Technische Universität München

Max Planck-Institut für Biochemie

Abteilung für Molekulare Strukturbiologie


# "Cryo-EM single particle analysis of ASCE-ATPases"


Jan Michael Schuller


Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.


Vorsitzender:             Univ.-Prof. Dr. Michael Sattler


Prüfer der Dissertation:    1. Hon.-Prof. Dr. Wolfgang Baumeister

                          2. Univ.-Prof. Dr. Sevil Weinkauf


Die Dissertation wurde am 09.05.2016 bei der Technischen Universität München

eingