and the ATP as a stick model. B: Detailed view of difference in the loop between residues D106 and L125 in chain B with missing electron density between them. C: Detailed view of difference in the loop between residues V108 and H121 in chain A. In the 3H8V model more residues could be modeled into the electron density. D: Detailed view of Y105 which can adopt different conformations.

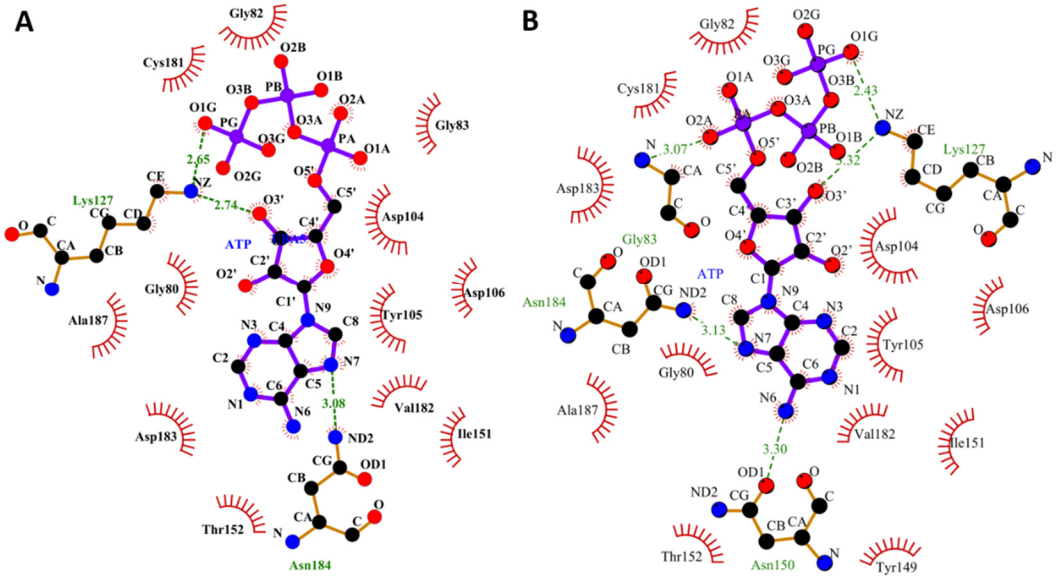


Fig. 2-26 Comparison of ATP binding site in UBA5 with involved residues identified by LigPlot. A: binding site of the structural model by Bacik at al.; B: binding site of own structural model.

In comparison of the binding site of both models (Bacik at al. and own model), the involved residues identified by LigPlot (**Fig. *2-26***) are nearly the same. Some small differences found by LigPlot are that additional H-Bonds between N150 and the N6 atom of ATP and an additional hydrophobic interaction with Y149 are found. Furthermore, in the model of Bacik et al. the G83 residue makes a hydrophobic interaction with ATP, while in the other model a H-bond is found between G83 and one oxygen atom of the first phosphate group.

As discussed in 3.1, the experiments show that although NPE-ATP can bind and does fit in the binding pocket, the preparation of crystal of UBA5 containing NPE-ATP was not possible.

2.2.5 Measurement of Affinity

The binding of NPE-ATP to UBA5 is determined and compared to the binding affinity of ATP. The experiments show a difference in signal generation. While the binding of ATP quenches the tryptophan fluorescence signal, the binding of NPE-ATP enhances it. The Hill fits of both nucleotides reveal some differences (**Fig. *2-27*** and **Tab. *2-10***). The application of the hill formula shows no measurable cooperation of ATP, but again reveals a negative cooperativity of around 0.5 of NPE-ATP. If this has a relevant biological effect is currently not known. Additionally, due to a higher variation of the average fluorescence intensity, the fit for the NPE-ATP is not as accurate as the fit for the ATP measurement; the cooperativity estimation therefore has also a higher uncertainty. Furthermore, the dissociation constant of ATP is around 0.8 µM and 3.4 for NPE-ATP. The difference can be estimated by the slopes of the hill fits. The curve of ATP is steeper, meaning that the affinity of the nucleotide is higher resulting in a lower K_D, while the curve of NPE-AT is less steep.

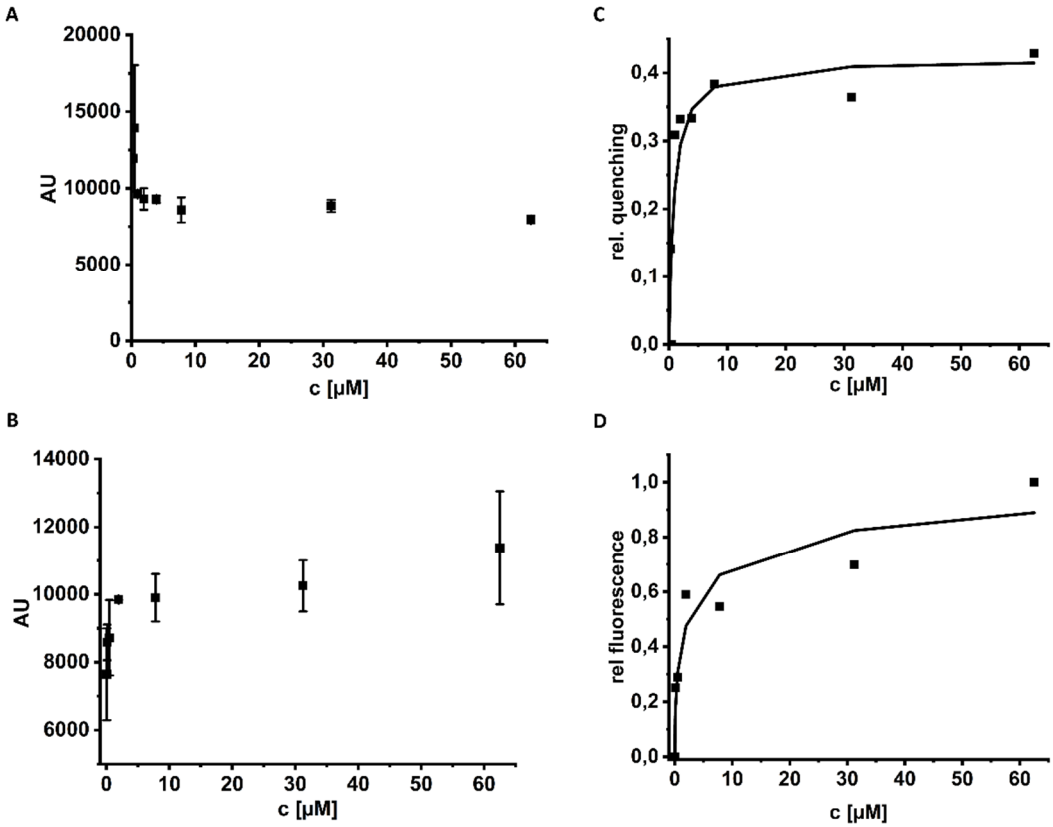


Fig. 2-27 Fluorescence measurements of binding affinity of nucleotides to UBA5. A: Fluorescence quenching of UBA5 at different ATP concentrations with error bars. B: Relative fluorescence quenching of UBA5 at different ATP concentrations with Hill fit. C: Fluorescence intensity of UBA5 at different NPE-ATP concentrations with error bars. D: Relative fluorescence intensity of UBA5 at different NPE concentrations with hill fit.

Tab. 2-10 Calculated parameters from the hill fit for the affinity measurement of ATP and NPE-ATP to UBA5

Parameter	ATP	NPE-ATP
K_d [µM]	0.83	3.41
B_{max}	0.42	1.10
n	1.00	0.49

The calculated K_D for ATP is in the same regime at it was found by the docking software SwissDock, which finds a value of 0.2 µM. The value determined by Sybyl is much lower and is nearly 0 and therefore unrealistic. The numerical value of the dissociation constant determined by fluorescence spectroscopy is similar to what was found by Mashahreh at al [59]. However, the values differ in their order of magnitude. The fluorescence spectroscopy determined a K_D of around 830 nM, while Mashahreh at al. found a K_D of ~ 730 µM.

2.2.6 Stability of NPE-ATP

Lastly, the stability of NPE-ATP against light illumination was checked to investigate if the NPE-group is not released during the experimental setups e.g. preparation of buffers, crystallization or crystal mounting before subsequent diffraction experiments. This was done to determine if special precautions are necessary when handling the compound during the experiments, e.g., work only with red light illumination. The measurements of ATP and NPE-ATP and dark and light conditions show that the stability and decaging of NPE-ATP is not much affected by visible day light illumination (**Fig. *2-28***). The peak at 17 min shows no significant difference over time. For the ATP measurement, different peaks are visible with the main ATP peak at 12 min. The peak at 15.5 min could come from ADP due to ATP hydrolyzation, the small peak at 17 min from AMP [147]. Two peaks appear also in the caged-ATP runs. On peak around 15 min and on small peak around 12 min, matching the ATP main peak and what is presumably ADP. The ADP peak could come from spontaneous hydrolyzation of the third phosphate group together with the nitrophenyl group. The ATP peak could either appear due to a decaging process or that the delivered NPE-ATP solution contains ATP impurities, as the purity is only specified as >95 % [148]. The different elution times compared to the exchange optimization experiments (2.2.1) with the RP HPLC originate from the different columns used for these runs.

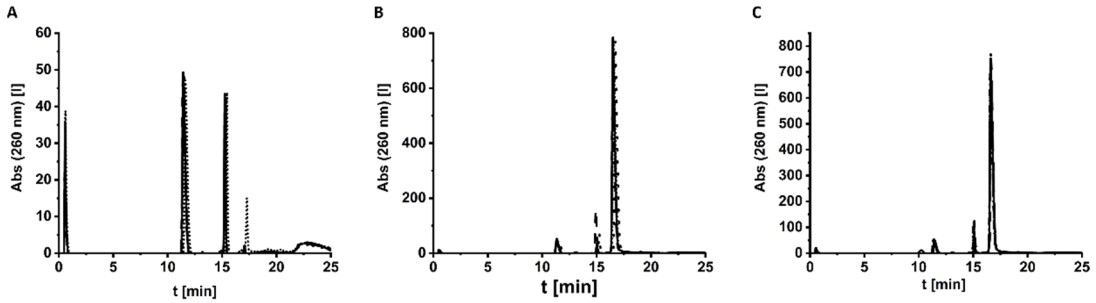


Fig. 2-28 Stability check of ATP and NPE-ATP. A: ATP under light illumination; dotted line: 30 min; black line: 1 h; long-dashed line: 2 h; short-dashed-line: 6 h. B: NPE-ATP incubated in the dark; dotted line: 30 min; black line: 1 h; long-dashed-line: 2 h; short dashed-line: 16 h. C: NPE under light illumination; dashed-dotted line: 30 min; black line: 1 h; long-dashed line: 2 h; short-dashed line: 6 h; dotted line: 16 h.

2.3 Oligomerization state and binding constant of the involved protein complexes

2.3.1 Determination of the oligomerization state applying DLS, AUC and SAXS

Although, the general activation process of UFM1 is known as depicted in **Fig. *1-5***, details like the involved oligomers are still unknown. After successful purification, the complex formation of UBA5 and UFM1, as well as UFC1 and UFM1 was investigated for better an understanding of UFM1 activation process. First, the proteins were mixed, and the behavior was analyzed with DLS and compared with UBA5 and UFC1 before the addition of UFM1. For the analysis of the UBA5-UFM1 complex, different mixtures with increasing concentrations were measured (**Fig. *2-29***). Both, single UBA5 samples and the complex samples show a broad peak coming from a polydisperse solution. The radius increases with increasing concentration. For UBA5, the software identified in mostly all measurements one radius and corresponding MW starting at 38 kDa at 0.5 mg/mL to around 60 kDa, between 4 and 8 mg/ml. The DLS signal is generally broader and noisier for the mixture. The software identified two or more radii in most measurements. Additionally, the molecular weight range of the most identified radius is smaller at higher concentrations compared to single dimeric UBA5 (**Fig. *2-30***). The radius of the scatterer is 3.0 nm at the lowest concentration and 3.3 nm at the highest concentration.

When UFC1 and UFM1 are mixed, the radius increases from 2.3 nm to 2.5, which corresponds to a molecular weight increase from 20 to 28 kDa (**Fig. *2-31***).

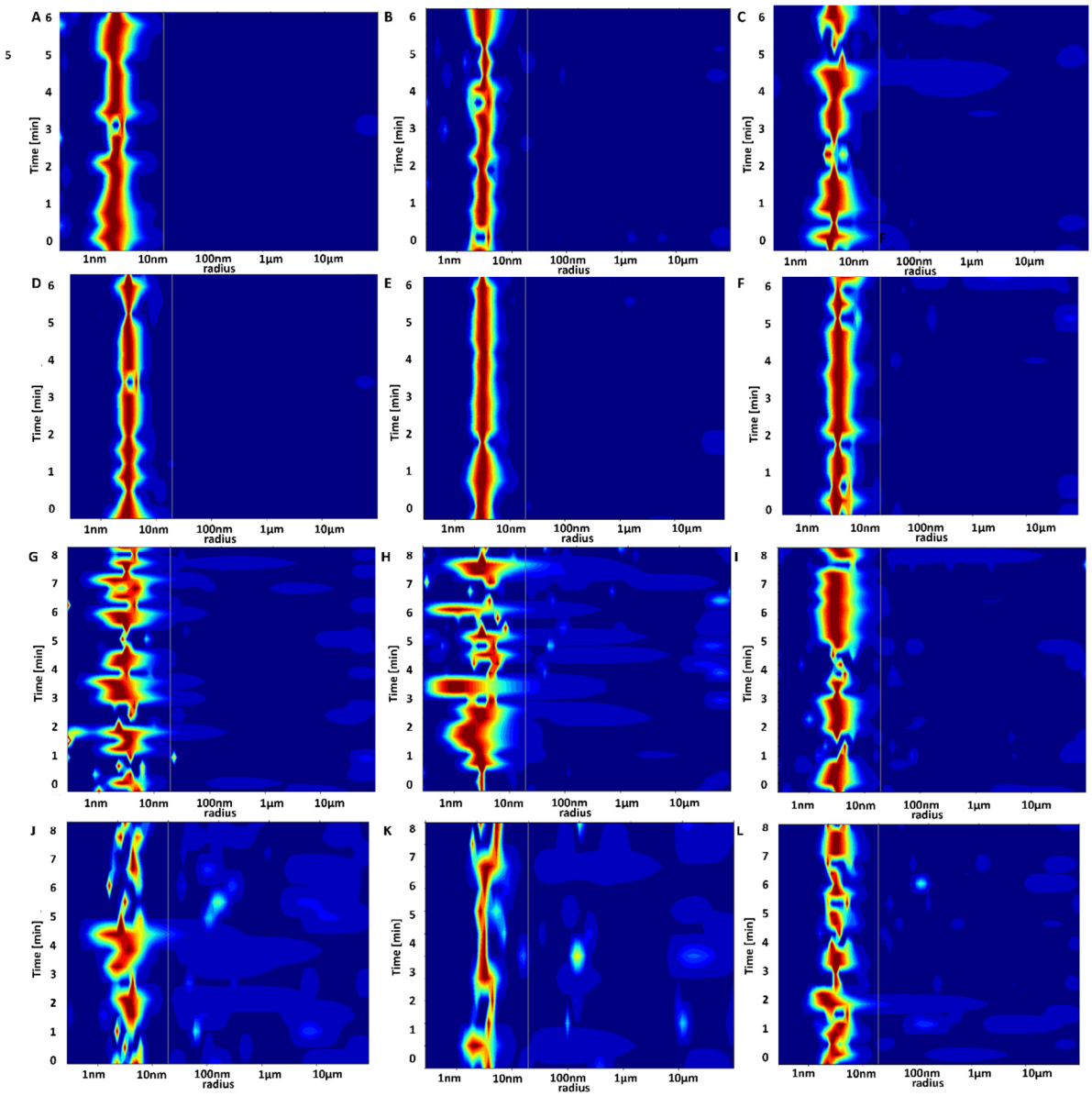


Fig. 2-29 DLS measurements of increasing concentrations of UBA5 57-346 over 6 minutes and UBA5-UFM1 complex (A-F) over 8 minutes (G-L) . A: 0.5 mg/mL; B: 1 mg/mL; C: 2 mg/mL; D: 4 mg/mL; E: 6 mg/mL; F: 8 mg/mL; G: 0.5 mg/mL; H: 1 mg/mL; I: 2 mg/mL; J: 4 mg/mL; K: 6 mg/mL; L: 8 mg/mL.

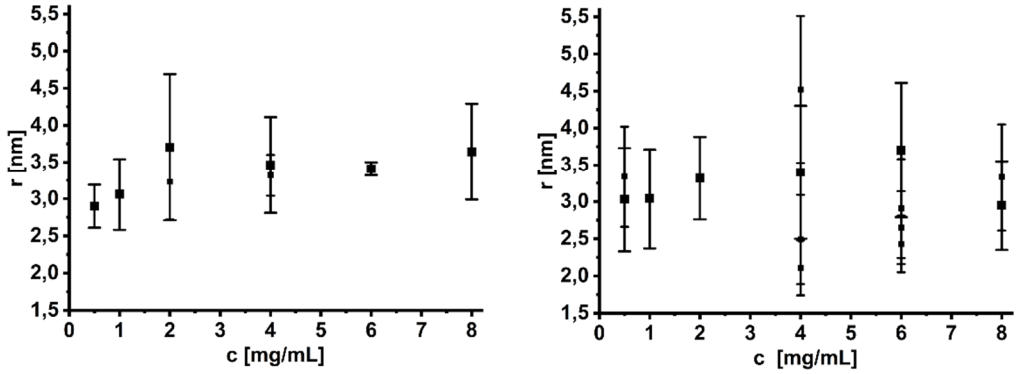


Fig. 2-30 Comparison of radii detected from DLS measurements (**Fig. 2-29**) between UBA5 (left) and UBA5-UFM1 mixture (right). The detected radii are plotted over the concentration with deviation are presented by error bars. Larger squares indicate the most detected radii during the measurement; smaller squares represent radii detected fewer times. Most measured radii of UBA5: 0.5 mg/mL: 2.9 +/-0.3 nm; 1 mg/mL: 3.1 +/- 0.5 nm; 2 mg/mL: 3.7 +/- 1.0 nm; 4 mg/mL: 3.5 nm +/- 0.7 nm; 6 mg/mL: 3.4 +/- 0.1 nm; 8 mg/mL: 3.6 nm +/- 0.7 nm. Most measured radii of UBA5-UFM1: 0.5 mg/mL: 3.0 +/- 0.7 nm; 1 mg/mL: 3.0 +/- 0.7 nm; 2 mg/mL: 3.3 +/- 0.6 nm; 4 mg/mL: 3.4 +/- 0.9 nm; 6 mg/mL: 3.7 +/- 0.9 nm; 8 mg/mL: 3.3 +/- 0.7 nm.

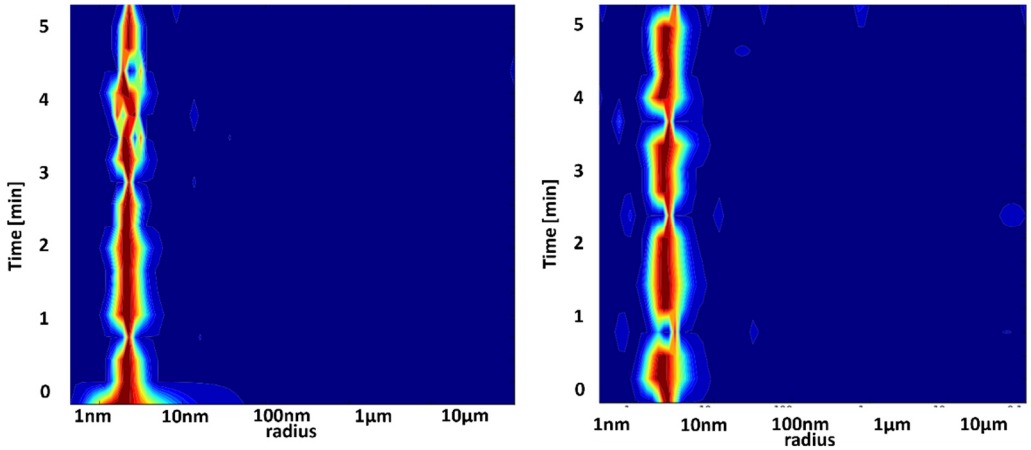


Fig. 2-31 Comparison of DLS measurements of UFC1 (left) and equimolar mixtures of UFC1 (right). The UFC1 has a radius of 2.3 nm resulting in a calculated molecular weight of 22 kDa. The mixture gives a radius of 2.5 nm, which corresponds to 28 kDa.

The concentration-dependent oligomeric state was further investigated with analytical ultracentrifugation (AUC) by analyzing the same concentration range (**Fig. *2-32*** and **Tab. 4-34**). The AUC revealed the sedimentation coefficients, where the peaks were integrated. The ratio of the peak areas revealed continuous decrease of the peak with

lower sedimentation constant while the area of the second peak increases. Below 2 mg/mL, UBA5 consists mainly as monomer; meanwhile the dimer is the main component at 2mg/mL and higher. The UBA5-UFM1 mixture shows a more complex transition. Like UBA5, at low concentrations (2 mg/mL and below), the mixture consists mainly as monomers (UBA5) or heterodimers (UBA5-UFM1) and with increasing concentration it shifts to higher stats (dimer and heterotrimer). In comparison to DLS, the AUC run also shows several peaks at different concentrations.

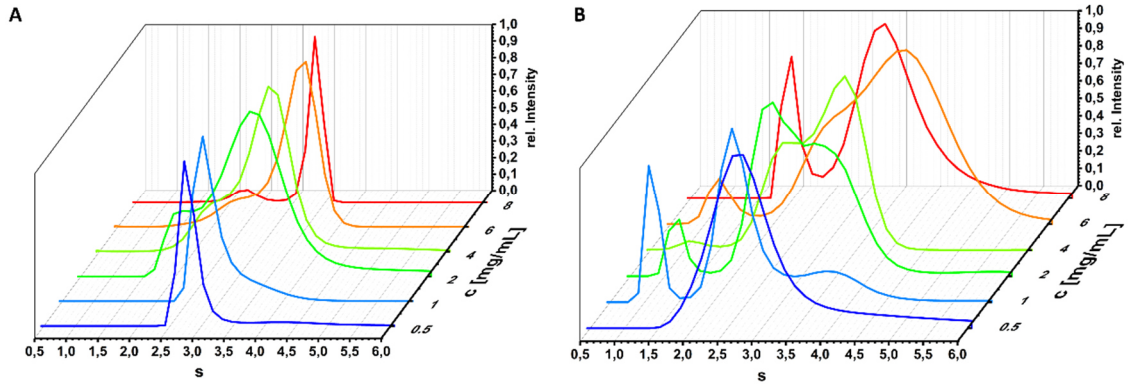


Fig. 2-32 Determined sedimentation coefficients of UBA5 (A) and UBA5-UFM1 (B) in rel. intensity over a concentration range of 0.5-8 mg/mL. For UBA5 at low concentrations a sedimentation coefficient of mainly 2.3.-2.7 S was determined, shifting to a coefficient of 3.2-3.5 S, and resulting in a calculated molecular weight of approx. 35 and 67 kDa respectively. For the complex also at low concentrations the analysis shows mainly a single sedimentation coefficient of around approx. 2.3-2.7 S. For higher concentrations a double or broad peak was identified predominantly with a sedimentation coefficient of 3.2-3.9 S.

Tab. 2-11 Calculated SEDFIT values from AUC analysis of UBA5 and UFM1 with frictional ratio (shape factor) with corresponding rmsd value of the fit, Svedberg constant of peaks in S and ratio of integrated peaks 1 and 2

sample	concentration [mg/mL] UBA5	frictional ratio	rmsd	Svedberg constant [S] [peak 1]	Svedberg constant [S] [peak 2]	ratio of integrated area peak1 vs peak 2
UBA5	0.5	1.1	0.0106	2,7	4,2	92,5 / 7,5
	1	1.1	0.0109	2,7	3,5	88,9 / 11,1
	2	1.2	0.0058	2,6	3,2	22,9 / 77,1
	4	1.3	0.0230	2,4	3,2	14,6 / 85,4
	6	1.3	0.0232	2,7	3,5	15,9 / 84,1
	8	1.3	0.0536	2,3	3,3	12,5 / 87,5
UBA5-UFM1	0.5	1.1	0.0025	2,7	-	100.0 / 0.0
	1	1.1	0.0039	2,3	3,6	75.0 / 25.0
	2	1.2	0.0024	2,6	3,2	53,6 / 46,4
	4	1.2	0.0053	2,4	3,3	32,1 / 67,9
	6	1.3	0.0085	3,1	3,9	26,4 / 73,6
	8	1.3	0.0059	2	3,3	18,6 / 81,4

The interaction of UFM1 and UFC1 was also investigated with the AUC by analyzing equimolar mixtures of UFM1 and UFC1 in comparison to UFC1 only (**Fig. *2-33***). Without UBA5 activation, UFM1 and UFC1 do not seem to interact and form a complex. Both AUC revealed very similar sedimentation coefficients and respective MW: 20.0 kDa for the UFC1 run and 20.2 kDa for the UFC1-UFM1 run. Furthermore, two additional peaks are visible. One 36.3 kDa peak in the UFC1 analysis, which might be small aggregates or a possible UFC1 dimer. In the mixture, the additional peak has a calculated MW of 3.5 kDa, which belongs to UFM1. This analysis is summarized in **Tab. *2-12***. For a quantification of the involved oligomers, SAXS experiments were performed.

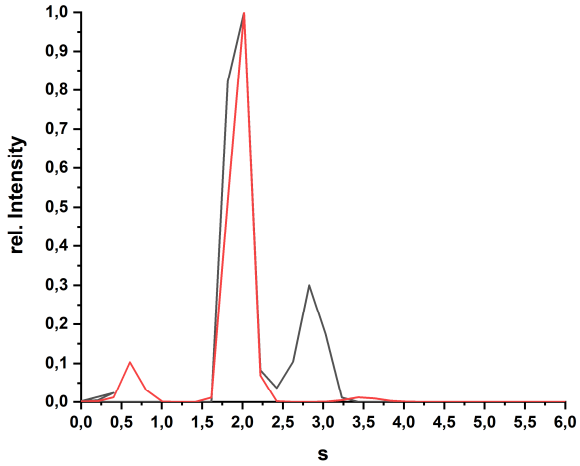


Fig. 2-33 Comparison of AUC analysis of UFC1 (black) and UFC1-UFM1 (red) mixture. Both calculated sedimentation coefficients are very similar, meaning that UFC1 does not form a complex with UFM1 without interaction with UBA5.

Tab. 2-12 Calculated SEDFIT values from AUC analysis of UFC1 and UFM1 molecular weight in kDa, frictional ratio (shape factor) of the protein and rmsd of the fit.

sample	molecular weight [kDa]	frictional ratio	rmsd
UFC1	21.0; 36.3	1.3	0.0017
UFC1-UFM1	3.5; 20.2	1.2	0.0027

After determination of the described typical SAXS values ($R_{g,}$ Porod Volume, distance distribution function), the program DAMMIN was used to calculate envelope models and compare them to structural models deposited in the pdb (**Fig. *2-34***). For UFC1 and UFM1, both models fit well to the envelope; this is also shown by low χ^2 values around 1.0 of the envelope curves compared to the experimental scattering curves. Although the UBA5 models also fit in the envelope, there are small extensions visible that cannot be explained by the crystal structure models. Additionally, the C-terminal helix of UBA5 57-346 necessary for UFM1 binding are outside of the envelope. A summary of the used models and χ^2 values is presented in **Tab. *2-13***.

Tab. 2-13 Models and corresponding data set concentrations used for DAMMIF model building

protein	pdb ID	c [mg/mL]	χ^2
UBA5 57-329	3H8V	7.1	1.198
UBA5 57-346	5IAA chain A+B	8.0	1.163
UFM1	1WXS	2.4	1.045
UFC1	2Z6O	1.1	0.723

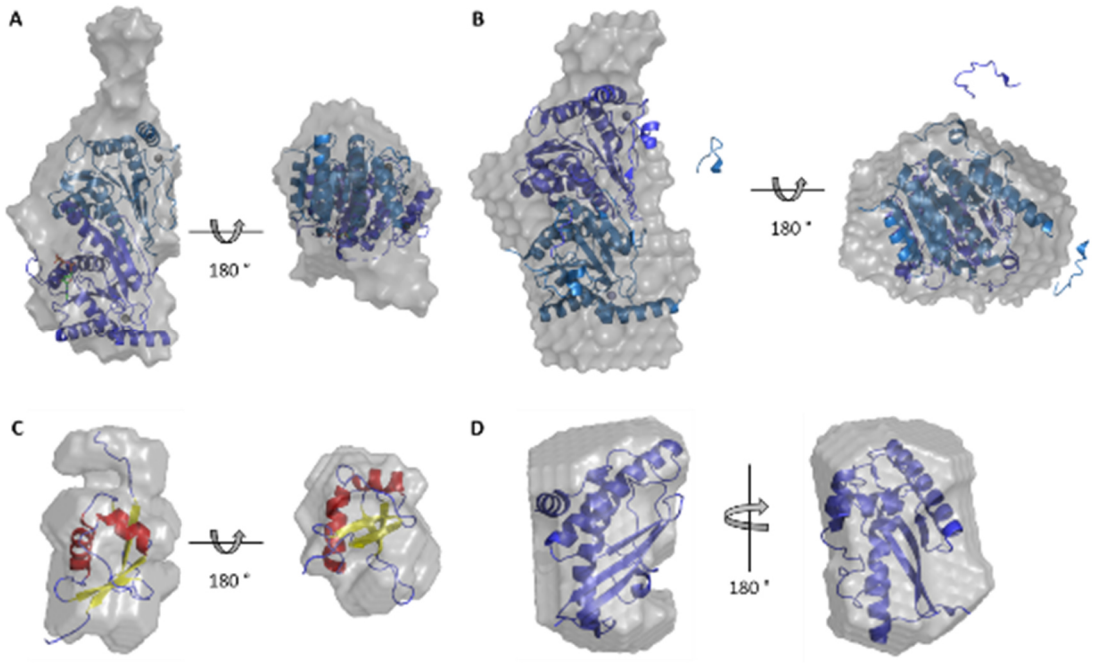


Fig. 2-34 SAXS envelopes of analyzed proteins calculated with DAMMIF. A. Envelope of UBA5 57-329 with respective ribbon model of the crystal structure; B: Envelope of UBA5 57-346 with respective ribbon model of the crystal structure; C: Envelope of UFM with respective ribbon model of NMR structure; D: Envelope of UFC1 with respective ribbon model of the crystal structure.

In order to gain a detailed insight in the complex formation of UBA5 and UFM1 and gather complementary data to the AUC and DLS experiments, the composition of the solution was investigated by comparing UBA5 and UFM1 with an equimolar mixture of UBA5 and UFM1 over a concentration range. Again, typical SAXS values (R_g, D_{max}, MW from Porod) were extracted from the experimental data. A comparison of these values is depicted in **Tab. *2-14*** as well as the increase in D_{max} in **Fig. *2-35***. To further investigate the solution composition, CRYSOL was used to find protein models that fit to the scattering curves. In particular, for the UBA5 solutions, a monomer and a dimer model were used. For the UBA5-UFM1 solutions, a UBA5 monomer and dimer model were used, as well as a UBA5-UFM1 heterotrimer (composed of a UBA5 dimer with an attached UFM1 monomer) and a UBA5-UFM1 heterotetramer. The UFM1 scattering curves was fitted with the pdb model 1WXS (**Fig. *2-37***). Except for the UFM1 data, no single model was sufficient to explain the scattering curve properly (**Fig. *2-36***). Especially in the range between a scattering vector of 0.15 to 0.25 s, the theoretical scattering curves shows minima that were not present in the measured scattering curve. Therefore, OLIGOMER was used to find the volume fraction of different models to fit the

scattering curve The final χ2 of the fits are also included in **Tab. *2-14***. A selection of the OLIGOMER fits to their respective scattering curves at different concentrations can be seen in **Fig. *2-38***, the volume fractions for each concentration are depicted in **Fig. *2-39***, the used models are depicted in **Fig. *2-40***.

The minimal ensemble necessary to find proper fits for the UBA5 solutions consisted of two monomers and three dimers. For the UBA5-UFM1 solutions, the minimal ensemble consisted of the UFM1 NMR model (pdb ID: 1wxs), two UBA5 monomers, two UBA5 dimers and 1 UBA5-UFM1 heterotrimer. It was also tried to fit the curves with a set consisting of fewer models, but the fits χ2 values for the fits were systematically higher than the ensemble used for the final fits (**Tab. *2-15*** and **Fig. *2-41***).

Tab. 2-14 Summary of the parameters for the UBA5 and UBA5-UFM1 experimental data: Radius of gyration R_g, interparticle distance D_{max}, the molecular weight MW, the χ2 goodness of the fits of the volume fractions and the SASBDB accession codes.

UBA5 concentration [mg/ml]	UFM1 concentration [mg/ml]	Guinier R_g [nm]	D_{max} [nm]	MW [kDa]	Final fit χ^2	SASBDB ID
0.3	-	2.8±0.1	9.0	48±5	1.07	SASDGG6
1.0	-	2.9±0.1	11.0	36±5	1.03	SASDGH6
1.9	-	2.9±0.1	9.8	48±3	1.01	SASDGM6
3.6	-	3.0±0.1	10.0	50±4	1.03	SASDGL6
4.4	-	3.1±0.1	12.0	62±3	1.11	SASDGK6
6.8	-	3.2±0.1	12.0	62±4	1.12	SASDGJ6
2.0	0.50	2.9±0.1	11.0	50±4	0.99	SASDL73
3.0	0.75	3.0±0.1	12.0	53±3	0.98	SASDL83
4.0	1.00	3.1±0.1	13.0	56±3	1.07	SASDL93
5.0	1.25	3.1±0.1	13.0	60±2	0.98	SASDLA3
6.0	1.50	3.1±0.1	13.0	64±3	1.07	SASDLB3
7.0	1.75	3.1±0.1	13.0	62±4	1.03	SASDLC3
8.0	2.00	3.1±0.1	13.0	62±4	1.11	SASDLD3
9.0	2.25	3.2±0.1	14.5	62±7	1.16	SASDLE3
-	1.56	1.5±0.1	5.1	7.3±2	1.01	SASDM82

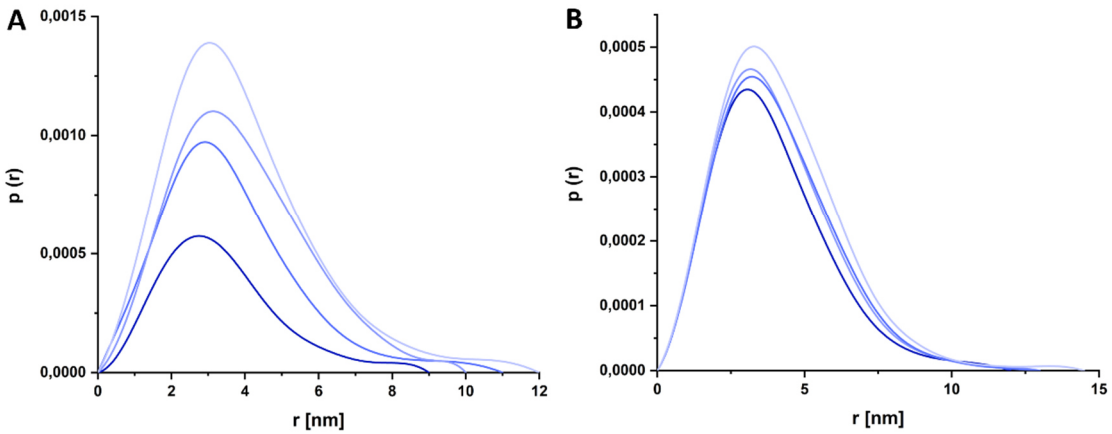


Fig. 2-35 Selection of D_{max} with increasing concentration of UBA5 (A) between 0.3 and 6.8 mg/mL and UBA-UFM1 (B) between 2 and 9 mg/mL, each from dark to light blue. The accuracy of the D_{max} estimation is around 10 %.

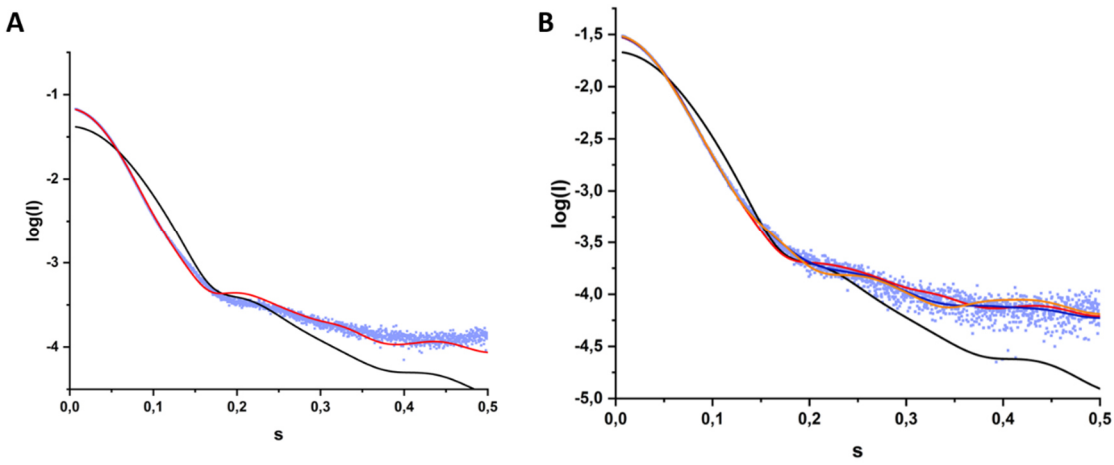


Fig. 2-36 CRYSOL fits of models to the scattering curve of UBA5 (A) and UBA5-UFM1 (B).: Black: monomer; red: dimer; orange: heterotrimer; blue: heterotetramer.

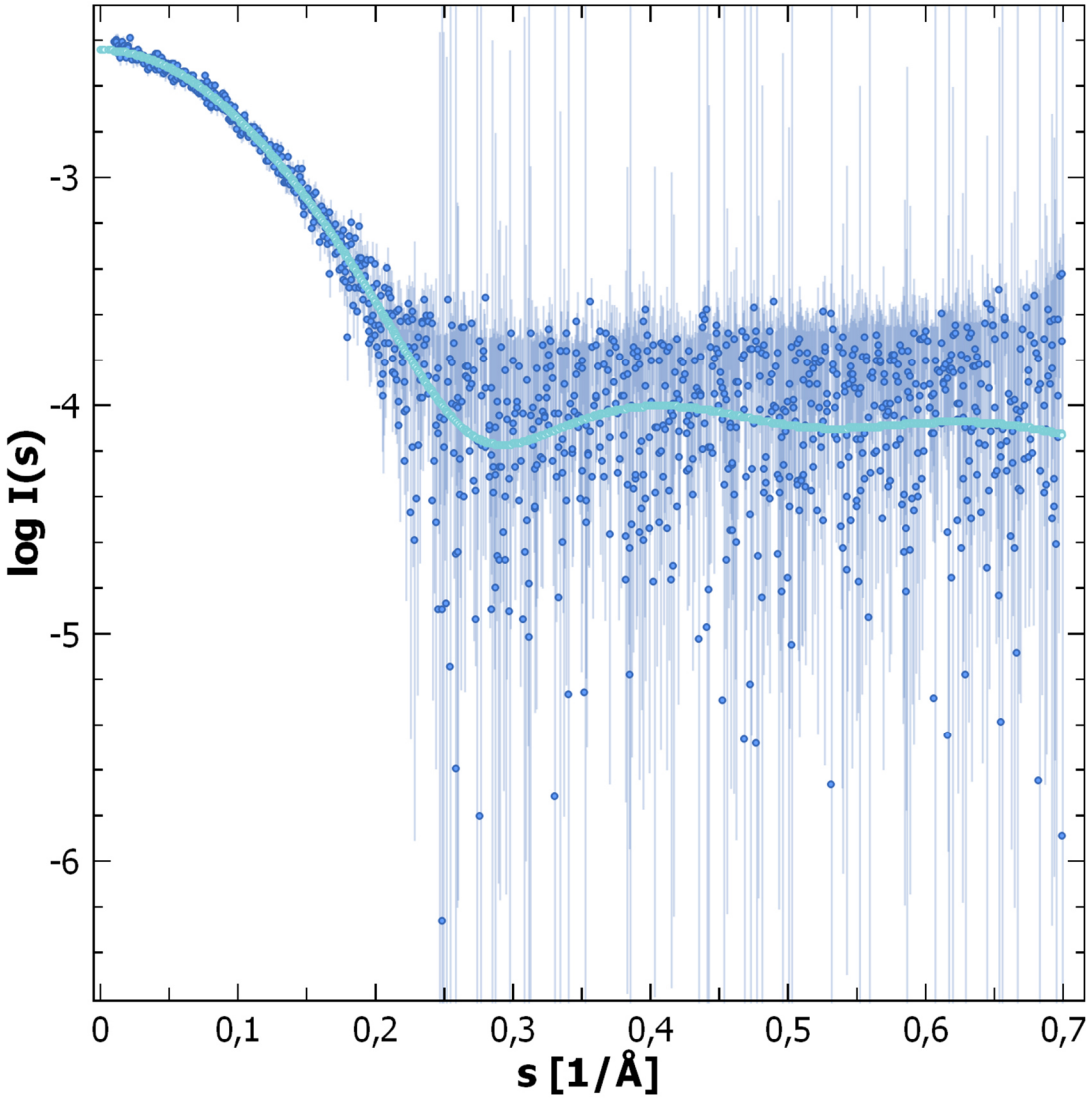


Fig. 2-37 UFM1 experimental SAXS data (dark-blue dots with error bars; SASBDB entry: SASDM82) fitted by the scattering computed from the UFM1 model (light-blue; PDB entry: 1wxs) with χ^2 of 1.01.

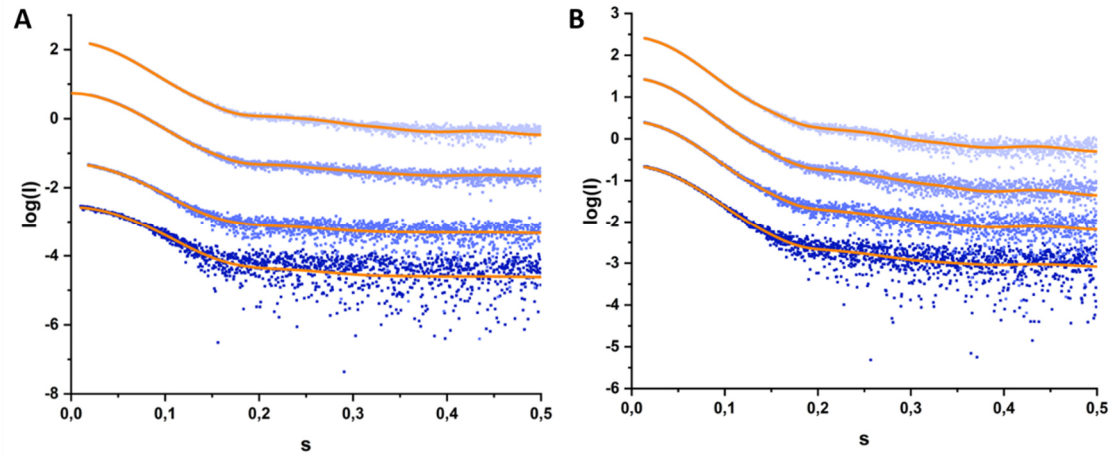


Fig. 2-38 OLIGOMER fits of models to the scattering curve of UBA5 (A) between 0.3 and 6.8 mg/mL and UBA5-UFM1 (B) between 2 and 8 mg/mL.

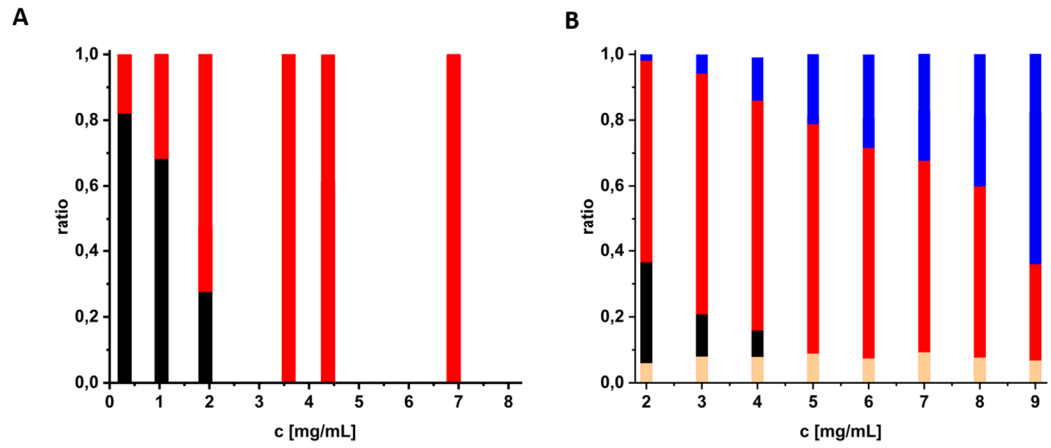


Fig. 2-39 OLIGOMER results depicting the ratio of the generated models against the concentration. A: OLIGOMER ratios with UBA5 monomers in black and dimers in red. B: OLIGOMER ratios with UFM1 in yellow, UBA5 monomers in black, UBA5 dimers in red and UBA5-UFM1 heterotrimers in blue.

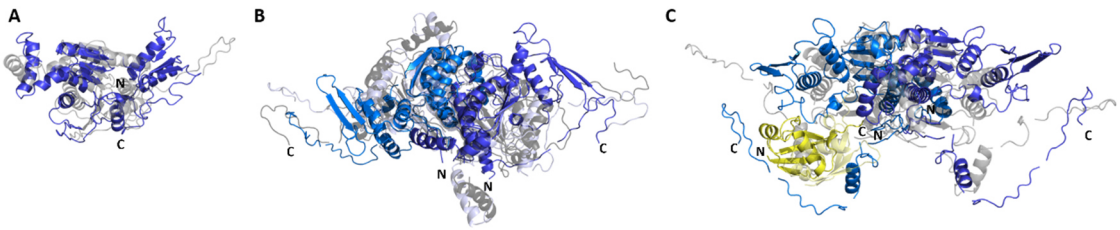


Fig. 2-40 Proteins models generated by ITASSER and SASREF/SASREFMX used for CRYSOL and OLIGOMER analysis. A: UBA5 monomers; B: UBA5 dimers; C: UBA5 heterotrimer with UFM1 in yellow.

Tab. 2-15 The χ^2 values of the fits using four-component mixtures (monomeric UFM1 and UBA5, as well as rigid body models of UBA5 dimer and UBA5:UFM1 heterotrimer) presented in the third column are systematically worse than the χ^2 values from the ensembles of partially flexible models (fourth column, same values as in Table 4). Here the best fitting models of UBA5 monomer (from SASDGG6), UBA5 dimer (from SASDGL6) and the UBA5:UFM1 heterotrimer (from SASDLE3) were used and PDB entry 1wxs was used as the UFM1 model. The data from UBA5 alone were fitted with the mixtures of UBA5 monomer and dimer only.

UBA5 concentration [mg/ml]	**UFM1 concentration [mg/ml]**	**Four models fit χ^2**	**Final fit χ^2**	**SASBDB ID**
0.3	-	1.07	1.07	SASDGG6
1.0	-	1.08	1.03	SASDGH6
1.9	-	1.05	1.01	SASDGM6
3.6	-	1.03	1.03	SASDGL6
4.4	-	1.43	1.11	SASDGK6
6.8	-	3.06	1.12	SASDGJ6
2.0	0.5	0.99	0.99	SASDL73
3.0	0.75	0.99	0.98	SASDL83
4.0	1.0	1.09	1.07	SASDL93
5.0	1.25	1.03	0.98	SASDLA3
6.0	1.5	1.18	1.07	SASDLB3
7.0	1.75	1.13	1.03	SASDLC3
8.0	2.0	1.26	1.11	SASDLD3
9.0	2.25	1.38	1.16	SASDLE3

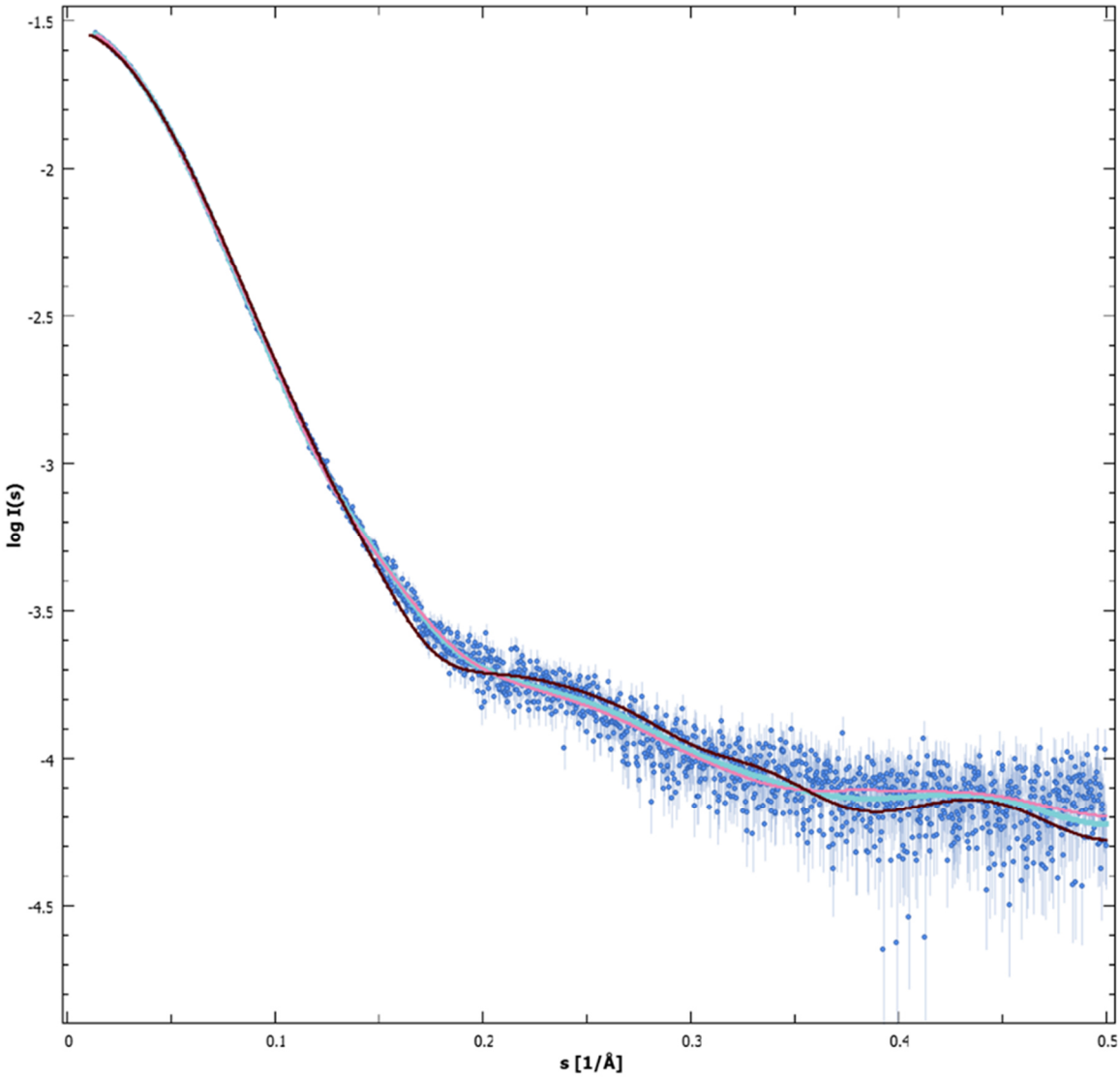


Fig. 2-41 Experimental SAXS data from the mixture of 9 mg/mL UBA5 and 2.25 mg/mL UFM1 (blue dots with error bars, same as in Fig. 3, bottom) fitted by different sets of models. Cyan: full set of models including two trimer models, one UBA5 dimer and one UFM1 monomer, $\chi^2 = 1.16$; pink: one trimer model, one UBA5 dimer and one UFM1 monomer model, $\chi^2 = 1.38$; maroon: only UBA5 dimer and UFM1 monomer models, $\chi^2 = 3.55$.

Furthermore, SAXS was used to determine the time scale of UBA5-UFM1 complex formation by using the stop-flow technique; this is described in section 4.2.4.7. For the first 60 frames, the parameters R_g and the molecular weight from the Porod volume and the I_0 with BSA as reference were calculated (**Fig. *2-42***) covering a time range of 15 s. A continuous increase was observed for all values. The radius of gyration increased from 3 to around 4.3. The molecular weight extracted from the Porod volume increased from 60 to 100 kDa, and the molecular weight calculated from the forward scattering increased from 28 to 48 kDa.

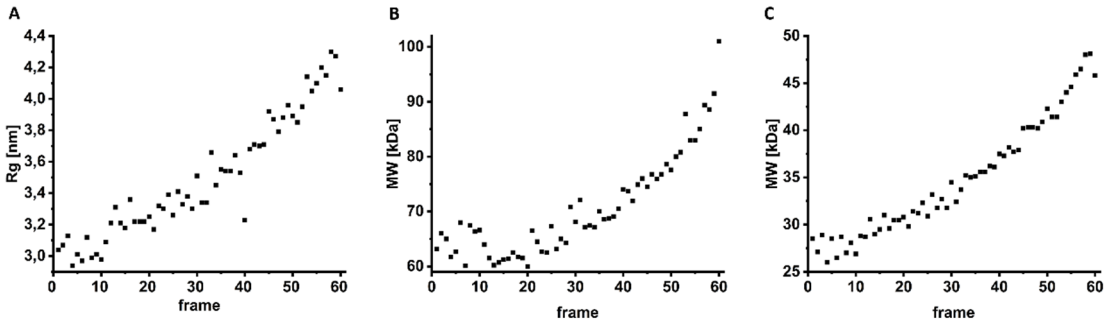


Fig. 2-42 Calculated parameters for the first 60 frames from stopped-flow SAXS experiments. A: Radius of gyration in nm; B: calculated molecular weight in kDa from Porod volume; C: calculated molecular weight in kDa from I_0 with BSA as reference.

2.3.2 Determination of the UBA5-UFM1 Binding Constant

One method to investigate and determine the binding constant of the complex formation is microscale thermophoresis; the results are shown in **Fig. *2-43***. The Prometheus NT label software calculated the dissociation constant from the data directly. For the UBA5-UFM1 complex, several measurements resulted in different K_D ranging from 2.61 to 14.00 µM, leading to an average of 7.84 µM. Furthermore, the mixture of UFM1 and UFC1 did not show any reaction during the experiment.

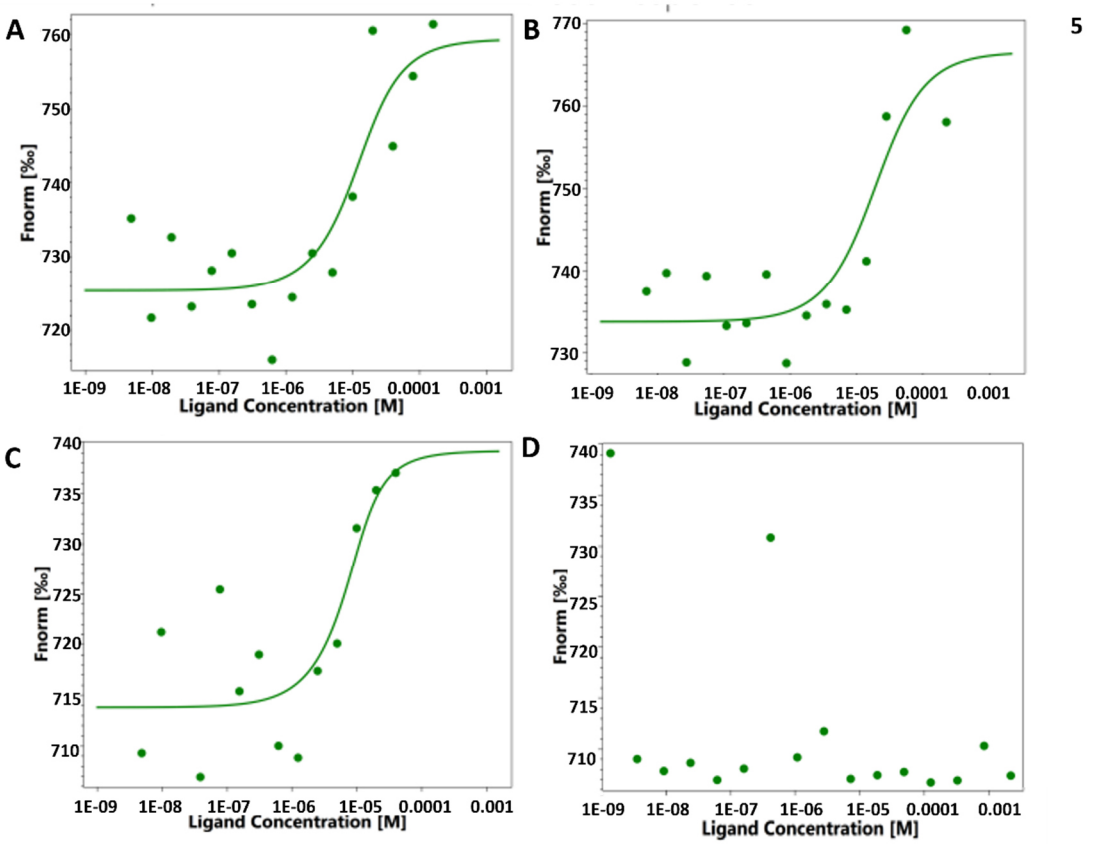


Fig. 2-43 MST measurements of UBA5-UFM1 (A-C) and UFC1-UFM1 (D) analysis. Depicted is the relative fluorescence signal over the ligand concentration UFM1. UBA5 and UFC1 both had a concentration of 10 µM. Calculated K_D were A: 6.92 µM; B: 14.00 µM; and C: 2.61 µM. The mixture of UFM1 and UFC1 (D) did not generate a strong fluorescence change, which leads to the assumption that there is no interaction between the binding partners.

Another complementary attempt to determine the binding affinity of the complexes was performed with fluorescence spectroscopy, as described in section 4.2.4.4. For UBA5-UFM1, a binding constant of around 5 µM, as well as a negative cooperativity of 0.5 could be determined (**Fig. *2-44*** and **Tab. *2-16***). The K_D is similar to the values found by MST. Again, for the mixture of UFM1 and UFC1, no fluorescence increase or decrease (quenching) was measured (data not shown).

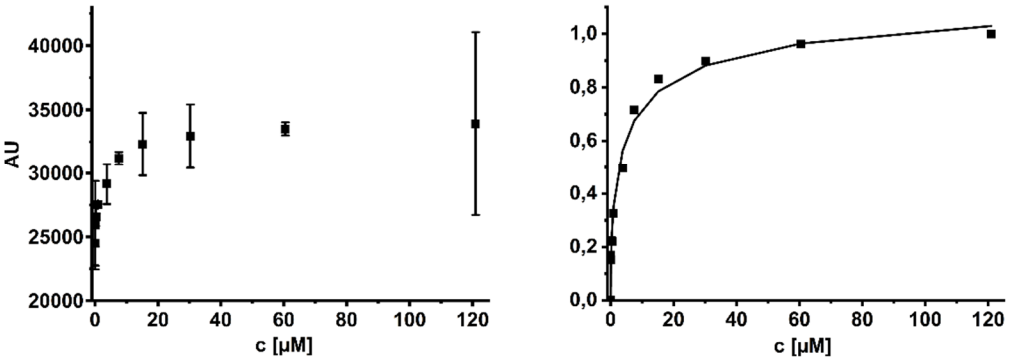


Fig. 2-44 Fluorescence intensity of UBA5 tryptophan residues against UFM1 concentration. Left: measured intensity in absolute values with error bars. Right: normalized fluorescence intensity against UFM1 concentration and plotted hill fit (black line).

Tab. 2-16 Calculated parameters from UBA5 UFM1 fluorescence measurements from hill fit

Parameter	**UBA5**
B_{max}	1.2
K_D [µM]	4.8
n	0.55

3 Discussion

3.1 Evaluation of UBA5 in complex with caged-ATP as a model system to be applied for serial crystallography

First step in the optimization of the exchange method was to check if the additional protecting group of NPE-ATP has sufficient space without sterically obstruction of the nucleotide part binding in ATP binding site. Both docking methods applied reveal that the additional group can be positioned outside the binding pocket as there is a large gap. Comparing the UAB5 binding pocket with ATP and the conformation found by SwissDock, several differences can be noticed (**Fig. 3-1** and

Tab. 3-1). Most residues interact with the ligand in the same way, which are H-Bonds between K127 and the second phosphate group, as well as hydrophobic interactions between the adenine and D104, I151, V182, D183 and N184. LigPlot indicates for ATP H-bonds between the first and second phosphate group and N184. Hydrophobic interaction between Y105 and adenine can occur via π-stacking [149]. Furthermore, the glycine residues G80, G82, and G83 have interactions to the phosphate groups (not shown in **Fig. 3-1** C). Depending on the hydrolyzation state of the phosphate groups and the residues, hydrogen bonds might also form between two aspartate residues (D106 and N184) and phosphate. To stabilize the nitrophenyl group, the residues M258 and E209 are involved via Van-der-Waals interactions. Comparing the binding affinity found by the fluorescence spectroscopy analysis of approx 0.8 µM and the value calculated from the SwissDock result of 0.2 µM, the result seems plausible, since both values are in the same range of magnitude. Although SYBYL finds a similar conformation, only seven hydrophobic and hydrophobic interactions are identified by LigPlot (s. **Tab. 2-7**) and calculated K_D is nearly 0. However, NPE-ATP should be able to bind to UBA5 without sterically obstruction through the nitrophenyl group. In comparison with published binding affinities of ATP to UBA5 [59], it can be noted that the numerical values are similar, but the order of potency differ in 10^3 (830 nM compared to the published 730 µM). The reason for this difference however is not clear, as the used protein and ATP concentrations are similar. The difference might be explained through the used buffer conditions, which differ mostly in the used buffer substance and pH (TRIS pH 7.5 compared to published Bis-TRIS pH 6.5).

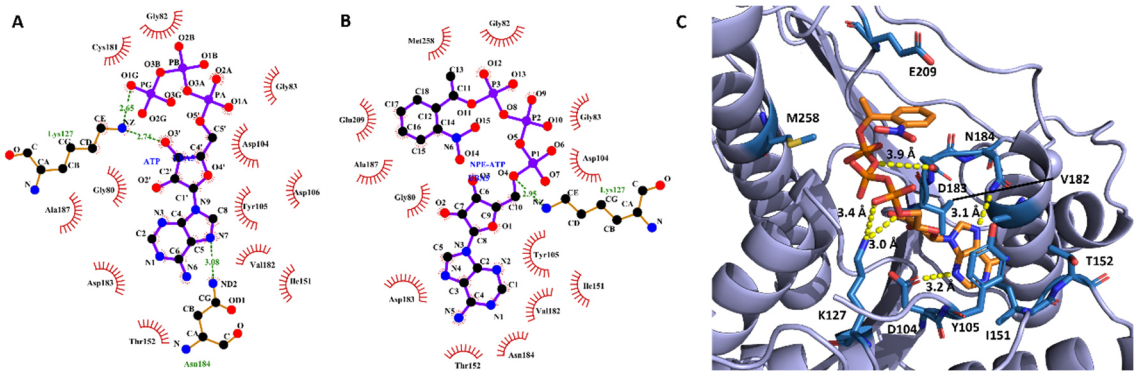


Fig. 3-1 Intermolecular interactions in the ATP binding pocket. A: Interactions of UBA5 residues with ATP determined by LigPlot [150]. B: Interactions of UBA5 residues with NPE-ATP determined by LigPlot C: Structural insight with possible interactions of NPE-ATP with surrounding residues with distance of H-bonds (yellow dotted lines) in Å.

Tab. 3-1 Summary of interactions (H-bonds and hydrophobic) between nucleotides found by LigPlot between UBA5 and ATP and NPE-ATP from docking results applying SwissDock. (A) marks the acceptor residue of the H-bond interaction.

Type of interaction	**UBA5 with ATP**	**UBA5 with NPE-ATP**
H-bond	K127 (A)	
	N184 (A)	
Hydrophobic	G80	
	G82	
	G83	
	D104	
	Y105	
	D183	
	D187	
	I151	
	V182	
	D106	
	T152	
	C181	
		N184
		E209
		M258

Determination of the saturation state of UBA5 with ATP before exchange reveals that even without additional ATP supplement during purification, there is a saturation of approx. 20 % ATP in the UBA5 binding site since the calculated concentration of ATP in the sample was approx. 5.5. µM, whereas the concentration of UBA5 used for this experiment was approx. 25 µM. This probable results from the ATP present in the E. *coli*

cells during expression [151]. This data indicates that it is necessary to develop and optimize an appropriate exchange method.

Different methods were tried to exchange ATP with NPE-ATP. Two of them use EDTA to remove magnesium ions for lowering the nucleotide affinity, the third uses an alkalic phosphatase to generate ADP or AMP, which have a lower affinity to the binding center. The three methods [152],[141] initially applied to exchange the bound ATP in UBA5 with NPE-ATP showed that the incubation with EDTA has the highest impact, especially Method B, if additional ammonium sulfate was not added. The $(NH_4)_2SO_4$ was added to have a small salting-out effect and therefore lower the ATP affinity to UBA5; however this did not prove to be successful [142]. The incubation with phosphatase showed nearly no effect, which might be based on the fact that the enzyme could not interact with ATP with its active site due to steric reasons. Therefore, the hydrolyzation to ADP or AMP was not successful. It might have been possible to combine the phosphatase with the EDTA incubation, but the phosphatase itself is magnesium dependent [153], which would interfere with the presence of EDTA and therefore inhibit its activity. To reach high exchange rates, the two-step dialysis is necessary: First dialysis to remove ATP, and then the removal of EDTA. During both incubations an excess of NPE-ATP was added to make sure to reach saturation of the binding site. After optimization, the saturation was high enough to permit crystallization trials. Furthermore, next experiments involved the evaluation of the dispersity and folding state of the protein after the exchange reaction and the affinity of NPE-ATP to UBA5.

DLS experiments before and after the exchange demonstrate, that there is an increase in radius of the proteins and the signal has a broader appearance, hinting that the protein becomes polydisperse by forming higher oligomers or microaggregates. Additional measurements show that this transition occurs during the first dialysis step with EDTA. Measuring the influence of the different additives ($MgCl_2$, $ZnCl_2$, NPE-ATP and EDTA) revealed no negative influence on the dispersity of the protein at first and for short incubation times.

Although it was expected that EDTA is responsible be either binding the zinc or magnesium ions. By consecutive measuring the protein with added EDTA over several hours confirms this suspicion as it shows that the signal becomes broader and increases

after 4 h. Therefore, the dialysis must be as short as possible to limit the negative influence of EDTA on the stability but long enough so that ATP can be removed by dialysis. Although CD spectroscopy of the protein with the different additives confirms, that the folding state does not alter much, except when a high concentration of zinc ions is added (**Fig. *2*-*17***), it is necessary to address exact effect of EDTA to the sample.

EDTA has the ability to bind divalent cations by forming chelates; therefore, it is possible that EDTA removes zinc from its binding site in addition to the binding of magnesium to lower the affinity of ATP. Determination of the concentration of zinc confirms this. It was not possible to determine the magnesium concentration before and after EDTA incubation. This is not surprising as magnesium is only loosely bound to the protein without specific bindings sites, as there are no magnesium ions visible in the electron density, even though they were added to the crystallization solution [56]. However, EDTA lowers the zinc concentration, which means that their bindings sites are unoccupied.

Zinc is bound by UBA5 through four cysteine residues in a tetradic arrangement. The binding of zinc by only cysteine is unusual, as the typical binding site consist of two histidine and two cysteine residues [154]. The removal of zinc can therefore have different effects. First the four cysteines become prone to oxidation or forming cysteine bridges to other molecules. The purpose of this binding site is so far unknown, as it is not in close proximity to the ATP binding site or has an enzymatic function. It might be possible that the binding site is necessary to keep the helix (residues 305-318) in place (s. **Fig. *3*-*2***). Without zinc, this might become flexible, which could lead to unspecific aggregations. Beside typical structural motives for DNA recognition (zinc finger) or transporters [155], zinc binding sites are also found on the binding sites of protein-protein complexes, especially with this four cysteine residue motive [156]. This may be a possible function since the zinc binding site is located near the C-terminus where UBA5 interacts with UFC1. Next, the possibilities for crystallization compared to native UBA5 were evaluated.

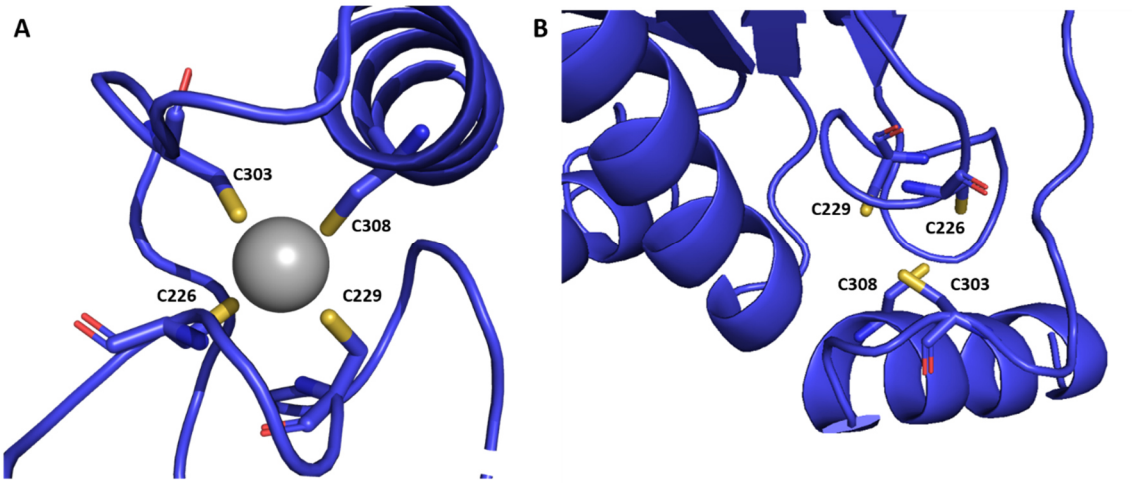


Fig. 3-2 Zinc binding site of UBA5. A: Tetradic arrangement of the four cysteine residues responsible for zinc binding. B: Zinc free binding site showing that the C-terminal helix (305-318) might become flexible.

Based on the optimized crystallization conditions, attempts were made to crystallize the constructs with NPE-ATP. However, the sensitivity of UBA5 to the removal of ATP and treatment with EDTA during the dialysis time prevented the successful crystallization of the derivate. As discussed in section 2.2.2, CD results show that the complex has a higher stability than the shorter UBA5 construct. Therefore, the complex was also used for further crystallization experiments Nevertheless, these attempts did also show no satisfying result, probably due to the same reasons: the sensitivity of the protein to EDTA and the removal of ATP before saturation with the nucleotide derivate.

Another consideration is the affinity of different nucleotides to the protein. In order to reach high saturation rates, the ligand must have preferably a high affinity. This was tested for ATP and NPE-ATP with fluorescence spectroscopy. One first observation of the experiment the different influence of the nucleotides to the fluorescence of tryptophan: ATP has a quenching effect while NPE-ATP increases it. The fluorescence potency of tryptophan is highly dependent on their chemical surroundings [157]. Hydrophobic residues increase the potency, hydrophilic decrease it. Therefore, the different effect of the nucleotides can be explained by their polarity. ATP has a high polarity, which leads to the observed quenching. The additional nitrophenyl group of NPE-ATP adds a hydrophobic property to the molecule, which increases the fluorescence of nearby tryptophan residues (**Fig. *3-3***). Comparison of affinity of both

nucleotides to UBA5 with the hill fit reveals, that they both have a dissociation constant in the low micromolar range. The affinity of NPE-ATP is four-fold lower than ATP.

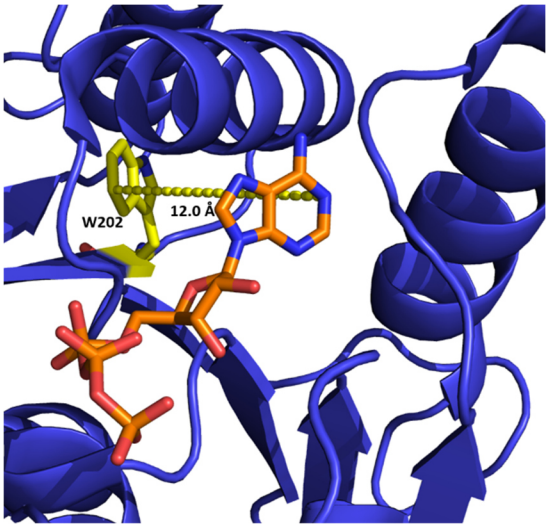


Fig. 3-3 Tryptophan residue W202 (yellow) closest to ATP in UBA5.

Although current experiments did not find crystallization conditions for UBA5 with NPE-ATP, it is also interesting to determine the stability of NPE-ATP under normal light illumination. This is necessary as several steps in the determination of the crystal structure are time consuming not feasible to perform in the dark, e.g., setting up the crystallization plates, inspection for crystals under the microscope or collecting the diffraction data. The measurements with the RP-HPLC system reveal that NPE-ATP is stable for at least 16 h with constant light illumination at ambient temperatures without measurable significant amounts of free ATP in the sample (**Fig. *2-28***). The same applies when the sample is stored in the dark. Additionally, ATP itself shows signs of self-hydrolyzation, as there are additional peaks visible. This means that NPE-ATP meets the requirements to perform crystallization experiments and time resolved diffraction experiments [158],[159].

3.2 The UBA5-UFM1 complex formation shows a stepwise binding mechanism

The DLS experiments of the purified proteins show, that UFM1 and UFC1 seem to be monodisperse in solution, but UBA5 shows a polydisperse behavior (probably mixture of mono- and dimer). Both constructs show also different ratios in the monomer/dimer distribution when eluted from the SEC (s. **Fig. *2-1*/Fig. *2-2***). UBA5 57-329 shows two peaks of similar size, whereas for 57-346 the monomer peak is distinctively larger than the dimer peak. This could mean, that the oligomeric state is influenced by the additional C-terminal residues. Therefore, the next experiments involved the investigation of the single proteins as well as mixtures of them.

UBA5, as well as the mixture of UFM1 and UBA5, showed a concentration dependent change of the oligomeric state (**Fig. *2-29***). With increasing concentration, the measured radius of UBA5 increases from 2.9 to 3.8 nm. This means at low concentrations (< 2mg/mL) UBA5 consists mainly as monomer and shifts to mainly dimer at a concentration of 2 mg/ml and above. If UFM1 is added to the protein, the measured mean radius decreases. Additionally, the signal shows more noise and is broader. The comparison of the measured radii at each concentration also reveals that the device detects more signals in the mixture then in the single UBA5 measurements (**Fig. *2-30***). Therefore, mixing UBA5 with UFM1 starts complex formation; however there is no complete tetramer formation as it is concluded from the crystal structure of the complex [58]. On the contrary, the mixture seems to form an equilibrium of different oligomers with also free monomeric and dimeric UBA5 and free UFM1. Other possible oligomers are heterodimers and heterotrimers. Since UFM1 is rather small compared to UBA5, it is possible, that the device is not able to resolve the signals from all oligomers and they overlap, generating the broad, noisy signal. Additionally, DLS measures every scatterer simultaneously, and thus it is also possible that the signal is an average of all components. This explains why the radius is generally smaller than the measured radius of pure, dimeric UBA5 of the same concentration.

To reveal if UFM1 can also form a complex with UFC1 by itself or if it needs UBA5 as activator1 before, DLS of a mixture of UFM1 and UFC1 was measured and compared to UFC1. The radius of the mixture is around 2.5 nm, and UFC1 has a radius of 2.3 nm. This corresponds to molecular weight of 28 kDa and 22 kDa, which would be the expected size of the heterodimer of UFM1 and UFC1. This could mean that UFM can form a complex with UFC1 without previous activation through UBA5. In order to confirm this results, additional complex studies were applied. The complex formation of UBA5 and UFM1 was also further investigated.

Analytical ultracentrifugation is a valuable complementary method for analyzing a mixture of macromolecules in solution. Therefore, it is very suitable to investigate the concentration dependent complex formation of the proteins.

The AUC analysis of UBA5 between 0.5 and 8 mg/ml shows that very similar results as the DLS experiments could be obtained. At low concentrations (< 2 mg/mL), UBA5 consists mainly as monomer and makes a nearly complete transition to dimer at a concentration of 2 mg/ml and above (**Fig. *2-32***). The lowest concentration (dark blue) shows only one monomeric peak. At 1 mg/ml (light blue), beside the monomer peak, a small shoulder is visible; hence there is a small amount of dimer in the solution. At 2 mg/ml and above (green to red) the dimer peak is the main peak and the shoulder beside the peak (monomer) becomes smaller with increasing concentration. The transformation from monomer to dimer can also be concluded from the ratio of the peak areas, as the area of the monomer peak decreases continuously from 93 % to 13 %, while the amount of dimer increases from 8 % to 87 % with increasing concentration. The mixture of UBA5 and UFM1 show again, that there is an equilibrium of different oligomers involved. The peaks are generally broader. At low concentrations (0.5 and 1 mg/mL) the analysis reveals mainly monomers (100 % of the peak area). With increasing concentrations, a mixture of different oligomers are involved, which could range from heterodimer via heterotrimers to the tetramer. The amount of these oligomers increases up to 81% at 8mg/mL with increasing concentration. Over the concentration range, the AUC also detects single UFM1. However, a UFM1 peak is not visible in all measurements since the absorption of UFM1 is rather low, which makes the detection of that protein difficult.

The AUC analysis of UFM1 and UFC1 cannot confirm the UFM1-UFC1 complex. Both runs show one peak with a sedimentation coefficient of around two, which fits to a molecular weight of 20 kDa (monomeric UFC1). In the mixture a peak with a molecular weight of around 4 kDa is also detected, which is the signal of UFM1. The calculated molecular weight is only half the expected size of 9 kDa, but the low absorbing UFM1 makes it difficult to assign an accurate MW to the weak signal. The UFC1 run also detects a peak at around 36 kDa, which could mean that UFC1 could form a heterodimer. Nevertheless, this kind of oligomer was not reported before and it could not be verified with other methods (DLS, SAXS). Therefore, this signal could be a form of small aggregate or an artifact [160]. For accurate determination of the involved oligomers, several SAXS experiments and applications were performed.

The first experiment involved the calculation of envelopes and comparison to the respective crystal or NMR solution structure. For all proteins, satisfying models could be calculated. Especially UFM1 and UFC1 envelope and models match very well. Although the low χ^2 of UFC1 is evidence for a slight overfitting of the model to the experimental curve, the crystal structure and the envelope model coincide without any unexplainable envelope extensions or parts of the crystal structure models which extend over the surface of the envelope.

Both, UBA5 constructs are also congruent to the envelope model. However, the envelope has some extension which could not be fitted with parts of the crystal structure. Additionally, the crystal structure extends in some small parts the envelope surface. Especially, the 57-346 crystal structure with its C-terminal extension (responsible for UFM1 binding) resides outside. This can be explained by the high flexibility of UBA5, which can be seen by the number of unmodeled amino acid residues (they could not be added to the model due to missing electron density). Notably, the residues 323-333 could not be modeled, which connects the C-terminal UFM1 binding domain with the rest of the protein. It can be assumed that without UFM1 the C-terminal residues will be unstructured and highly flexible and will not reside at specific conformation as it is depicted by the crystal structure model. Therefore, it is possible that they will have other conformations that could explain the envelope extensions. The

same might apply for the other flexible, unstructured elements in the proteins, like loops.

Next, the concentration dependent behavior of UBA5 and the UBA5-UFM1 complex in solution was further analyzed and compared to the previous AUC, fluorescence spectroscopy and DLS results. For the concentration dependent oligomeric state of UBA5, similar conclusions can be drawn from the SAXS experiments. With increasing concentrations, the R_g and molecular weight increase, which corresponds to the transition of UBA5 from monomer to dimer. By using CRYSOL, it was tested if the scattering behavior could be explained by a single model. For UBA5 only, a monomer or a dimer model was used. For the UBA5-UFM1 complex, a UBA5 monomer and dimer and a UBA5-UFM1 heterotrimer and -tetramer were used. For UBA5, the monomer model did not fit properly to the scattering curve at all concentrations and the quality decreased even more with increasing concentrations. The $\chi 2$ for the dimer ranged from 1.3 to 2.8 with increasing concentrations. Similar behavior could be observed for the CRYSOL fits to the scattering of the UBA5-UFM1 complex. The monomer model did not fit to the scattering curve at all concentrations. The dimer and tetramer showed similar quality at low concentrations ($\chi 2 \approx 1.5$). The only improvement of the fits was observed for the heterotrimer with increasing concentrations. However, none of the single models were sufficient to explain the scattering behavior of both, UBA5 and the complex. Especially the theoretical scattering curves of the models showed minima between a scattering vector of 0.15 and 0.25 s, while these features were not present in the measured data. Additionally, the general higher quality of the fits at low concentrations can also be explained in the lower signal to noise ratio of the measured sample.

Therefore, OLIGOMER was used to model the composition and ratios of different models in solution. These are also very similar to the results obtained from AUC and explains the calculated molecular weight and polydisperse appearance of the DLS experiments. However, in order to generate satisfying fits, two similar monomers and three similar dimers generated with the combination of ITASSER, SASREF and SASREFMX were necessary, showing the high flexibility of UBA5. At concentrations between up to around 2 mg/mL, OLIGOMER shows an equilibrium between monomers and dimers. With

increasing concentration, a mixture of dimers can be found. The Chi² of the fits range from 1.07 to 1.12. This shows that UBA5 is flexible in solution, especially the C- and N-termini, which can be concluded from the elongated appearance of distance distribution plots (**Fig. *2-35***).

The analysis of the complex reveals again a more complex situation. Similar to pure UBA5, R_g and molecular weight increase with increasing concentrations but not in the same stepwise manner. Also, the mean values stay at similar values as UBA5, although due to the addition of UFM1, it was expected that these values should be higher compared to single UBA5. There are possible explanations for this result. Firstly, the values (molecular weight from Porod volume, maximum distance distribution and R_g) are the average of all components contributing to the scattering. This includes also free UFM1, as well as monomers or dimers. Secondly, without UFM1, the UBA5 oligomers might be more flexible, e.g., the C-terminal residues responsible for the binding of UFM1 are free. This could lead to a higher maximum distance distribution and similar R_g or Porod volumes as the complex, which is more compact when UFM1 is bound. Nevertheless, the type of oligomers changes with increasing concentrations.

It shows that there is a mixture of homodimer, free UFM1 and heterotrimer. UBA5 monomer is also found at the lowest concentration of 2 mg/mL, similar to what was found in the same concentration range of UBA5 only. In comparison to the OLIGOMER analysis of UBA5 only, an ensemble of two trimer models, one UBA5 monomer and two dimer models with the addition of the NMR model of UFM1 (pdb ID: 1wxs) was necessary to achieve satisfying fits with $\chi < 1.2$. Using a set of fewer models both of the UBA5 data and the UBA5-UFM1 data shows that the quality of the OLIGOMER fits where systematically lower than the ensembles used for fitting (**Tab. *2-15***and **Fig. *2-41***). Since the fluorescence spectroscopy experiments show a negative cooperativity of UBA5 towards UFM1, it could mean that the complex forms in a stepwise manner, where the UFM1 molecules bind and are adenylated to the UBA5 dimer consecutively instead of simultaneously. This fits to the results found by Mashahreh et al.[59] which showed that UBA5 is a weak dimer and is stabilized by UFM1. UFM1 also enhances the affinity of ATP and that the binding of ATP enhances the affinity of UFM1 to UBA5.

Lastly, the time scale of complex formation was analyzed by time resolved SAXS. It shows that the complex forms in the second time regime. However, over the time range of the analysis of about 15 s does not plateau. A R_g and molecular weight increase is observed. This could be mean, that the complex formation up to a tetramer is not finished within the observed time. Alternatively, the consecutive increase in the measured values could also come from radiation damage that appears within the collected 60 frames [161]. The transition between increase of values by complex formation or radiation damage is hereby hard to discern, which makes an accurate estimation difficult.

The following experiments were conducted to determine the kinetic and the dissociation constant of the complex formation. Two methods were used to study the kinetics and determine the dissociation constant: microscale thermophoresis and fluorescence spectroscopy. Both methods are suitable without further protein labeling since the tryptophan fluorescence in UBA5 and UFM1 can be measured. This is feasible because UFM1 does not contain any tryptophan residues, which excludes any possible cross reactions.

The three measurements of the UBA5-UFM1 complex formation by MST extract an average dissociation constant of 7.8 µM. But the spread is between 2.6 and 14 µM. Additionally, the measurement does not seem to reach a saturation, even though an excess of 22-fold was applied at the highest UFM1 concentration. This makes it difficult to determine an accurate K_D. Therefore, fluorescence spectroscopy was applied as a complementary method. The analysis reveals a similar K_D of around 5 µM similar to what was found by Padala et al [54]. Additionally, the application of the hill formula reveals a negative cooperativity of around 0.5 involved in the complex formation. This means that the binding of one UFM1 molecule lowers the affinity of the complex for the second UFM1 molecule. This could explain why in the DLS and AUC experiments the tetrameric form was not detected.

Both methods also confirm the missing complex formation by mixing UFM1 and UFC1. Both do not detect a change in the fluorescence signal with increasing UFM1 concentration, which confirms that both proteins do not interact before the UBA5 activation step. This supports the AUC data and contradicts the result of the DLS.

The analysis of the behavior of the proteins in solution can bring in more details in the UFM1 activation process. While it is reported from crystal structures that the active unit of UBA5 is a dimer [56,58], the analysis reveals that the UBA5 is also present as monomer depending on the concentration. However, it could also be shown that the dimer is the most important from.

In addition, it was observed that both UFM1 and UFC1 are present, as monodisperse, monomeric solutions. While it was already shown that UFM1 interacts with UFC1 if it was adenylated by UBA5 before [61], these results proved that no interaction occurs without this step as it is shown by the fluorescence spectroscopy, MST, and AUC. This is a way for the activation mechanism to regulate UFM1 very tightly. This ensures that only adenylated UFM1 can be further activated by UFC1 without the possibility that inactive UFM1 blocks the binding site. This can be additional control mechanism to the already existing transformation of UFM1 from it 85 residues pro-UFM1 by UfSP1. The details how the recognition of the right UFM1 is achieved is currently unknown.

Another activation control mechanism can be concluded from the analysis of the UBA5-UFM1 complex with DLS, AUC, kinetic analysis, and SAXS. All data show the quick transition of monomeric UBA5 to homodimer with increasing concentration. If UFM1 and UBA5 are mixed, an equilibrium of different oligomers can be found with varying fraction depending on the concentration. Beside mono- and dimeric UBA5, as well as heterotrimers can be found. With increasing concentrations, the higher oligomers become more prominent with the UBA5-UFM1 heterotrimer as the main proportion. Although the crystal structure of the complex showed that the tetramer activates UFM1 via a cross-binding reaction, the analysis proved that it is not the preferred state in solution. It seems that it is an equilibrium of trimers and UBA5 homodimers. This is supported by the kinetic data, which showed that the dissociation constant is in the low µM range. While similar values of this were already reported [54], cooperativity of the different subunits were not considered. The analysis of the hill formula revealed a negative cooperativity of around 0.5, which means, that the affinity of UBA5 to a second UFM1 is lowered if one UFM1 subunit is already bound to the complex. This could mean that the activation process is slightly different from the already reported tetrameric cross-binding mechanism. The data shows that activation can happen in a stepwise

manner, where on UFM1 is first bound and adenylated, before a second UFM1 molecules binds to UBA5 and would explain why in the known crystal structures only one subunit contains ATP in the binding site. This kind of mechanisms are already known [162]. This activation procedure could also be a third way to control the UFM1 activity, as UFM1 would be activated step by step before it is transferred to UFL1 by UFC1.

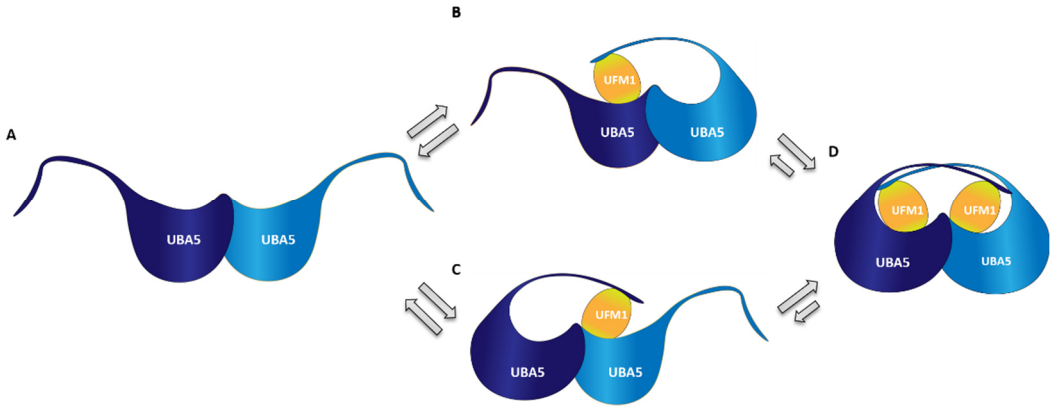


Fig. 3-4 Possible binding mechanism of UFM1 (yellow) to UBA5 (blue) in a stepwise manner. This binding mechanism might occur as there is a measured negative cooperativity when UFM1 binds.

Although concentrations in the low mg/ml range are for enzymes unusually high [163], it might be possible for the activation of UFM1. Since UFM1 is involved in the stress response of cells, it is necessary to give a fast reaction to stress stimuli [164]. This can be achieved by fast and high expression of the protein and a tight regulation of the activity by three means: processing from its pro-form [165], step wise activation [166], and binding of only adenylated UFM1 by UFC1.

3.3 Outlook

3.3.1 Time resolved experiments with caged compounds

One of the goals of this work was to evaluate whether UBA5 in complex with caged ATP may be a suitable model system for time resolved structural studies. After docking experiments showed the possibility of UBA5 also to bind NPE-ATP without steric complications, one nucleotide exchange method could be developed to exchange the nucleotide up to a satisfying saturation. However, crystallization conditions of UBA5 with NPE-ATP could not be found yet. The main reason for this is the sensitivity of the protein to EDTA, which is necessary to remove ATP. This is not unusual, as it is known that apoenzymes have often a lower stability than holoenzymes [167]. Additionally, the removal of zinc out of its binding site could also influence the stability negatively, which could lead to unspecific aggregation and a polydisperse solution and therefore prevent successful crystallization.

So far, other methods for exchanging the nucleotide seem not to be practicable. The usage of phosphatases is not working, probable due to inaccessibility of the catalytic center of the phosphatase to the ATP binding pocket. Simple soaking of the crystal is difficult, because the determination of the nucleotide loading status of UBA5 reveals, that there is still approx. 20 % of ATP bound to the active site, even without addition during protein purification. The determination of the nucleotide affinity reveals that ATP has four-fold lower dissociation constant. This would prevent the saturation of the binding pocket with NPE-ATP to rates high enough for the application in TR crystallization (> 90 %).

It could be proved that the exchange of nucleotides with a caged derivate is possible in proteins. The stability under normal conditions is high enough to prepare experiments without many precautions to prevent photoactivated release of the protecting group. However, for the UBA5 system, no optimal crystallization conditions have been found so far. Therefore, it is necessary to optimize the procedure to prevent aggregation of the protein during nucleotide exchange. This might be achievable by adjusting the EDTA concentration, time of dialysis or amount of NPE-ATP. The amount of NPE-ATP that can be used so far is limited. The commercially available NPE-ATP is delivered as a solution in an appropriate buffer with a concentration of 10 mM. This determines, how much

nucleotide can be added to the sample to reach specific excess ratios. Additionally, the amount of pure protein, which is necessary for crystallization is quite high, especially for time resolved experiments. Therefore, high amount of caged nucleotide is also necessary, which is quite expensive if purchased from a commercial distributor. A possible workaround would be the synthesis in the own lab, what would also opening the possibility to set the concentration of the NPE-ATP stock solution. A higher concentration of the stock solution would minimize dilution effects and also enable to reach higher excess rates, which would lead to higher saturation rates. With this possibility, simple soaking of the caged compound to exchange ATP in the crystal might be feasible. Furthermore, it is possible to synthesize caged ATP with a variety of different protecting groups differing in its properties like solubility, extinction coefficient, or quantum yield. Another approach to evaluate the applications of caged compounds in serial crystallography is the usage of another model protein, which is known to be very stable and has a low sensitivity to environmental changes.

However, on order to perform such experiments on a regular basis, several parameters must be evaluated. These include to evaluate the optimal crystal size for the laser pulse, which decages the caged compound in all asymmetric units simultaneously. Furthermore, the optimal laser intensity and wavelength must be known to ensure high penetration depths in the crystal in context of short decaging times. Additionally, the time duration of the laser pulse is also an important parameter to achieve an optimal decaging reaction.

3.3.2 UFM1 activation cascade

Next experiments should investigate in more detail in the whole activation procedure. This involves the study of the transfer reaction to UFC1 and finally to UFL1. It might be possible to crystallize the complex of UFM1, UBA5 and UFC1 if a full-length UBA5 construct is used. It might be necessary to used mutants to stall the reaction at specific points, similar to what was done in the crystallization experiment of the UBA5-UFM1 complex, where the catalytic cysteine C250 was mutated to an alanine. It is also necessary to get more details about UFL1 and the final targets since currently there is little known about these proteins. Especially, detailed structural information is missing.

Furthermore, it would be interesting to perform in vivo experiments under ER stress conditions to reveal more findings about the regulation of UFM1 activation. The expression rates in intracellular concentrations would be important to compare with the corresponding data presented in this work. In addition, experiments investigating the detailed mechanism of the interaction of UFM1 with its assigned proteases should reveal more details about the regulation, either via the transition to its active from pro-UFM1 or via cleavage from the target protein.

4 Material and Methods

4.1 Material

4.1.1 Chemicals and Enzymes

Tab. 4-1 List of chemicals

name	**distributor**
acetic acid	VWR
acetonitrile	FischerChemical
acrylamide	Carl Roth
adenosine triphosphate	Carl Roth
agar agar	Carl Roth
agarose	Carl Roth
ammonium chloride	Carl Roth
ammonium persulfate	Carl Roth
ammonium sulfate	Carl Roth
ampicillin	Carl Roth
bromophenol blue	Biorag
coomassie brilliant blue R250	AppliChem
dimethyl sulfoxide	Sigma-Aldrich
dithiothreitol	Carl Roth
dipotassium hydrogen orthophosphate	Carl Roth
ethylenediaminetetraacetic acid disodium salt	Sigma-Aldrich
ethylene glycol	Merck
glucose	Carl Roth
glutathione	Carl Roth
glycerol	Carl Roth
glycine	Carl Roth
(4-2-hydroxyethyl)-1-piperazineethnesulfonic acid	VWR
hydrochloric acid	VWR
imidazole	Carl Roth
isopropanol	VWR
isopropyl β-D-1-thiogalctopyranoside	Carl Roth
kanamycin	Carl Roth
lactose	Sigma-Aldrich
lithium chloride	Merck
magnesium chloride	Carl Roth
2-methyl-2,4-pentanediol	Merck
nitrophenyl adenosine triphosphate	Jena Bioscience
paraffin oil	VWR
peptone	Carl Roth
polyethylene glycol 3500	Fluka
polyethylene glycol 6000	Fluka
potassium chloride	Carl Roth

name	**distributor**
potassium dihydrogen phosphate	Merck
silicon oil	Carl Roth
sodium chloride	Carl Roth
sodium citrate	Carl Roth
sodium dihydrogen phosphate	VWR
sodium dodecyl sulfate	Carl Roth
sodium fluoride	Merck
sodium hydroxide	Th. Geyer
sucrose	Fluka
tascimate	Hampton Research
tetra-n-butylammonium bromide	Carl Roth
tetramethyl ethylenediamine	Merck
tris(hydroxymethyl)aminomethane	Carl Roth
xylene cyanol	Sigma-Aldrich
yeast extract	Carl Roth
zinc chloride	Merck

Tab. 4-2 List of used enzymes

enzyme	**distributor**
DreamTaq DNA Polymerase	Thermo Fisher Scientific
EcoRI endonuclease	Thermo Fisher Scientific
HRV protease	in house production
NdeI endonuclease	Thermo Fisher Scientific
BamHI endonuclease	Thermo Fisher Scientific
XhoI endonuclease	Thermo Fisher Scientific
T4 Ligase	Thermo Fisher Scientific
Phusion High Fidelity DNA Polymerase	Thermo Fisher Scientific
proteinase K	Sigma Aldrich

4.1.2 Buffers and Stock Solutions

All buffers were dissolved in deionized water, filtered and stored at RT.

Tab. 4-3 Composition of Complex Buffer

500	mM	NaCl
20	mM	Tris
	pH	7.5

Tab. 4-4 Composition of PBS Buffer

140	mM	NaCl
2.7	mM	KCl
10	mM	Na_2HPO_4
1.8	mM	KH_2PO_4
	p	7.4

Tab. 4-5 Composition of PBS Glutathione Buffer

140	mM	NaCl
2.7	mM	KCl
10	mM	Na_2HPO_4
1.8	mM	KH_2PO_4
20		glutathione
	pH	7.4

Tab. 4-6 Composition of Wash Buffer

500	mM	NaCl
20	mM	Tris
30	mM	imidazole
	pH	7.5

Tab. 4-7 Composition of Elution Buffer

150	mM	NaCl
20	mM	Tris
200	mM	imidazole
1	mM	DTT
	pH	7.5

For purification of UBA5 57-329, 2 mM ATP and 5 mM $MgCl_2$ were added to the buffer.

Tab. 4-8 Composition of SEC Buffer

150	mM	NaCl
20	mM	Tris
2	mM	DTT
	pH	7.5

For purification of UBA5 57-329, 2 mM ATP and 5 mM $MgCl_2$ were added to the buffer.

Tab. 4-9 Composition of Complex Buffer

50	mM	NaCl
20	mM	Tris
2	mM	DTT
	pH	7.5

Tab. 4-10 Composition of TBAB Buffer

5.0	mM	TBAB
8.3	mM	K_2HPO_4
17.1	mM	KH_2PO_4

4.1.3 Buffers and Solutions for Agarose Gels

Tab. 4-11 Composition of TAE Buffer. pH was adjusted until EDTA was completely dissolved

1	mM	EDTA
40	mM	Tris
20	mM	acetic acid

Tab. 4-12 Composition of TAE agarose gels. Solution was heated until agarose was completely dissolved

1	mM	EDTA
40	mM	Tris
20	mM	acetic acid
1.5	%	agarose

Tab. 4-13 Composition of Loading Buffer

10	mM	Tris-HCl (pH 7.6)
60	mM	EDTA
0.03	%	bromophenol blue
0.03	%	xylene cyanol

4.1.4 Plasmids, Strains, and Oligonucleotides

Tab. 4-14 List of used plasmids

plasmid	**distributor**
pEt-28a(+)	Invitrogen
pGEx-6P-1	GE Healthcare Life Science

Tab. 4-15 List of used E. *coli* strains

strain	**purpose**
DH5α	amplification of ligation products
BL21 (DE3)	expression of UFM1 and UFC1
BL21 Star	expression of UBA5

Tab. 4-16 List of used oligonucleotide primers for PCR reactions

name	**sequence**
Uba5_57_fwd_NdeI	GATCCATATGATGGCGCTGAAACGGATG
Uba5_346_rev_EcoRI	CATGGAATTCCTAGACAAGTTCGATACCCCATTC
Ufm1_fwd	CTAGGGATCCATGAGCAAAGTGTC
Ufm1_rev	CTAGCTCGAGCTAGCCTACACGATC
Ufc1_fwd	CTAGGGATCCATGGCGGATG
Ufc1_rev	CTAGCTCGAGCTATTGGTTGCATTTC

4.1.5 Media

All media were dissolved in deionized water, autoclaved and stored at RT.

Tab. 4-17 Composition of 1 L LB media

10	g	tryptone
5	g	yeast extract
5	g	NaCl
	pH	7

Tab. 4-18 Composition of 1 L LB-agar for agar-plates

10	g	tryptone
5	g	yeast extract
5	g	NaCl
15	g	agar-agar
	pH	7

Tab. 4-19 Composition of 1 L 2xYT media

16	g	tryptone
10	g	yeast extract
5	g	NaCl
	pH	7

4.1.6 Buffers and Solutions for SDS-PAGE

Tab. 4-20 Composition of 2x sample buffer

25	% (v/v)	glycerol
12.5	% (v/v)	stacking gel buffer
10	% (w/v)	SDS
0.5	% (w/v)	bromphenol blue
0.2	% (w/v)	DTT

Tab. 4-21 Composition of Coomassie staining solution

25	% (v/v)	isopropanol
10	% (v/v)	acetic acid
0.1	% (w/v)	coomassie brilliant blue R250

Tab. 4-22 Composition of Destaining solution

20	% (v/v)	acetic acid

Tab. 4-23 Composition of Electrode Buffer

192	mM	glycine
25	mM	Tris
0.1	% (w/v)	SDS

Tab. 4-24 Composition of gel buffers

stacking gel (pH 6.8)	0.5	M	Tris
separating gel (pH 8.8)	1.5	M	Tris

Tab. 4-25 Composition of gels for SDS-PAGE

Stacking gel (4 %)			
9.2	mL	H_2O	
3.8	mL	stacking gel buffer	
2	mL	acrylamide/bisacrylamide (37.5:1)	30 %
150	µL	SDS	10 % (w/v)
75	µL	APS	10 % (w/v)
15	µL	TEMED	
Separating gel (12 %)			
10.2	mL	H_2O	
7.5	mL	stacking gel buffer	
12	mL	acrylamide/bisacrylamide (37.5:1)	30 %
300	µL	SDS	10 % (w/v)
150	µL	APS	10 % (w/v)
15	µL	TEMED	
Separating gel (15 %)			
7.2	mL	H_2O	
7.5	mL	stacking gel buffer	
15	mL	acrylamide/bisacrylamide (37.5:1)	30 %
300	µL	SDS	10 % (w/v)
150	µL	APS	10 % (w/v)
15	µL	TEMED	

4.1.7 Commercially available Kits and Screens

name	**purpose**	**distributor**
GeneJet Gel Extraction Kit	agarose gel extraction	Thermo Fisher Scientific
Plasmid miniprep kit I (C-line)	plasmid isolation	peqlab
JCSG	Crystallisation Screen	Molecular Dimensions
ComPas Suite	Crystallisation Screen	Qiagen
Gel Filtration Calibration Kit	SEC column calibration	GE Healthcare

4.1.8 Devices

Tab. 4-26 Devices

device	name	distributor
agarose gelelectrophoresis system	BlueMarine 100	SERVA
agarose gel scanner	Bio-1000F	MicroTek
CD spectrometer	J-815	JASCO
centrifuge	Multifuge X3R	Thermo Fisher Scientific
centrifuge	minispin plus	Eppendorf
centrifuge	Multifuge X1R	Thermo Fisher Scientific
centrifuge	5424R	Eppendorf
centrifuge	5415R	Eppendorf
DLS	SpectroSize300	Xtal Concepts
fluorimeter	XSpark	Tecan
HPLC	1200 series	Agilent
incubator	Thermixer comfort	Eppendorf
incubator	KS3000 I control	IKA
incubator	innova 44	New Brunswick
microscale thermophoresis	Monolith NT label free	nanotempertech
PCR cycler	^{3}Prime	Techne
size exclusion chromatography	ÄKTApurifier	GE healthcare
photometer	CO8000 Cell density Meter	WPA biowave
polycarylamide gelectrophoresis system	EV231	peqlab
sonicator	soniprep 150	Haake Mess-Technik
ultracentrifuge	Optima	Beckman Coulter

4.1.9 Software

Tab. 4-27 Software

name	purpose
JASCO spectra manager	analysis of cd spectral data
ATSAS	software package for analysis of SAXS data
CCP4	software package for analysis of diffraction data
ChemStation for LC 3D Systems	Corntol software for HPLC
COOT	model building
ORIGIN	plotting software
SEDFIT	fitting of analytical ultracentrifugation data
SEDNTERP	calculation of buffer values for analysis of analytical ultracentrifugation data
SwissDock	online docking software
SYBYL	docking software
Unicorn	control software for ÄKTA
XDS	integration of diffraction data

4.2 Methods

4.2.1 Molecular Biology Methods

4.2.1.1 Cloning of UBA5

For most experiments, either a construct containing residues 57-346 or residues 57-329 were used. Amplification of the DNA encoding residues 57-346 was performed by PCR using UBA5 wild type gene isoform 2 encoding residues 57-404 send by eurofines as template. The forward primer encoded 15 complementary nucleotides of the 5‘-end of UBA5 gene sequence and nucleotides encoding a NdeI cleavage site. The reverse primer encoded 15 complementary nucleotides of 3‘-end of UBA5 gene sequence, as well as an EcoRI cleavage site. The PCR program for amplification is shown in **Tab. *4-28***. Composition of the PCR reaction mixture is shown in **Tab. *4-29***.

Tab. 4-28 PCR program for amplification of UBA5 57-346. The steps 2-4 were repeated 30 times.

step	duration [s]	temperature [°C]
initial melting	240	98
melting	30	98
annealing	30	58
elongation	70	72
final elongation	240	72
cooling	-	4

Tab. 4-29 Composition of the PCR reaction mixture for amplification of UBA5 gene

Component	Volume [µL]
Phusion Buffer	5
Phusion High Fidelity DNA Polymerase	0.5
Forward Primer	2
Reverse Primer	2
DNA template	5
H_2O	35.5

For cleavage with restriction enzymes of both the PCR product and an empty pET-28a(+) the mixture shown in **Tab. 4-30** was incubated for 2 h at 37 °C. After incubation the mixtures were purified with the Gel Extraction Kit.

After purification the products were ligated by mixing 0.5 µL T4 ligase, 0.5 µL H_2O, 10 µL cleaved pET-28a(+) and 5 µL cleaved UBA5 PCR product. The mixture was incubated for 1 h at 20 °C. The ligase was inactivated by heating the mixture for 10 min at 80 °C.

Tab. 4-30 Mix for restriction enzyme cleavage of UBA5 DNA sequence

amount [µL]	additive
5	DNA
5	Buffer Orange
1	EcoRI
1	NdeI
38	H_2O

4.2.1.2 Cloning of UFM1 and UFC1

Amplification of theUFM1 DNA encoding residues 1-83 was performed with a PCR from the template send by eurofines. Forward primer encoded 15 complementary nucleotides of the 5‘-end of UFM1 gene sequence and nucleotides encoding a BamHI cleavage site. The reverse primer encoded 15 complementary nucleotides of 3‘-end , as well as an XhoI cleavage site. The PCR program for amplification is shown in **Tab. *4-31***. Composition of the PCR reaction mixture is shown in **Tab. *4-32***.

Tab. 4-31 PCR program for amplification of UFM1 and UFC1. The steps 2-4 were repeated 30 times.

step	duration [s]	temperature [°C]
1. Denaturation	240	98
2. denaturation	30	98
3. annealing	30	58
4. elongation	70	72
5. elongation	360	72
6. hold	-	4

Tab. 4-32 Composition of the PCR reaction mixture for amplification of UFM1 and UFC1 gene

Component	**Volume [µL]**
DreamTaq Buffer	5
DreamTaq DNA Polymerase	0.5
Forward Primer	2
Reverse Primer	2
DNA template	5
H_2O	35.5

For cleavage with restriction enzymes of both the PCR products and an empty pGEx-6p-1 vector, the mixture shown in **Tab. *4-33*** was incubated for 2 h at 37 °C. After incubation the mixtures were purified with the Gel Extraction Kit.

Tab. 4-33 Mix for enzymatic cleavage of UFM1 DNA sequence

amount [µL]	**additive**
1	DNA template
20	Tango Buffer
1	BamHI
2	XhoI
26	H_2O

After purification the products were ligated by mixing 0.5 µL T4 ligase, 3.5 µL H_2O, 5 µL cleaved pGEx-6P-1 and 10 µL cleaved PCR product. The mixture was incubated for 90 min at 20 °C. The ligase was inactivated by heating the mixture for min at 80 °C.

4.2.1.3 Transformation

For amplification and verification of the ligated product, DH5α cells were used. For expression the strains BL21 (DE3) or BL21 Star were used. 5 µl Plasmid were mixed with 50 µL competent cells and incubation the cells for 30 min on ice. After incubation heat shock was performed by incubating the cell for 45 s at 42 °C. After 10 min resting on ice, 450 µL LB was added to the cells and they were incubated for 1 h at 37 °C and 300 rpm. 100 µL of the cell suspension was plated on agar plates containing 100 µg/mL ampicillin. The plates were incubated for 17 h at 37 °C.

For Sequencing, single colonies were picked and incubated in 5 mL LB media containing 50 µg/mL kanamycin for 17 h at 37 °C. The plasmids were isolated following the protocol form the Plasmid Isolation Kit. The sequence was verified by sending 5 µL plasmid sample added with 5 µL sequencing primer to EZseq.

4.2.1.4 Agarose Gelelectrophoresis

Electrophoresis is the movement of charged particles through an electric field. It can be used to separate DNA molecules of different sizes, as the negative charged DNA moves with different velocity through a polymer gel like agarose, depending on the molecule size.

Samples were prepared by mixing 1 µL loading buffer with 9 µL DNA sample. The sample was transferred to a pocket of a 1 % agarose gel. As reference, 5 µL of low range marker were loaded in one pocket of the gel. Electrophoresis was performed by applying constant current of 150 mA and voltage of 125 V for 20 min. DNA bands were visualized with the gel scanner.

4.2.2 Expression

4.2.2.1 Preculture

For expression of all UBA5 constructs the following preculture was prepared: one single colony of transformed cells were picked from an agar plate containing 50 µg/mL kanamycin and transferred to 200 mL LB media containing 50 µg/mL kanamycin. The media was incubated for 16 h at 37 °C.

For expression of UFM1 and UFC1 constructs the following preculture was prepared: one single colony of transformed cells were picked from an agar plate containing 100 µg/mL ampicillin and transferred to 200 mL LB media containing 100 µg/mL ampicillin. The media was incubated for 16 h at 37 °C.

4.2.2.2 Expression of UBA5

Expression of UBA5 57-329 was performed with BL21 Star as host in an autoinduction media. 10 mL of preculture were transferred to 1 L LB Media containing 50 µg/mL Kanamycin, 0.5 % glycerol, 0,05 % glucose, 0.2 % lactose, 50 mM Na_2HPO_4, 50 mM K_2HPO_4, 100 mM NH_4Cl, 10 mM $(NH_4)_2SO_4$. The media was incubated for 5 h at 37 °C,

followed by incubation for 16 h at 20 °C. Cells were harvested by centrifugation of the cell suspension at 17000 g for 10 minutes at 4 °C. The pellet was resuspended in 30 mL lysis buffer and stored at -20 °C.

The expression of UBA5 57-346 was performed with BL21(DE3) as host in 2xYT media. 10 ml of preculture were transferred in 1 L 2xYT media containing 50 µg/mL Kanamycin and incubated at 37°C. When an optical density OD_{600} reached 0.6-0.8, expression was induced by addition of 1 mM IPTG. After incubation for 16 h at 20 °C, cells were harvested by centrifugation of the cell suspension at 17000 g for 10 minutes at 4 °C. The pellet was resuspended in 30 mL lysis buffer and stored at -20 °C.

4.2.2.3 Expression of UFM1 and UFC1

The expression of both proteins was performed with BL21(DE3) as host in LB media. 10 mL of preculture were transferred to 1 L LB media containing 100 µg/mL ampicillin. The suspension was incubated at 37 °C. When an optical density OD_{600} reached 0.6-0.8, expression was induced by addition of 1 mM IPTG. After incubation for 3.5 h at 37 °C, cells were harvested by centrifugation of the cell suspension at 17000 g for 10 minutes at 4 °C. The pellet was resuspended in 30 mL PBS buffer and stored at -20 °C.

4.2.2.4 SDS-PAGE

The sodium dodecyl sulfate polyacrylamide gel electrophoresis can be used to separate protein molecules by size. Sodium dodecyl sulfate binds with its hydrophobic chain to the core of proteins and gives them a negative charge from the sulfate group. By applying a current, the protein molecules move to the positive. The velocity of the molecules through a polyacrylamide gel is then dependent by the size of the molecules.

Samples were prepared by mixing 10 µL sample buffer with 10 µL protein solution and incubation for 10 min at 98 °C. 10 µL of sample were added on 12 % polyacrylamide gel (fractions containing UBA5) or 15 % polyacrylamide gel (fractions containing UFM1 or UFC1). As reference, 5 µL of protein marker were added to one pocket of the gel. For electrophoresis, a constant current of 25 mA per gel and voltage of 250 V was applied for 100 minutes. Gels were stained for at least 1 h in coomassie staining solution and destained in destaining solution until protein bands were clearly visible.

4.2.3 Protein Purification

4.2.3.1 Affinity Chromatography

Cell lysis was usually performed by sonication of 30 mL cell suspension for ten times 30 s with 30 s pause on ice. For crystallization purposes, the cell lysis was performed by continuously freezing the sample at -80 °C and thawing it in a water bath at RT for 5 times. The lysate was centrifuged for 20 min at 17000 g at 4 °C. Samples containing UBA5 were applied on 5 mL Ni-NTA matrix equilibrated with lysis buffer. After 2 h incubation at RT, the flow through was collected and the matrix washed four times with 15 mL wash buffer. The protein was eluted with 8 times 5 mL elution buffer. To yield high amounts of proteins several optimization steps were necessary to improve the yield. Firstly, the concentration of imidazol was increased from 20 to 30 mM in the Wash Buffer. Then the number of elution steps was increased to 8. Furthermore, fresh DTT was added to every Elution buffer, as well as to the SEC Buffer. Fractions containing UBA5 were pooled and concentrated with a 10 kDa cutoff concentrator to 2-4 mL.

Samples containing UFM1 or UFC1 were applied on 5 mL GST-agarose matrix equilibrated with PBS buffer. After 2 h incubation at RT, the flow through was collected and the matrix washed four times with 15 mL PBS buffer. The protein was eluted with 8 times 5 mL PBS glutathione buffer

Fractions containing protein were pooled and concentrated with a 3 kDa cutoff concentrator to 2-4 mL. For cleavage of GST-tag, 100 µL of 1 mg/mL HRV protease was added and the sample was incubated for 16 h at RT while dialyzing against 1 L PBS buffer.

The sample was applied on a 16/60 Sup75 column equilibrated with SEC buffer. Elution was collected in 2 mL fractions with a flow rate of 0.8 ml/min. Fractions containing UBA5 were pooled and checked for monodispersity and folding state with DLS and CD spectroscopy, respectively. All fractions from all purification steps were controlled with SDS-PAGE.

4.2.3.2 Size Exclusion Chromatography

Size exclusion chromatography or gel filtration is a method for separating molecules by their size. The sample is solved in the mobile phase and passing the stationary phase: a highly porous material with a large surface like superose. Large molecules do not fit into

the pores and pass the material fast without interaction. Smaller molecules cover a longer distance by flowing into the pores of the stationary, leading to longer elution times.

The sample was applied on a 16/60 Sup75 column equilibrated with SEC buffer. Elution was collected in 2 mL fractions with a flow rate of 0.8 ml/min. Fractions containing protein of interest were pooled and checked for monodispersity and folding state with DLS and CD spectroscopy, respectively. All fractions from all purification steps were controlled with SDS-PAGE.

For the calibration of Sup 75 column, the Gel Filtration Calibration Kit from GE Healthcare was used as described in the corresponding manual. In one run, Blue Dextran was applied to determine the void volume V_0. The following proteins were used for the calibration: ribonuclease A, proteinase K, ovalbumin, albumin, conalbumin. The partition coefficient was calculated with the following formula:

$$K = \frac{V_e - V_0}{V_c - V_0}$$

Eq. 21 Calculation of the partition coefficient K with V_e: protein elution volume; V_0: void volume; V_c: column volume (120 mL).

4.2.4 Biophysical Methods

4.2.4.1 Atomic Absorption Spectroscopy

The atomic absorption spectroscopy (AAS) is spectroscopic method for quantification of specific elements (mostly metals) in a solution. The sample is vaporized and illuminated with light from a hollow cathode lamp with a resonance frequency for the specific element. The light is absorbed by the element proportional to the concentration of atoms in the sample. By measuring a standard, the concentration of the element can be determined down to the ppm range.

This method was used to determine magnesium and zinc concentrations bound to UBA5 before and after treatment with EDTA during ATP exchange experiments. For analysis of native UBA5, 1500 µL of a 4.28 mg/mL UBA5 solution in SEC was transferred in a 25 ml Kjeldahl flask and diluted with deionized water. For analysis of the EDTA treated UBA5,

2000 µL of a 2.79 mg/mL UBA5 solution were dialyzed for 4 h against SEC buffer with 20 mM EDTA. After a second dialysis against SEC buffer to remove EDTA the sample was transferred to a 25 mL Kjeldahl flask. As reference, SEC buffer was measured as well. Each sample was measured as a doublet with a Solaar S Series AAS from Thermo by the central elemental analytic facility of the chemistry department of the university of Hamburg.

4.2.4.2 Dynamic Light Scattering

Dynamic light scattering is a method to determine the size of macromolecular molecules in solution, as well as the dispersity of the solution. Both values are derived from the translational diffusion coefficient D_t. The relationship of the diffusion coefficient and particle hydrodynamic radius is described by the Stokes-Einstein equation [168].

$$D_t = \frac{k_B T}{6\pi\eta R_h}$$

Eq. 22 Stokes-Einstein Equation; k_B: Boltzmann's constant; T: temperature in K; η: solvent viscosity; R_h: hydrodynamic radius

During a DLS measurement, the sample is illuminated with laser light, which is scattered by the macromolecules in the solution. The scattered light is detected at a fixed angle. Due to the Brownian motion the interparticle distance changes, leading to constructive or destructive interference and scattering intensity fluctuations over time. The particle movement can be described with an autocorrelation function, which decays over time, depending on motion of the particles and their diffusion coefficient. From the autocorrelation function the diffusion coefficient can be extracted.

20 µL of protein sample were centrifuged for 60 minutes at 21400 g transferred to a DLS cuvette. For a short-term measurement, each interval was repeated 20 times. Long term measurements had a duration of at least 16 h with one interval each hour. Each measurement had a duration of 30 s, temperature was set to 20 °C. The scattering intensity was measured at an angle of 90 ° to the incoming beam, which had a wavelength of 660 nm and a power of 100 W.

4.2.4.3 CD Spectroscopy

Circular dichroism spectroscopy can be used to determine the folding state and ratio of secondary structure elements of a protein. Circular dichroism is the different absorption of left or right circular polarized light by chiral molecules. Since proteins contain chiral amino acids, they absorb circular polarized light. Additionally, different secondary structures absorb at different parts of the UV spectra, which can be used to determine the conformation and secondary structure of a protein sample.

For secondary structure analysis, 100 µL of protein sample in CD buffer were added to a 1 mm quartz glass cuvette and measured at 20 °C. Each Spectra was recorded between 260 and 190 nm with a scan rate of 50 nm/min. Spectra of ten cycles were accumulated. The measured m° were transformed into mean residue ellipticity (MRE.

$$MRE = \frac{m° * MW}{AA * 10 * l * c}$$

Eq. 23 Calculation of mean residue ellipticity in °*cm²*dmol^{-1}. m°: ellipticity in millidegree; molecular weight of the sample; AA: number of residues; l: pathlength in cm; c: concentration in g/L.

The secondary structure was calculated by uploading the spectrum to the BeStSel server as described in the manual of the website.

As reference, ten cycles of CD buffer were recorded with a scan rate of 100 nm/min. Melting curves were recorded with temperature gradient of 1 °C/min between 20 °C and 95 °C and CD signal was recorded at 208 and 220 nm.

4.2.4.4 Fluorescence Spectroscopy

The method was used to determine the dissociation constant of the UBA5-UFM1 interaction, as well as to compare the binding affinity of ATP and NPE-ATP to UBA5. For the experiment, the native fluorescence of tryptophan in the UBA5 sequence was measured. Upon binding of UFM1 or the nucleotides, the chemical environment was altered. This resulted in a change of the tryptophan fluorescence intensity with increasing concentration of the titrated binding partner.

For UFM1 binding experiments, 10 µL of a 10 µM UBA5 in complex buffer were mixed with 10 µL of a UFM1 solution in complex buffer in a concentration range between 0.03 and 242 µM in 384 well microtiterplate. The mixtures were incubated for 10 minutes at

room temperature and measured with a TECAN Spark at 25 °C. Samples were excited by illuminating the wells at 280 nm (20 nm bandwidth) and detection the emission at 360 nm (35 nm bandwidth). As reference, complex buffer, 5 µM UBA5 without added UFM1 and UFM1 without UBA5 in the same concentration range were measured under the same conditions.

For calculation of the dissociation constant K_D of the complex, the average of three independent measurements were plotted against the UFM1 concentration and a fit was calculated based on the Hill equation. For the fit, the fluorescence intensity was normalized by calculating the ratio of the fluorescence intensity from one concentration and the fluorescence concentration of the highest UFM1 concentration.

$$FI = \frac{B_{max} * c^n}{K_D^n + c^n}$$

Eq. 24 Hill equation. FI: Fluorescence intensity, B_{max}: maximal protein concentration able to bind ligand, c: concentration of UFM1, n: cooperativity, K_D: dissociation constant

For nucleotide binding experiments, 10 µL of a 20 µM UBA5 in complex buffer were mixed with 10 µL of either ATP or NPE-ATP solution in complex buffer in a concentration range between 0.2 and 62.5 µM in a 384 well microtiter plate. The mixtures were incubated for 10 minutes at room temperature and measured with a TECAN Spark at 25 °C. Samples were excited by illuminating the wells at 280 nm (20 nm bandwidth) and detection the emission at 360 nm (35 nm bandwidth). As reference, complex buffer, 10 µM UBA5 without added nucleotide and nucleotide solutions without UBA5 in the same concentration range were measured under the same conditions.

For calculation of the dissociation constant K_D of the nucleotide binding affinity, the average of three independent measurements were plotted against the nucleotide concentration and a fit was calculated based on the Hill equation. For the fit of ATP, the relative quenching was calculated by dividing the measured fluorescence intensity by the highest measured fluorescence intensity and the hill fit was applied to the relative quenching. For the fit of NPE-ATP, the fluorescence intensity was normalized by calculating the ratio of the fluorescence intensity from one concentration and the fluorescence concentration of the highest NPE-ATP concentration.

4.2.4.5 Microscale Thermophoresis

Another to analyze molecular interactions is microscale thermophoresis (MST). Thermophoresis is the movement of particles in a temperature gradient. During a MST experiment a capillary is filled with a solution of the analyte and ligand and illuminated for a short time-period with an infrared laser on a specific spot, which generates a temperature gradient inside the capillary. The heated spot leads to a depletion of the analyte, which can be quantified by the Soret coefficient S_T (**Eq. *25***). The movement of the particles is detected by measuring the fluorescence signal of covalently attached or intrinsic fluorophores. Since the thermophoresis of analyte and analyte-ligand complex differ due to changes in solvation energy, size or charge, it can be used to quantify molecule interactions[169].

$$\frac{c_{hot}}{c_{cold}} = e^{S_T \Delta T}$$

Eq. 25 Calculation of the Soret coefficient. c_{hot}/c_{cod}: normalized concentration; S_T: Soret coefficient; ΔT: Temperature difference

For determination of the dissociation constant of the UBA5-UFM1 complex, 10 µM of UBA5 solution in complex buffer was mixed with varying concentrations of UFM1 ranging from 318 µM and 4.8 nM. Measurements were performed with the Monolith NT label free. Microphoresis was detected my measuring the fluorescence intensity of tryptophans inside the capillary. Binding constants were determined by the internal software of the device.

4.2.4.6 Analytical Ultracentrifugation

For determination of the oligomeric state of UBA5 and UBA5-UFM1 complex, sedimentation velocity experiments were performed on Beckman Coulter Optima analytical ultracentrifuge. 400 µL of sample were filled in the sample chamber of the Double Sector cell (12 mm centerpiece). 410 µL of complex buffer were filled in reference chamber. The sedimentation was observed by measuring the absorption along the sample chamber at 280 nm for 8 h at 20 °C and 50000 g.

Data analysis was performed with the programs SEDNTERP for calculating the sample parameters and SEDFIT for determination of sedimentation coefficients and corresponding molecular weights. Sample parameters were calculated with SEDNTERP.

Protein parameters used for the fitting are depicted in ***Tab. 4-34***. Sample density was 1.005 g/mL, dynamic viscosity was $1.002*10^{-2}$ Poise.

Tab. 4-34 Parameters for protein calculated with SEDNTERP. MW: molecular weight, VBar: partial specific volume; Hydration: Hydration shell; A (280 nm): extinction coefficient at 280 nm in $g/l*cm^{-1}$.

Parameter	UBA5	UBA5-UFM1
MW [kDa]	34194.7	43122.1
VBar [mL/g[	0.73642	0.73981
Hydration [nm]	0.39	0.38
Charge at pH 7.4	- 9.45	2.85
A (280 nm)	0.534	0.134

4.2.4.7 Small Angle X-Ray Scattering

To gain insight in structure and oligomer composition of the proteins of interest, small angle X-ray scattering experiments were performed, at EMBL beamline P12, PETRA III, Hamburg. For data collection, a PILATUS 6M detector was used, which had a distance of 3 m to the sample chamber. The X-ray beam had a wavelength of 1.24 Å.

For batch mode experiments each measurement was performed at 15°C. For each sample, 50 frames were collected with an exposure time of 0.45 s. Additionally, before and after each measurement the exact same buffer was measured and subtracted to reduce background contribution.

To study the time resolved complex formation of UBA5, UFM1 and UFC1 the stop-flow syringe apparatus from SAXS facility of EMBL Hamburg was used. This device enables very fast mixing of two samples. Hereby the mixing of each binding partner was used to start the complex formation reaction. For this experiments the following proteins samples solved in Complex buffer were measured: 6.03 g/mL UBA5, 1.40 mg/mL UFM1 and 2.18 mg/mL UFC1. Before the measurement all samples were centrifuged at 22000 g for 10 minutes. For each measurement 100 µL of two components (UBA5 and UFM1 or UFC1 and UFM1) were mixed with the stopped flow device at a flow rate of 4 mL/min. Each measurement was repeated 8 times by collecting 60 frames each frame for 50 µs. with a delay of 250 µs between each frame to cover a time range of 15 s.

Data was analyzed with the ATSAS software package version 2.8[170] from EMBL. With primusqt, the guinier approximation and Kratky plot was extracted to calculate radius of gyration, maximal distance distribution D_{max} and the molecular weight from the Porod volume, which is calculated by dividing the Porod volume by 1.66[171]. For generation of ab initio structural models, the software DAMMIF was used with default parameters and 20 cycles of refinement. Further analysis was performed with CRYSOL[109] and OLIGOMER[111] to investigate the oligomer composition at different concentrations. For this, different models were created with ITASSER[172] and CORAL[173] from the UBA5-UFM1 crystal structure (pdb entry: 5iaa): a UBA5 monomer and dimer, as well as a UBA5-UFM1 heterodimer, -trimer and -tetramer.

For the CRYSOL analysis default settings were used, except the number of harmonics were set to 50 and during the fit a constant was subtracted. Additionally, the first 50 data points from the scattering curves were removed. The program calculates a theoretical scattering curve from the used structure model and fits it to the experimental scattering curve. The quality of the fit is evaluated by the Chi^2 value for each fit.

For the OLIGOMER analysis default parameters were used, except the following: number of harmonics were set to 50, addition of additional constant for better fits and performing a comparative analysis to take into account that specific models seem not to be present at higher concentrations (UBA5 monomer and UBA5 heterodimer at a concentration higher than 1 mg/ml and 2 mg/ml respectively).

4.2.5 Exchange of Bound Nucleotides and Loading Status Check

4.2.5.1 Docking

Docking is a bioinformatic tool for identifying and predicting binding interactions of proteins and small molecules. This is essential for drug discovery and protein engineering as it can reduce screening times for potential targets and can give important information to optimize the ligand or binding site properties [143] [174] This method requires both the structure of the ligand and the protein [175]. The different programs vary in their approach and used algorithm. One approach is the blind docking, where the whole protein surface is scanned for possible binding sites [176] by taking several parameters, for example solvent molecules and the van der Waals radius, into account

[177]. The possible identified binding sites are then scored by a scoring functions, mostly be determining the free binding energy ΔG of the ligand to the binding pocket [178]. This approach is used e.g. by the software SwissDock [179].

Another approach can be used if the binding site is known, and it is investigated if a specific ligand or a derivative of a known ligand fits into the binding pocket. Hereby, the binding site with important residues is defined before the docking procedure, as well as a ligand in an idealized conformation [180]. During the docking, the ligand is broken into fragments with rotatable bonds. The conformation of each fragment is optimized against a scoring function before the ligand is reconstructed into the binding pocket [181]. The scoring functions also includes the determination of the optimal free energy value [182]. This approach is used e.g. in the program SYBYL [180].

Other methods use a sequence based approach by identifying binding sites through sequence similarity [183] or with matrices constructed by superimposing known amino acid residues and ligand interactions and using it as a library for docking procedures and binding site prediction [184].

For docking studies of NPE-ATP to UBA5, two different software were used. At first the online tool SwissDock was used to generate different models with possible binding sites of NPE-ATP. For this purpose, the UBA5 57-329 pdb model (3H8V) was used and NPE as 3D mol2 file with added hydrogens. For docking default parameters were used.

As an alternative, the program SYBYL was used to study the possible binding of NPE-ATP to the ATP binding pocket. For this, also the pdb model (3H(V) was loaded into the software. Water molecules present in the model were removed, hydrogens added to the residues, as well as possible charges (protein and ligand). For both, the ligand and the protein, a conformation with a minimum energy was calculated by the software with a dielectric function with 20 as constant and 500 iterations (50 iterations for the ligand. To perform the docking analysis, a 3 Å radius around the binding pocket was selected and the ligand placed inside. SYBYL calculated from the initial setting a possible conformation of the NPE-ATP in the binding pocket.

4.2.5.2 Nucleotide Exchange

Caged compounds are inert derivatives of biomolecules carrying a photocleavable protecting group. Upon illuminating the compound with light, the compound is promoted to an excited state. This highly reactive state can either be returned to the ground state through release of energy or release of stable products: this is called the dark reaction[185]. While the excitation is in the range of ns, the dark reaction can be rather slow in the range of ms [185]. This can be quantified by the quantum yield, which is defined as the amount of product formed in relation to the number of photons absorbed [185]. The reaction is usually based on photoisomerization and one of the most common protecting group is the 1-(2-nitrophenly)-ethyl group [186].

One of the first caged molecules studied was ATP (**Fig. *2-10***); it is also one of the best-studied caged compounds when it comes to quantum yield and release rates [187]. Moreover, ATP is one of the most used ligands in biochemical reactions, which makes it an interesting target for studying dynamics and kinetics with caged ATP: examples range from examining the force generation and time cycle of kinesin movement [188] on microtubule over ATPase activation [189] up to arterial contraction [190]. Furthermore, caged compounds can also be used in crystallography since time-resolved experimental setups become increasingly available.

The exchange of nucleotides bound to protein with a specific derivative depends on incubation of the protein with EDTA and its ability to bind divalent cations. Since nucleotide binding proteins also natively contain magnesium ions as counterions for the nucleotides, EDTA can be used to bind magnesium and lowering the affinity of the nucleotide to the protein. By adding an excess of the derivative, the nucleotide can be exchanged.

Three different methods were tested at the beginning. In Method A UBA5 was incubated in SEC buffer with addition of 15 mM EDTA, 150 mM $(NH_4)_2SO_4$, 1 mM NPE-ATP for 1 h at RT. In method B UBA5 was incubated with SEC buffer with addition of 5 mM $MgCl_2$, 3 mM DTT, 20 mM EDTA and 1 mM NPE-ATP for 1To exchange the ATP in UBA5 against NPE-ATP, 20µM of a UBA5 sample was incubated with 20 mM EDTA and 100 µM NPE-ATP, while dialyzing against 150 mM NaCl, 20 mM Tris pH 7.5, 20 mM EDTA and 1 µM NPE-ATP for 4 h. To remove the EDTA afterwards, 100 µM NPE-ATP were added to the

sample and it was dialyzed against 150 mM NaCl, 20 mM Tris pH 7.5 and 1 µM NPE-ATP. After dialysis, 100 µM NPE-ATP, 5 mM $MgCl_2$ and 100 µM $ZnCl_2$ were added to the UBA5 sample.

Furthermore, the stability of NPE-ATP against light illumination was tested. Therefore, NPE-ATP was illuminated with daylight at RT over a consecutive time of 16 h. After several timepoints, samples were prepared for analysis with RP-HPLC. As comparison, an ATP solution was treated in the same way. Additionally, NPE-ATP was also stored for 16 h at RT in the dark with sample preparation at the same time points as the light treatment.

4.2.5.3 RP-HPLC

High Performance Liquid Chromatography (HPLC) is an analytical method for separation of mixtures by a specific property, usually by the polarity of the involved analytes. The separation is accomplished by the different solubility of the analytes in a mobile and a stationary phase. With a higher solubility in the mobile phase, molecules are eluted faster from the chromatographic system. Molecules with a higher solubility in the stationary phase have higher elution times. The Reversed Phase High Performance Liquid Chromatography (RP-HPLC) uses a stationary phase with aliphatic chains with a low polarity. The mobile phase consists of a buffer with a high polarity.

The analysis of nucleotide concentration bound to UBA5 was performed on an Agilent System with a 1 mL RP18 column. The gradient used for elution is shown in. Prior to analysis of nucleotide concentration bound to UBA5, the protein sample was denaturated for 10 minutes at 95 °C. After the denaturation step, the sample was centrifuged for 10 minutes at 17000 g. After transferring the supernatant to new tube, the centrifugation step was repeated. 50 µL of the supernatant were applied on a RP18 column. The nucleotide elution was detected at 260 nm. For determination of nucleotide concentration, ATP or NPE-ATP in a range of 10 to 100 µM were dissolved directly in TBAB buffer and 50 µL were applied on the column.

Tab. 4-35 Gradient program for the separation of nucleotides with RP-HLPC

time [min]	acetonitrile [%]	flow [ml/min]
0.0	0	0.15
0.5	0	
10.0	30	
15.0	30	
15.1	0	
25.0	0	

4.2.6 Structural Analysis of Proteins

4.2.6.1 Protein Crystallization

UBA5 57-329 was crystallized with the hanging drop method on 24 well Linbro plates. For crystallization, a concentration of 8 mg/ml was used. The protein was solved in SEC Buffer. Before setting up crystallization trials, the sample was centrifuged for 1 h at 17000 g. The reservoir solution had a volume of 1000 µL. and consisted of 0.3 M $(NH_4)_2SO_4$, 0.1 M Na-citrate with varying pH from 5.9 to 6.2 and $LiSO_4$ with varying concentration from 0.8 to 1.3 M. 1 µL of protein solution was mixed with 1 µL of reservoir solution and 0.5 µL of seed stock or by seek streaking from seed stock on a siliconized coverslip. Seed stock was prepared from previous crystallization trials from precipitate by adding 10 µL of mother liquor to the drop and crushing the precipitate with a glass tip. The seed stock was stored at -20 °C. To minimize evaporation the coverslip and reservoir was sealed with silicon grease. Plates were stored at 16 °C for 4 d.

For screening of crystallization conditions of UBA5-UFM1 complex with NPE-ATP, commercial screens Compass and JCSG were used. Sitting drop 96 well plates were pipetted with the Oryx pipetting robot witch 10 µL reservoir solution. One the first position 0.3 µL reservoir solution were mixed with 0.3 µL protein solution with concentration of 15 mg/mL. On the second position 0.3 µL reservoir solution were mixed with 0.2 µL protein solution and 0.1 µL seed stock, prepared from UBA5 57-329 crystals. Additionally, with JCSG and Compass screens microbatch crystallization under oil were pipetted by mixing 0.3 µL screen condition with 0.3 µL protein solution and covering the drop with 100 µL Al's Oil. Additionally, Linbro plates were prepared with varying conditions of tascimat pH 7 between 6 and 12 % and PEG3350 between 12 and 22 % on

one plate. A second plate was prepared with 0.1 M HEPES pH 7.5, PEG6000 between 6 and 16 % and MPD between 3 and 9 %. All plates were sealed and stored at 16 °C. To protect the NPE-ATP from decaging, all plates were additionally covered with aluminum foil.

4.2.6.2 X-Ray Diffraction and structure solution

Prior to illuminating the crystal with X-rays, the crystal was cryoprotected by shortly dipping it in mother liquor with addition of 9 % sucrose, 2 % glucose, 8 % glycerol and 8 % ethylene glycol. Dataset of 3600 images was recorded on PILATUS6M detector at P13 beamline, EMBL, Hamburg with a wavelength of 0.9763 Å and a detector distance of 2.89 m at 100K. For each image, the mounted crystal was turned by 0.1 °.

Datasets were processed with XDS program package[191] at resolutions of 2.2 Å in the space group $P3_221$. The structures was determined by Molecular replacement with the program MOLREP[192] with a solvent free structure of UBA5 (pdb id: 3H8V) as a search model. With MATTHEWS[193] of the ccp4 software package, the Matthews coefficient and numbers of search models were extracted by crystal content analysis. The structure was refined with REFMAC5[194] and manual model rebuilding with COOT[195]. Finally, the quality of the model was checked with the validation tools of the program COOT. Figures were generated using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) [196].

Literature

[1] S. Fuchs, A. G. Kikhney, R. Schubert, C. Kaiser, E. Liebau, D. I. Svergun, C. Betzel, M. Perbandt, *Journal of structural biology* **2021**, *213*, 107796.

[2] G. Goldstein, M. Scheid, U. Hammerling, E. A. Boyse, D. H. Schlesinger, H. D. Niall, *Proceedings of the National Academy of Science* **1975**, *72*, 11.

[3] H. Busch, I. L. Goldknopf, *Proceedings of the National Academy of Science* **1977**, *74*, 864.

[4] A. Hershko, *Trends in Biochemical Sciences* **1996**, *21*, 445.

[5] A. Hershko, A. Ciechanover, H. Heller, A. L. Haas, I. A. Rose, *Proceedings of the National Academy of Science* **1980**, *77*, 1783.

[6] A. S. Baldwin, *Annual Reviews Immunology* **1996**, *14*, 649.

[7] C.-M. Fan, T. Maniatis, *Nature* **1991**, *354*, 395.

[8] V. J. Palombella, O. J. Rando, A. L. Goldberg, T. Maniatis, *Cell* **1994**, *78*, 773.

[9] M. R. Campanero, E. K. Flemington, *Proceedings of the National Academy of Science* **1997**, *94*, 2221.

[10] F. Hofmann, F. Martelli, D. M. Livingston, Z. Wang, *Genes & development* **1996**, *10*, 2949.

[11] Maki Carl G., J. M. Huibregtse, P. M. Howley, *Cancer Research* **1996**, *56*, 2649.

[12] R. Y. Hampton, J. Rine, *Journal of Cell Biology* **1994**, *125*, 299.

[13] J. Delic, M. Morange, H. Magdelenat, *Molecular and Cellualr Biology* **1993**, *13*, 4875.

[14] V. Chau, J. W. Tobias, A. Bachmair, D. Marriot, D. J. Ecker, A. K. Gonda, A. Varshasky, *Science* **1989**, *243*, 1576.

[15] A. Ciechanover, H. Heller, R. Katz-Etzion, A. Hershko, *Proceedings of the National Academy of Science* **1981**, *78*, 761.

[16] A. Hershko, H. Heller, S. Elias, A. Ciechanover, *Journal of Biological Chemistry* **1983**, *258*, 8206.

[17] M. Rolfe, *Ubiquitin and the Biology of the Cell*, **1998**.

[18] R. D. Mueller, H. Yasuda, C. L. Hatch, W. M. Bonner, E. M. Bradbury, *Journal of Biological Chemistry* **1985**, *260*, 5147.

[19] A. G. van der Veen, H. L. Ploegh, *Annual review of biochemistry* **2012**, *81*, 323.

[20] Matunis, Michael J., Coutavas, Elias, G. Blobel, *Journal of Cell Biology* **1996**, *135*, 1457.

[21] F. Golebiowski, I. Matic, M. H. Tatham, C. Cole, Y. Yin, A. Nakamura, J. Cox, G. J. Barton, M. Mann, R. T. Hay, *Science Signaling* **2009**, *2*, 1.

[22] A. Carbia-Nagashima, J. Gerez, C. Perez-Castro, M. Paez-Pereda, S. Silberstein, G. K. Stalla, F. Holsboer, E. Arzt, *Cell* **2007**, *131*, 309.

[23] Y. Han, C. Huang, X. Sun, B. Xiang, M. Wang, E. T. H. Yeh, Y. Chen, H. Li, G. Shi, H. Cang et al., *The Journal of biological chemistry* **2010**, *285*, 12906.

[24] W. Wei, P. Yang, J. Pang, S. Zhang, Y. Wang, M.-H. Wang, Z. Dong, J.-X. She, C.-Y. Wang, *Biochemical and biophysical research communications* **2008**, *375*, 454.

[25] C.-Y. Wang, P. Yang, M. Li, F. Gong, *Biochemical and biophysical research communications* **2009**, *381*, 477.

[26] Z.-Q. Pan, A. Kentsis, D. C. Dias, K. Yamoah, K. Wu, *Oncogene* **2004**, *23*, 1985.

[27] D. T. Huang, O. Ayrault, H. W. Hunt, A. M. Taherbhoy, D. M. Duda, D. C. Scott, L. A. Borg, G. Neale, P. J. Murray, M. F. Roussel et al., *Molecular cell* **2009**, *33*, 483.

[28] D. J. Lenschow, C. Lai, N. Frias-Staheli, N. V. Giannakopoulos, A. Lutz, T. Wolff, A. Osiak, B. Levine, R. E. Schmidt, Garcia-Sastre et al., *Proceedings of the National Academy of Science* **2007**, *104*.

[29] A. Snyder, Z. Alsauskas, P. Gong, P. E. Rosenstiel, M. E. Klotman, P. E. Klotman, M. J. Ross, *Journal of virology* **2009**, *83*, 11983.

[30] M. Nakamura, S. Shimosaki, *The FEBS journal* **2009**, *276*, 6355.

[31] T. Kirisako, Y. Ichimura, H. Okada, Kabeya, N. Mizushima, T. Yoshimori, M. Ohsumi, T. Takao, T. Noda, Y. Ohsumi, *Journal of Cell Biology* **2000**, *151*, 263.

[32] I. Tanida, N. Mizushima, M. Kiyooka, M. Ohsumi, T. Ueno, Y. Ohsumi, E. Kominami, *Molecular Biology of the Cell* **1999**, *10*, 1367.

[33] L. M. Iyer, A. M. Burroughs, L. Aravind, *Genome Biology* **2006**, *7*, 1.

[34] S. Leidel, P. G. A. Pedrioli, T. Bucher, R. Brost, M. Costanzo, A. Schmidt, R. Aebersold, C. Boone, K. Hofmann, M. Peter, *Nature* **2009**, *458*, 228.

[35] Vijay-Kumar, Senadhi, C. E. Bugg, W. j. Coo, *Journal of Molecular Biology* **1987**, *194*, 531.

[36] M. Komatsu, T. Chiba, Tatsumi, kanako, Iemura, Shun-ichiro, I. Tanida, N. Okazaki, T. Ueno, E. Kominami, K. Tanaka, *EMBO Journal* **2004**, *23*, 1977.

[37] K. Tatsumi, Y. Sou, N. Tada, E. Nakamura, S. Iemura, T. Natsume, S. H. Kang, C. H. Chung, M. Kasahara, E. Kominami et al., *The Journal of biological chemistry* **2010**, *285*, 5417.

[38] S. H. Kang, G. R. Kim, M. Seong, S. H. Baek, J. H. Seol, O. S. Bang, H. Ovaa, K. Tatsumi, M. Komatsu, K. Tanaka et al., *The Journal of biological chemistry* **2007**, *282*, 5256.

[39] C. Chen, E. Itakura, K. P. Weber, R. S. Hegde, M. de Bono, *PLoS genetics* **2014**, *10*, e1004082.

[40] H.-J. Kim, J.-Y. Yi, H.-S. Sung, D. D. Moore, B. H. Jhun, Y. C. Lee, J. W. Lee, *Molecular and Cellular Biology* **1999**, *19*, 6323.

[41] K. A. Green, J. S. Carroll, *Nature Reviews Cancer* **2007**, *7*, 713.

[42] C. m. Klinge, *Nucleic Acids Research* **2001**, *29*, 2905.

[43] H. M. Yoo, J. H. Park, Y. J. Jeon, C. H. Chung, *Frontiers in endocrinology* **2015**, *6*, 36.

[44] A. Azfer, J. Niu, L. M. Rogers, F. M. Adamksi, P. E. Kolattukudy, *American Journal of Physiology-Heart and Heart Circulatory Physiology* **2006**, *291*, 1411.

[45] Y. Zhang, M. Zhang, J. Wu, G. Lei, H. Li, *PloS one* **2012**, *7*, e48587.

[46] K. Lemaire, R. F. Moura, M. Granvik, M. Igoillo-Esteve, H. E. Hohmeier, N. Hendrickx, C. B. Newgard, E. Waelkens, M. Cnop, F. Schuit, *PloS one* **2011**, *6*, e18517.

[47] X. Hu, Q. Pang, Q. Shen, H. Liu, J. He, J. Wang, J. Xiong, H. Zhang, F. Chen, *International journal of molecular medicine* **2014**, *33*, 1539.

[48] Y. Cai, W. Pi, S. Sivaprakasam, X. Zhu, M. Zhang, J. Chen, L. Makala, C. Lu, J. Wu, Y. Teng et al., *PLoS genetics* **2015**, *11*, e1005643.

[49] K. Tatsumi, H. Yamamoto-Mukai, R. Shimizu, S. Waguri, Y. Sou, A. Sakamoto, C. Taya, H. Shitara, T. Hara, C. H. Chung et al., *Nature communications* **2011**, *2*, 181.

[50] S. Gannavaram, P. S. Connelly, M. P. Daniels, R. Duncan, P. Salotra, H. L. Nakhasi, *Molecular microbiology* **2012**, *86*, 187.

[51] H. Liu, J. Li, B. Tillman, B. A. French, S. W. French, *Experimental and molecular pathology* **2014**, *97*, 81.

[52] M. D. Rubio, K. Wood, V. Haroutunian, J. H. Meador-Woodruff, *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **2013**, *38*, 1910.

[53] H. Sasakawa, E. Sakata, Y. Yamaguchi, M. Komatsu, K. Tatsumi, E. Kominami, K. Tanaka, K. Kato, *Biochemical and biophysical research communications* **2006**, *343*, 21.

[54] P. Padala, W. Oweis, B. Mashahreh, N. Soudah, E. Cohen-Kfir, E. A. Todd, C. E. Berndsen, R. Wiener, *Scientific reports* **2017**, *7*, 508.

[55] T. Dou, S. Gu, J. Liu, F. Chen, L. Zeng, L. Guo, Y. Xie, Y. Mao, *Molecular biology reports* **2005**, *32*, 265.

[56] J.-P. Bacik, J. R. Walker, M. Ali, A. D. Schimmer, S. Dhe-Paganon, *The Journal of biological chemistry* **2010**, *285*, 20273.

[57] B. A. Schulman, J. W. Harper, *Nature reviews. Molecular cell biology* **2009**, *10*, 319.

[58] W. Oweis, P. Padala, F. Hassouna, E. Cohen-Kfir, D. R. Gibbs, E. A. Todd, C. E. Berndsen, R. Wiener, *Cell reports* **2016**, *16*, 3113.

[59] B. Mashahreh, F. Hassouna, N. Soudah, E. Cohen-Kfir, R. Strulovich, Y. Haitin, R. Wiener, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **2018**, *32*, 2794.

[60] S. Habisov, J. Huber, Y. Ichimura, M. Akutsu, N. Rogova, F. Loehr, D. G. McEwan, T. Johansen, I. Dikic, V. Doetsch et al., *The Journal of biological chemistry* **2016**, *291*, 9025.

[61] S. Xie, *Acta crystallographica. Section F, Structural biology communications* **2014**, *70*, 1093.

[62] E. Colin, J. Daniel, A. Ziegler, J. Wakim, A. Scrivo, T. B. Haack, S. Khiati, A.-S. Denommé, P. Amati-Bonneau, M. Charif et al., *American journal of human genetics* **2016**, *99*, 695.

[63] R. Duan, Y. Shi, L. Yu, G. Zhang, J. Li, Y. Lin, J. Guo, J. Wang, L. Shen, H. Jiang et al., *PloS one* **2016**, *11*, e0149039.

[64] T. Mizushima, K. Tatsumi, Y. Ozaki, T. Kawakami, A. Suzuki, K. Ogasahara, M. Komatsu, E. Kominami, K. Tanaka, T. Yamane, *Biochemical and biophysical research communications* **2007**, *362*, 1079.

[65] M. Homrich, H. Wobst, C. Laurini, J. Sabrowski, B. Schmitz, S. Diestel, *Experimental cell research* **2014**, *324*, 192.

[66] A. M. Burroughs, M. Jaffee, L. M. Iyer, L. Aravind, *Journal of structural biology* **2008**, *162*, 205.

[67] M. Zhang, X. Zhu, Y. Zhang, Y. Cai, J. Chen, S. Sivaprakasam, A. Gurav, W. Pi, L. Makala, J. Wu et al., *Cell death and differentiation* **2015**, *22*, 1922.

[68] H. Shiwaku, N. Yoshimura, T. Tamura, M. Sone, S. Ogishima, K. Watase, K. Tagawa, H. Okazawa, *The EMBO journal* **2010**, *29*, 2446.

[69] B. Rupp, *Biomolecular crystallography*. *Principles, practice, and application to structural biology*, Garland Science, New York, NY, **2010**.

[70] P. G. Vekilov, A. Feeling-Taylor, S.-T. Yau, D. Petsev.

[71] A. McPherson, *John Wiley & Sons* **1982**.

[72] W. Zielenkiewicz, *J Therm Anal Calorim* **2008**, *92*, 105.

[73] a) A. McPherson, *Journal of Biological Chemistry* **1976**, *251*, 6300; b) B. R. Reid, G. L. E. Koch, Y. Boulanger, B. S. Hartley, D. M. Blow, *Journal of Biological Chemistry* **1973**, *80*, 199.

[74] J. R. Kimmel, E. L. Smith, *Journal of Biological Chemistry* **1954**, *207*, 515.

[75] F. R. Salemme, *Archives of the Biochemistry and Biophysics* **1972**, *151*, 533.

[76] C. Sauter, J. D. Ng, B. Lorber, G. Keith, P. Brion, M. W. Hosseini, J.-M. Lehn, R. Giegé, *Journal of Crystal Growth* **1999**, *196*, 365.

[77] C. Ostermeier, H. Michel, *Current Opinion in Structural Biology* **1997**, *7*, 697.

[78] W. Friedrich, P. Knipping, M. Laue, *Ann. Phys.* **1913**, *346*, 971.

[79] A. H. Compton, *Second Series* **1923**, *21*.

[80] M. von Laue, *Concerning the detection of X-ray interferences*, **1920**.

[81] W. H. Bragg, W. L. Bragg, *Proceedings of the Royal Society* **1913**, *88*, 428.

[82] P. P. Ewald, *Dispersion und Doppelbrechung von Elektronengittern (Kristallen)*, Müller, Marburg, **1912**.

[83] W. H. Miller, *A treatise on crystallography*, J. & J. J. Deighton, Cambridge, **1839**.

[84] C. Colliex, J. M. Cowley, S. L. Dudarev, M. Fink, J. Gjønnes, R. Hilderbrandt, A. Howie, D. F. Lynch, L. M. Peng, G. Ren et al. in *International Tables for Crystallography* (Hrsg.: H. Fuess), Springer Netherlands, Amsterdam, **2006**, S. 259–429.

[85] S. C. Wallwork, *International Union of Crystallography* **1980**, 1.

[86] G. Rhodes, *Crystallography made crystal clear*. *A guide for users of macromolecular models,* 3. Aufl., Elsevier, Amsterdam, **20]09**.

[87] C.-J. Chen, J. Rose, M. Gary Newton, Z.-J. Liu, B.-C. Wang in *Modern Protein Chemistry* (Hrsg.: G. Howard, W. Brown), CRC Press, **2001**, S. 137.

[88] S. Garcia-Granda, J. M. Montejo-Bernardo in *Reference module in chemistry, molecular sciences and chemical engineering* (Hrsg.: J. Reedijk), Elsevier, Amsterdam, **2013**.

[89] R. W. Harrison, *Journal of the Optical Society of America A* **1993**, *10*, 1046.

[90] D. W. Green, V. M. Ingram, M. F. Perutz, *The Royal Society* **1954**, *225*.

[91] J. M. Guss, E. A. Merrit, R. P. Phizackerley, B. Hedman, M. Murata, K. O. Hodgson, H. C. Freeman, *Science* **1988**, *241*, 806.

[92] S. Caticha-Ellis, *International Union of Crystallography* **1981**, 1.

[93] G. Friedel, *Comptes Rendus* **1913**, *157*, 1533.

[94] W. A. Hendrickson, M. M. Teeter, *Nature* **1981**, *290*, 107.

[95] W. A. Hendrickson, J. R. Horton, H. M. K. Murthy, A. Pahler, J. L. Smith in *Synchrotron Radiation in Structural Biology* (Hrsg.: E. H. Y. Chu, R. M. Sweet, A. D. Woodhead), Springer US, Boston, **1989**, S. 317–324.

[96] P. Evans, A. McCoy, *Acta crystallographica. Section D, Biological crystallography* **2008**, *64*, 1.

[97] A. L. Patterson, *Zeitschrift für Kristallographie - Crystalline Materials* **1935**, *90*.

[98] M. V. Petoukhov, D. I. Svergun, *The international journal of biochemistry & cell biology* **2013**, *45*, 429.

[99] A. G. Kikhney, D. I. Svergun, *FEBS letters* **2015**, *589*, 2570.

[100] C. A. Brosey, J. A. Tainer, *Current Opinion in Structural Biology* **2019**, *58*, 197.

[101] Guinier, André, Fournet, Gérard, *John Wiley & Sons* **1955**.

[102] A. Guinier, *Ann. Phys.* **1939**, *11*, 161.

[103] J. Trewhella, A. P. Duff, D. Durand, F. Gabel, J. M. Guss, W. A. Hendrickson, G. L. Hura, D. A. Jacques, N. M. Kirby, A. H. Kwan et al., *Acta crystallographica. Section D, Structural biology* **2017**, *73*, 710.

[104] G. Porod, *Kolloid Zeitschrift* **1951**, *124*, 83.

[105] P. Debye, *Ann. Phys.* **1915**, *351*, 809.

[106] Glatter, O., Kratky, O., *Small angle X-ray scattering - Scientific applications*, Academic Press, New York, NY, **1982**.

[107] P. Chacón, F. Morán, J.F. Díaz, E. Pantos, J.M. Andreu.

[108] D. I. Svergun, *Biophysical Journal* **1999**, *76*, 2879.

[109] D. Svergun, C. Barberato, M. H. J. Koch, *Journal of applied crystallography* **1995**, *28*, 768.

[110] C. L. Lawson, R. J. Hanson, *Solving least squares problems*, Society for Industrial and Applied Mathematics (SIAM 3600 Market Street Floor 6 Philadelphia PA 19104), Philadelphia, Pa, **1995**.

[111] P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, D. I. Svergun, *Journal of applied crystallography* **2003**, *36*, 1277.

[112] A. W. Malaby, S. Chakravarthy, T. C. Irving, S. V. Kathuria, O. Bilsel, D. G. Lambright, *Journal of applied crystallography* **2015**, *48*, 1102.

[113] N. M. Kirby, N. P. Cowieson, *Current Opinion in Structural Biology* **2014**, *28*, 41.

[114] J. B. Hopkins, R. E. Thorne, *J Appl Crystallogr* **2016**, *49*, 880.

[115] D. Tatchev, *J Appl Crystallogr* **2010**, *43*, 8.

[116] C. G. Shull, L. C. Roess, *Journal of Applied Physics* **1947**, *18*, 295.

[117] G. G. Stokes, *The Royal Society* **1852**, *142*.

[118] D. Bass, *Fluorescence Spectroscopy*, **2000**.

[119] A. Jablonski, *Nature* **1933**, *131*, 839.

[120] L. Stryer, *Science* **1968**, *162*.

[121] J. R. Lakowicz, *Principles of fluorescence spectroscopy,* 4. Aufl., Springer, New York, NY, **2010**.

[122] N. Shanker, S. L. Bane in *Methods in Cell Biology* (Hrsg.: J. Correia, H. W. Detrich), Elsevier, **2008**, S. 213–242.

[123] C. P. Toseland, *Journal of chemical biology* **2013**, *6*, 85.

[124] L. Reinhard, H. Mayerhofer, A. Geerlof, J. Mueller-Dieckmann, M. S. Weiss, *Acta crystallographica. Section F, Structural biology and crystallization communications* **2013**, *69*, 209.

[125] M. J. Sanderson, I. Smith, I. Parker, M. D. Bootman, *Cold Spring Harbor protocols* **2014**, *2014*, pdb.top071795.

[126] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, D. C. Prasher, *Science (New York, N.Y.)* **1994**, *263*, 802.

[127] A. B. T. Ghisaidoobe, S. J. Chung, *International journal of molecular sciences* **2014**, *15*, 22518.

[128] T. Svedberg, R. Fåhraeus, *Kolloid Zeitschrift* **1926**, *48*.

[129] D. A. Albright, J. W. Williams, *Journal of Physical Chemistry* **1967**, *71*, 2780.

[130] W. F. Stafford III, *Analyical Biochemistry* **1992**, *203*, 295.

[131] K. E. van Holde, R. L. Baldwin, *Journal of Physical Chemistry* **1958**, *62*.

[132] F. C. Bancroft, D. Freifelder, *Journal of Molecular Biology*, *54*.

[133] D. C. Teller, *Methods in Enzymology* **1973**, *27*, 349.

[134] A. P. Minton, *Analyical Biochemistry* **1990**, *190*, 1.

[135] G. Ralston, *Beckman* **1993**, 1.

[136] C. L. A. Schmidt, *J. Chem. Educ.* **1943**, *20*, 415.

[137] O. Lamm, *Zeitschrift für Physikalische Chemie* **1929**, *143A*.

[138] J. Lebowitz, M. S. Lewis, P. Schuck, *Protein Science* **2002**, *11*, 2067.

[139] P. Schuck, *Biophysical Journal* **2000**, *78*, 1606.

[140] R. M. Bock, N.-S. Ling, S. A. Morell, Lipton S. H., *Archives of the Biochemistry and Biophysics* **1956**, *62*, 253.

[141] S. Veltel, R. Gasper, E. Eisenacher, A. Wittinghofer, *Nature structural & molecular biology* **2008**, *15*, 373.

[142] P. Wingfield, *Current protocols in protein science* **2001**, *Appendix 3*, Appendix 3F.

[143] R. M. Anderson, K. J. Bitterman, J. G. Wood, O. Medvedik, D. A. Sinclair, *Nature* **2003**, *423*, 181.

[144] V. Zoete, M. A. Cuendet, A. Grosdidier, O. Michielin, *Journal of computational chemistry* **2011**, *32*, 2359.

[145] A. Grosdidier, V. Zoete, O. Michielin, *Journal of computational chemistry* **2011**, *32*, 2149.

[146] S. T. Harrison, *Biotechnological Advances* **1991**, *9*, 217.

[147] M. Bonora, S. Patergnani, A. Rimessi, E. de Marchi, J. M. Suski, A. Bononi, C. Giorgi, S. Marchi, S. Missiroli, F. Poletti et al., *Purinergic signalling* **2012**, *8*, 343.

[148] Jena Biosciences **2020**.

[149] M. Brylinski, *Chemical biology & drug design* **2018**, *91*, 380.

[150] A. C. Wallace, R. A. Laskowski, J. M. Thornton, *Protein engineering* **1995**, *8*, 127.

[151] D. A. Schneider, R. L. Gourse, *The Journal of biological chemistry* **2004**, *279*, 8262.

[152] J. Korlach, D. W. Baird, A. A. Heikal, K. R. Gee, Hoffman Grefory R., W. W. Webb, *Proceedings of the National Academy of Science* **2004**, *101*, 2800.

[153] R. A. Anderson, W. F. Bosron, Kennedy, F. Scott, Vallee, Bert, L., *Proceedings of the National Academy of Science* **1975**, *72*, 2989.

[154] I. L. Alberts, K. Nadassy, S. J. Wodak, *Protein Science* **1998**, *7*.

[155] N. J. Pace, E. Weerapana, *Biomolecules* **2014**, *4*, 419.

[156] T. O. Fischmann, A. Hruza, X. D. Niu, J. D. Fossetta, C. A. Lunn, E. Dolphin, A. J. Prongay, P. Reichert, D. J. Lundell, S. K. Narula et al., *Nature Structural Biology* **1999**, *6*, 233.

[157] James T. Vivian, Patrik R. Callis.

[158] Jeffery W. Walker, Gordon P. Reid, James A. McCray, and David R. Trentham, *Journal of the American Chemistry Society* **1988**, *110*, 7170.

[159] P. Klán, T. Šolomek, C. G. Bochet, A. Blanc, R. Givens, M. Rubina, V. Popik, A. Kostikov, J. Wirz, *Chemical reviews* **2013**, *113*, 119.

[160] J. L. Cole, J. W. Lary, T. P. Moody, T. M. Laue in *Methods in Cell Biology*, Elsevier, **2008**, S. 143–179.

[161] C. M. Jeffries, M. A. Graewert, D. I. Svergun, C. E. Blanchet, *Journal of synchrotron radiation* **2015**, *22*, 273.

[162] T. H. Kim, P. Mehrabi, Z. Ren, A. Sljoka, C. Ing, A. Bezginov, L. Ye, R. Pomès, R. S. Prosser, E. F. Pai, *Science (New York, N.Y.)* **2017**, *355*.

[163] J. R. Wiśniewski, M. Y. Hein, J. Cox, M. Mann, *Molecular & cellular proteomics : MCP* **2014**, *13*, 3497.

[164] E. de Nadal, G. Ammerer, F. Posas, *Nature reviews. Genetics* **2011**, *12*, 833.

[165] G. Thomas, *Nature reviews. Molecular cell biology* **2002**, 753.

[166] B. N. Khlodenko, *European Jorunal of Biochemistry* **2000**, *267*, 1583.

[167] F. J. Plou, J. L. Iborra, P. J. Halling, *Stability and Stabilization of Biocatalysts,* 1. Aufl., Elsevier textbooks, s.l., **1998**.

[168] B. Lorber, F. Fischer, M. Bailly, H. Roy, D. Kern, *Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology* **2012**, *40*, 372.

[169] M. Jerabek-Willemsen, T. André, R. Wanner, H. M. Roth, S. Duhr, P. Baaske, D. Breitsprecher, *Journal of Molecular Structure* **2014**, *1077*, 101.

[170] D. Franke, M. V. Petoukhov, P. V. Konarev, A. Panjkovich, A. Tuukkanen, H. D. T. Mertens, A. G. Kikhney, N. R. Hajizadeh, J. M. Franklin, C. M. Jeffries et al., *Journal of applied crystallography* **2017**, *50*, 1212.

[171] T. D. Grant, J. R. Luft, L. G. Carter, T. Matsui, T. M. Weiss, A. Martel, E. H. Snell, *Acta crystallographica. Section D, Biological crystallography* **2015**, *71*, 45.

[172] Jianyi Yang, Renxiang Yan, Ambrish Roy, Dong Xu, Jonathan Poisson, Yang Zhang.

[173] M. V. Petoukhov, D. Franke, A. V. Shkumatov, G. Tria, A. G. Kikhney, M. Gajda, C. Gorba, H. D. T. Mertens, P. V. Konarev, D. I. Svergun, *Journal of applied crystallography* **2012**, *45*, 342.

[174] B. K. Shoichet, S. L. McGovern, B. Wei, J. J. Irwin, *Current Opinion in Chemical Biology* **2002**, *6*, 439.

[175] A. M. Davis, S. A. St-Gallay, G. J. Kleywegt, *Drug discovery today* **2008**, *13*, 831.

[176] C. Hetényi, D. van der Spoel, *FEBS letters* **2006**, *580*, 1447.

[177] M. Hendlich, F. Rippmann, G. Barnickel, *Journal of Molecular Graphics and Modelling* **1997**, 359.

[178] C. Hetényi, D. van der Spoel, *Protein science : a publication of the Protein Society* **2011**, *20*, 880.

[179] A. Grosdidier, V. Zoete, O. Michielin, *Nucleic Acids Research* **2011**, *39*, W270-7.

[180] A. N. Jain, *Journal of medicinal chemistry* **2003**, *46*, 499.

[181] W. Welch, J. Ruppert, A. N. Jain, *Chemistry & Biology* **1996**, *3*, 449.

[182] A. R. Leach, B. K. Shoichet, C. E. Peishoff, *Journal of medicinal chemistry* **2006**, *49*, 5851.

[183] J. A. Capra, M. Singh, *Bioinformatics (Oxford, England)* **2007**, *23*, 1875.

[184] L. G. Trabuco, S. Lise, E. Petsalaki, R. B. Russell, *Nucleic Acids Research* **2012**, *40*, W423-7.

[185] A. M. Gurney, *Microelectrode Techniques* **1994**, 389.

[186] S. R. Adams, R. Y. Tsien, *Annual Reviews of Physiology* **1993**, *55*, 755.

[187] A. Barth, K. Hauser, W. Mäntele, j. E. T. Corrie, D. R. Trentham, *Journal of the American Chemistry Society* **1995**, *117*, 10311.

[188] H. Higuchi, M. Etsuko, Y. Inoue, T. Yanagida, *Proceedings of the National Academy of Science* **1997**, *94*.

[189] Lüpfert, Christian, Grell, Ernst, V. Pintschoivius, H.-J. Apell, F. Cornelius, R. J. Clarke, *Biophysical Journal* **2001**, *81*, 2069.

[190] N. Sjöblom-Widfeldt, A. Arner, H. Nilsson, *Journal of vascular research* **1993**, *30*, 38.

[191] W. Kabsch, *Acta crystallographica. Section D, Biological crystallography* **2010**, *66*, 125.

[192] A. Vagin, A. Teplyakov, *Journal of applied crystallography* **1997**, *30*, 1022.

[193] K. A. Kantardjieff, B. Rupp, *Protein science : a publication of the Protein Society* **2003**, *12*, 1865.

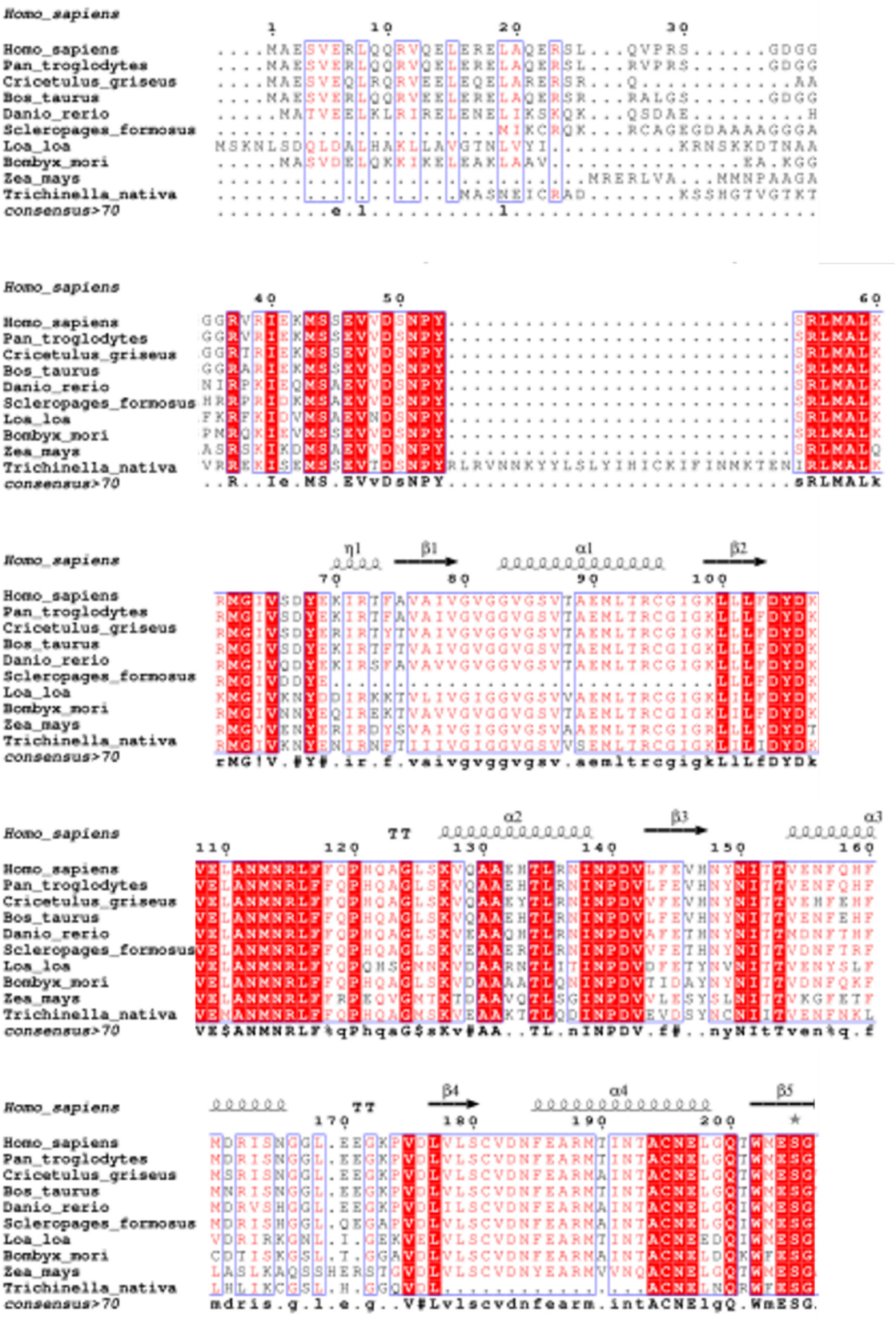
[194] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy et al., *Acta crystallographica. Section D, Biological crystallography* **2011**, *67*, 235.

[195] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta crystallographica. Section D, Biological crystallography* **2010**, *66*, 486.

[196] L. L. Schrödinger, *The PyMOL Molecular Graphics System, Version 1.8*.

Supplemental Material

Alignments



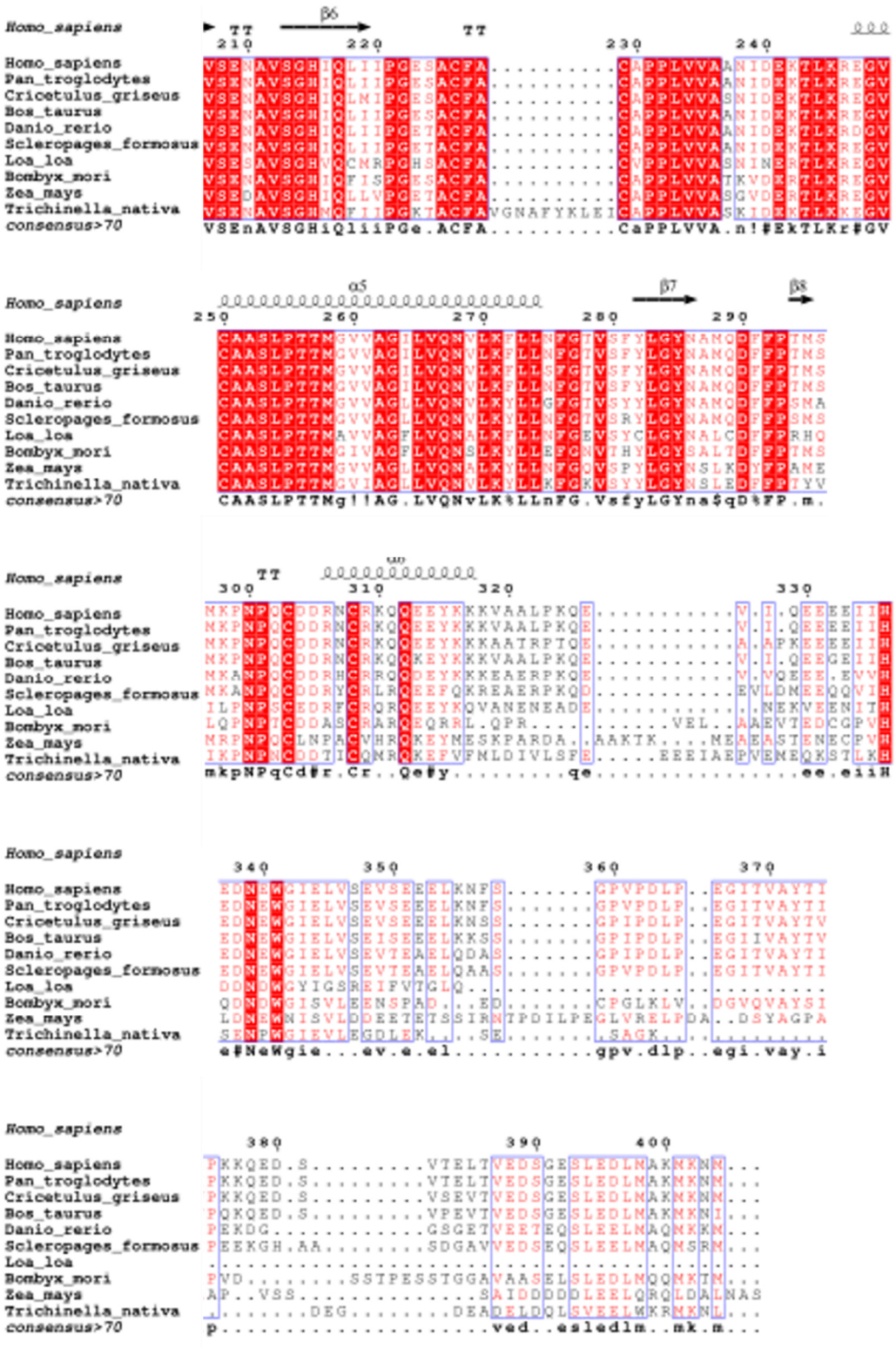


Fig. S 1 Multiple sequence alignment with Clustal Omega of UBA5 with secondary structure assignments and highlighted conserved residues in red. The proteins sequences from the following organisms were used: H. *sapiens*, P. *troglodytes*, C. *griseus*, B. *taurus*, D. *rerio*, S. *formosus*, L. *loa*, B. *moir*, Z. *mays*, T. *nativa*. Last line shows the consensus sequence.

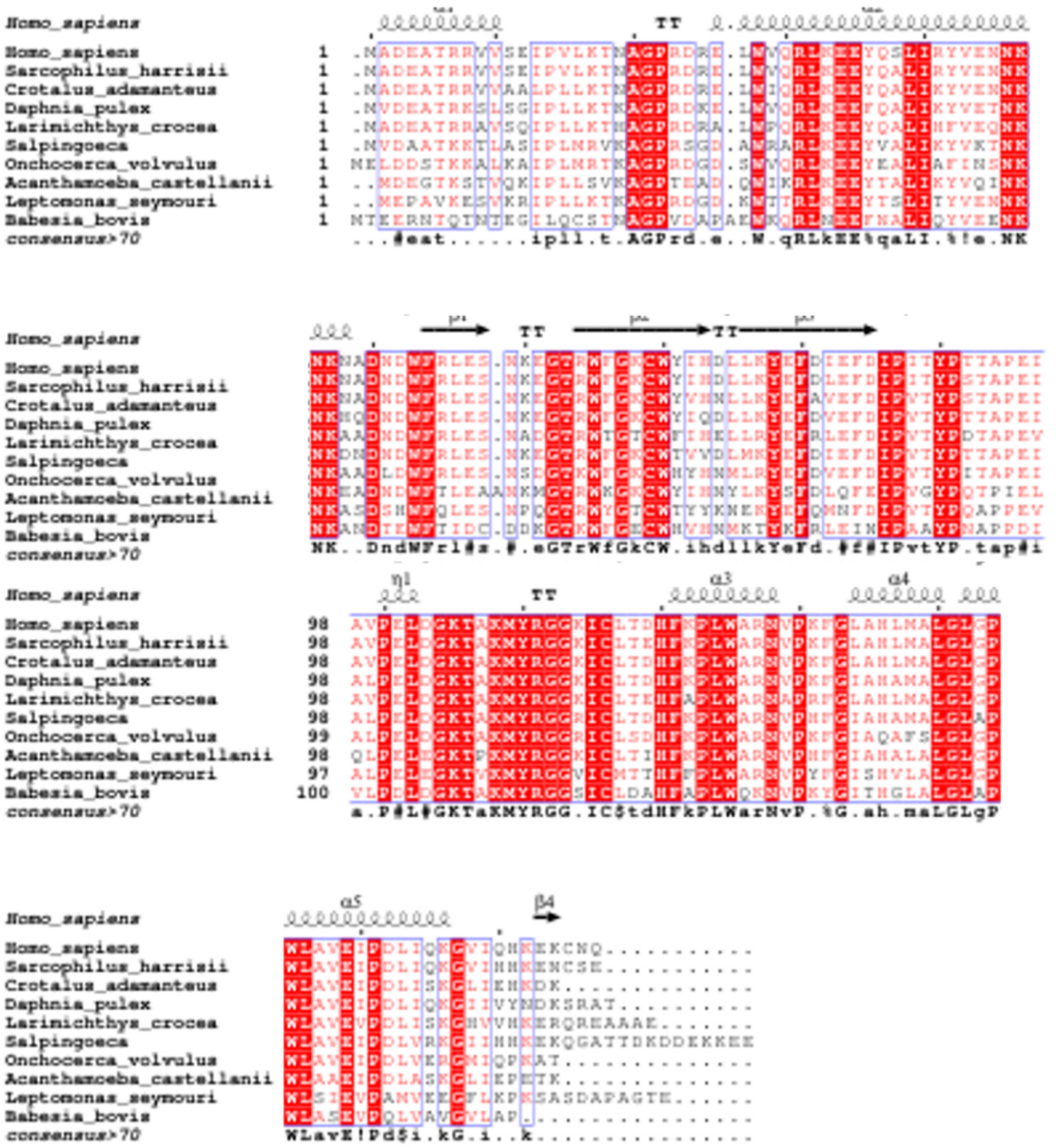


Fig. S 2 Multiple sequence alignment with Clustal Omega of UBA5 with secondary structure assignments and highlighted conserved residues in red. The proteins sequences from the following organisms were used: H. *sapiens*, S. *harrisii*, C. *adamanteus*, D. *pulex*, L. *crocea*, S. *salpingoeca*, O. *vo*lvulus, A. *castellanii*, L. *seymouri*, B. *bovis*. Last line shows the consensus sequence.

BestSel CD spectras

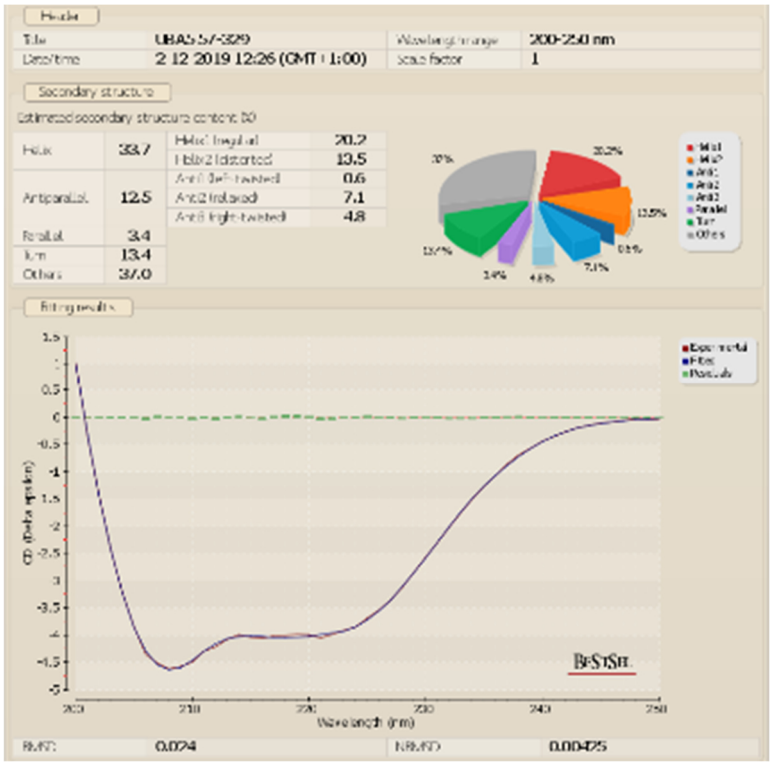


Fig. S 3 Summary of BeStSel result analysis of UBA5 57-329 CD spectrum.

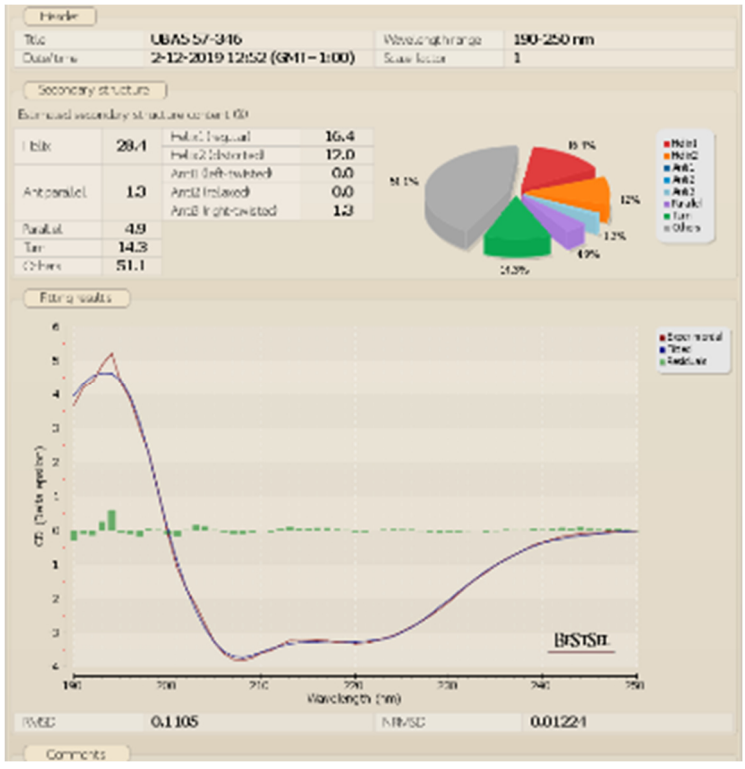


Fig. S 4 Summary of BeStSel result analysis of UBA5 57-346 CD spectrum.

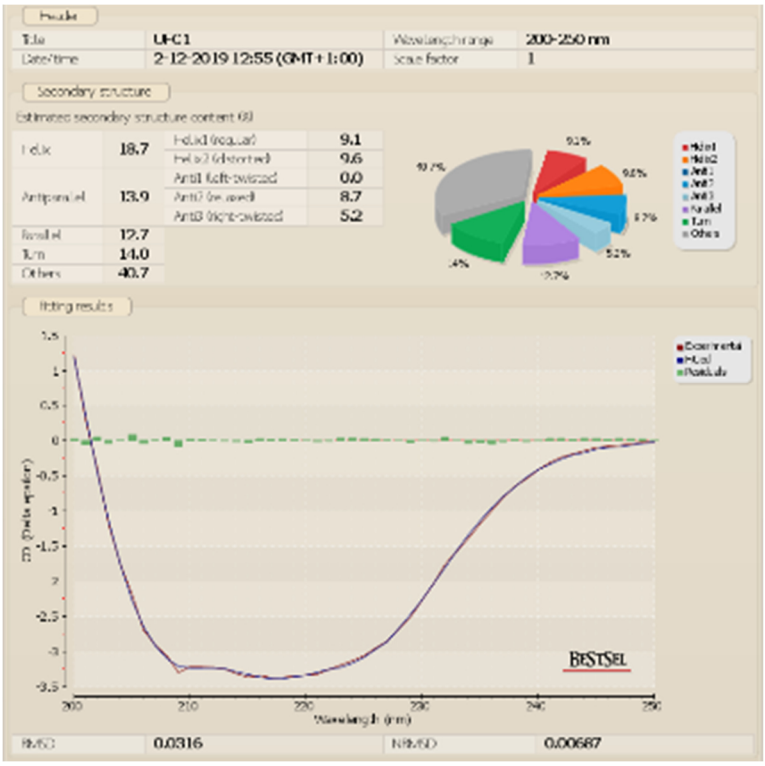


Fig. S 5 Summary of BeStSel result analysis of UFC1 CD spectrum.

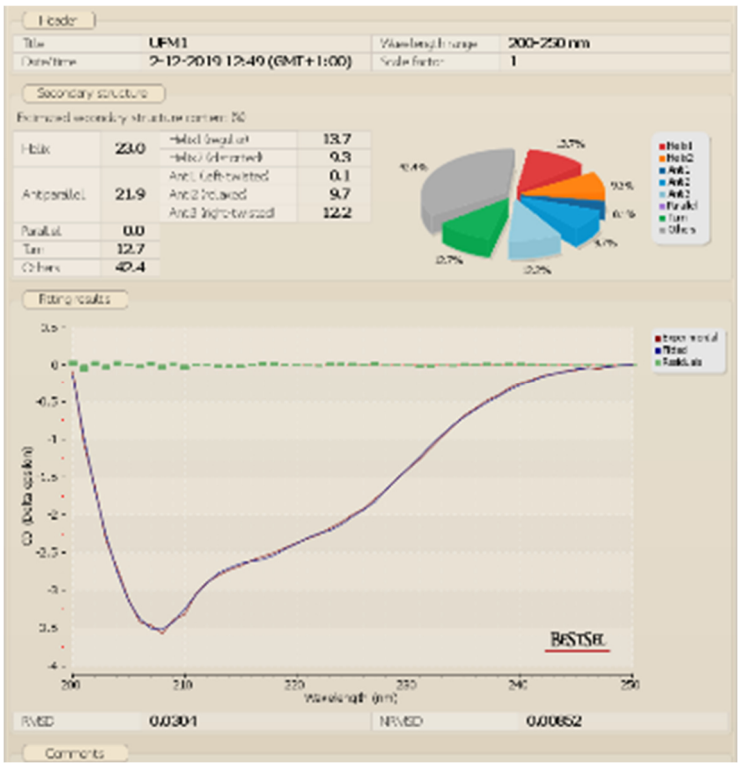


Fig. S 6 Summary of BeStSel result analysis of UFM1 CD spectrum.

SEC calibration

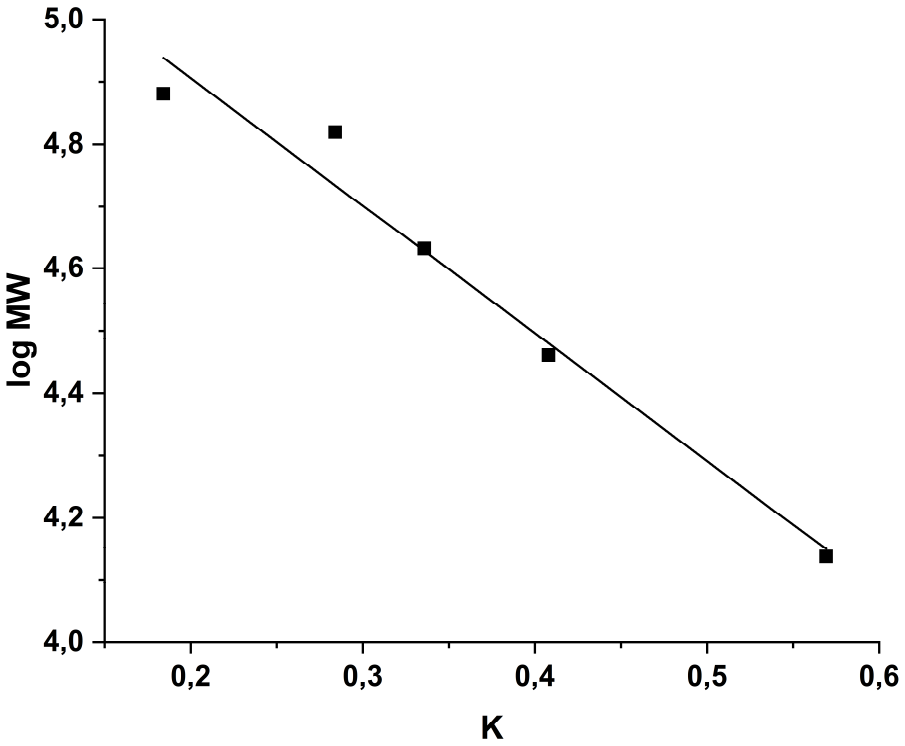


Fig. S 7 Calibration of SEC column Sup75. Depicted is logarithmic molecular weight over the partition coefficient K. For the calibration the following proteins were used: ribonuclease A, proteinase K, ovalbumin, albumin, conalbumin.

List of Chemicals

name	CAS-Nr.	GHS hazard	H-Statements	P-Statements
acetic acid	64-19-7	GHS02, GH205	226, 314	210, 280, 301, 330, 331, 303,361, 353, 305, 351, 38
acetonitrile	75-05-8	GH02	225, 302, 312,332, 319	210, 240, 302,352, 305, 351, 338, 403, 233
acrylamide	79-06-1	GHS06, GHS08	301, 312, 332,315, 317, 319, 340, 350, 361f, 372	201, 280, 302, 352, 304, 340, 305, 351, 338, 308, 310
adenosine triphosphate	56-65-5	-	-	-
agar agar	9002-18-0	-	-	-
agarose	9012-36-6	-	-	-
ammonium chloride	12125-02-9	GHS07	302, 319	305, 351, 338
ammonium persulfate	7727-54-0	GHS03, GHS07, GHS08	272, 302. 315. 317, 319, 334, 335	220, 261, 280, 305, 351, 338, 342, 311
ammonium sulfate	7783-20-2		-	-
ampicillin	69-52-3	GHS08	315, 317, 319, 334, 335	261, 280, 305, 351, 338, 342 , 311
bromophenol blue	62625-28-9		-	-
coomassie brilliant blue R250	6104-59-2		-	-
dimethyl sulfoxide	67-68-5		-	-
dithiothreitol	3483-12-3	GHS07	302, 315, 319, 335	261, 305, 351, 338
dipotassium hydrogen orthophosphate	7758-11-4		-	-
ethylenediaminetetra acetic acid disodium salt	60-00-4	GHS07	319, 332, 373	280, 304, 340, 312, 305, 351, 338, 337, 313
ethylene glycol	107-21-1	GHS02	302, 373	301, 312, 330
glucose	50-99-7		-	-
glutathione	70-18-8		-	-
glycerol	56-81-5		-	-
glycine	56-40-6		-	-
(4-2-hydroxyethyl)-1-piperazineethnesulfonic acid	7365-45-9		-	-
hydrochloric acid	7647-01-0	GHS05, GHS07	290, 314, 335	280, 303, 361, 353, 305, 351, 338, 310
imidazole	288-32-4	GHS05, GHS07, GHS08	360D, 302, 314	201, 280, 301, 330, 331, 305, 351, 338, 308, 310
isopropanol	67-63-0	GHS02, GHS07	225, 319, 336	210, 240, 305, 351, 338, 403, 233

isopropyl β-D-1-thiogalctopyranoside	367-93-1		-	-
kanamycin	64013-70-3	GHS08	360	201, 202, 280, 308, 313
lactose	63-42-3		-	-
lithium chloride	7447-41-8	GHS07	302, 315, 319	302, 352, 305, 351 338
magnesium chloride	7786-30-3		-	-
2-methyl-2,4-pentanediol	107-41-5	GHS07	315, 319	280, 305, 351, 338, 337, 313
nitrophenyl adenosine triphosphate	171800-68-3	GHS07	319	264, 280, 305, 351, 338, 337, 313
paraffin oil	8012-95-1	GHS08	304	301, 310, 331
peptone	91079-40-		-	-
polyethylene glycol	25322-68-3		-	-
potassium chloride	7447-40-7		-	-
potassium dihydrogen phosphate	7778-77-0		-	-
silicon oil	63148-62-9		-	-
sodium chloride	7647-14-5		-	-
sodium citrate	68-04-2		-	-
sodium dihydrogen phosphate	7558-80-7		-	-
sodium dodecyl sulfate	151-21-3	GHS02, GHS06	228, 302, 332, 315, 318, 335, 412	210, 261, 280, 301, 312, 330, 305, 351, 338, 310,370, 378
sodium fluoride	7681-49-4	GHS06	301, 315, 319	302, 352, 305, 351, 338,308,310
sodium hydroxide	1310-73-2	GHS05	290, 314	280, 310, 330, 331, 305, 351, 338, 308, 310
sucrose	57-50-1		-	-
tascimate	-	GHS05, GHS07	302, 315, 318, 319, 335	261, 280, 305, 351, 338
tetra-n-butylammonium bromide	1643-19-2	GHS07	302, 315, 319, 412	264, 270, 273, 301, 312, 302, 352, 305, 351, 338
tetramethyl ethylenediamine	110-18-9	GHS03, GHS05, GHS07	225, 332, 302, 314	210, 280, 305, 351, 338, 310
tris(hydroxymethyl)aminomethane	77-86-1	GHS07	315, 319, 335	261, 305, 351, 338
xylene cyanol	2650-17-1	GHS07	315, 319, 335	261, 305, 351, 338
yeast extract	8013-01-2		-	-
zinc chloride	7646-85-7	GHS05, GHS07, GHS09	302, 314, 410	273, 280, 301, 330, 331, 305, 351, 338, 308, 310

Crystalliaztion Screens

name	GHS hazard	H-Statements	P-Statements
ComPAS-Suite	GHS02, GH205, GHS06, GHS07, GHS08, GHS09	225, 301, 302, 315, 319, 331, 332, 335, 340, 350, 260FD, 373, 411	101, 201,270, 273, 280, 305, 351, 338, 309, 311, 313
acetonitrile	GH02, GHS05, GHS06, GHS07, GHS08, GHS09	225, 310, 312, 315, 318, 331, 335, 350, 411	101, 201, 270, 273, 280, 305, 351, 338, 309, 311, 313

List of Hazard Symbols

GHS02 *GHS03* *GHS05* *GHS06* *GHS07* *GHS08* *GHS09*

Danksagung

Zuallererst möchte ich mich sehr herzlich bei Prof. Dr. Christian Betzel bedanken. Zum einen für die Überlassung des hochinteressanten Themas sowie der Möglichkeit diese Dissertation innerhalb seiner Arbeitsgruppe durchführen zu können. Zum anderen bedanke ich mich für die konsequente und motivierende Betreuung während des gesamten Betreuungszeitraums.

Ein weiter großer Dank geht an Prof. Henning Tidow für die Übernahme des Zweitgutachtens und Prof Holl, sowie Prof Torda für die Übernahme der Dissertationsgutachten.

Ein besonderes Dankeschön geht an Dr. Markus Perbandt für seine konstruktive Betreuung, dass er mir über die gesamte Zeit stets mit Rat und Tat zur Seite stand, mich mit seinem hochqualitativen Feedback und den wissenschaftlichen Diskussionen gefördert und gefordert hat.

Ein Dankeschön auch an Prof. Dr. Eva Liebau von der Universität Münster, die zusammen mit Dr. Perbandt eine der Ideengebenden dieses Forschungsthemas war und mich ebenfalls jederzeit freundlich und kompetent unterstützt hat.

Weiterhin möchte ich mich beim Exzellenzcluster „Center for Ultra Fast Imaging“ bedanken, womit diese Arbeit und ich zu einem großen Teil finanziert wurden. Außerdem ermöglichte das Exzellenzcluster mir die Teilnahme an diversen spannenden wissenschaftlichen Tagungen und Workshops, die Möglichkeit mich zu vernetzen und mit anderen Wissenschaftlern (Jung und Alt) auszutauschen.

Ich bedanke mich bei Prof. Dr. Dmitri Svergun und Dr. Alexey Kikhney für die ausdauernde und überaus hilfreiche Unterstützung bei der Auswertung der herausfordernden SAXS-Daten. Ohne euch wäre diese Arbeit nie so weit gekommen.

Auch ohne die Hilfe von Dr. Robin Schubert wäre diese Arbeit niemals möglich gewesen. Vielen Dank noch einmal für deine Hilfe bei der Vorbereitung, Generierung und Auswertung der AUC Daten. Danke dafür, dass ich jederzeit die Geräte und Ausstattung des XFEL Labors nutzen durfte

Eine weitere große Hilfe kam von Mitgliedern des UKE Eppendorf. Damit meine ich Prof. Hartmut Schlüter und seine Mitarbeiter Dr. Stefan Veltel und Dr. Laura Heikaus. Zum einen, da sie mich in die Methodik und in die Analyse von Nukleotiden mittels RP-HPLC eingeführt haben und zum anderen mich mit ihren Geräten und Fachwissen unterstützt haben.

Bei den Mitarbeitern der Protein Characterization Facility des CSSB bedanke ich mich für die Unterstützung mit Wissen und Geräten bei den Microscale Thermophoresis Experimenten.

Vielen Dank an sämtliche aktuelle und ehemalige Mitglieder der Arbeitsgruppe Betzel für jedwede Hilfe, die ihr mir habt zu Teil werden lassen und ganz besonders für die Quality Time im und außerhalb des Labors. Ganz besonders möchte ich Dr. Christina Schmidt danken für all den Spaß, die witzigen und leckeren Mittagspausen, die ausgiebigen Grillabende und Weihnachtsfeiern, dem einen oder anderen Feierabendgetränk und deine tolle Unterstützung während der ganzen Zeit.

Nicht zu vergessen ist mein Dank an Dr. Lida Rostock. Bereits während unseres Studiums in Greifswald warst du für mich nicht wegzudenken. Während der Promotionszeit konnte ich immer auf dich zählen. Vielen Dank für die tolle Zeit, die langen Telefongespräche und deine immer offene und fröhliche Art, die mir über manch schwierige Forschungsphase hinweggeholfen hat.

Schließlich möchte ich mich bei meiner Familie und meinen Freunden bedanken. Unendliche Male Danke an euch alle, dass ihr mich stets unterstützt, mir zugehört und mit Rat und Tat zur Seite steht. Danke Mama, Papa, Stefan, Oma Hannelore und Opa Dieter. Danke Flori, Kim, Julian, Tina, Wolle, Claudi Kai, Tine und Anica sowie alle, die ich vielleicht vergessen habe.

Das Beste kommt zum Schluss. Daher geht mein letzter Dank an meine Partnerin Steffi. Wofür ich alles dankbar bin, kann ich hier gar nicht aufzählen. Ich versuche es trotzdem. Vielen Dank für deine Unterstützung, deine tolle Art mich immer wieder zu motivieren, wenn ich es brauche, mich aufzubauen und mich zum Lachen zu bringen, wenn es nötig ist und zu jeder Zeit einfach eine tolle Partnerin zu sein. Vielen Dank für deine Hilfe,

deine Art, die vergangenen und kommenden Abenteuer und einfach nur danke, dass du der Mensch bist, der mich perfekt ergänzt.

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.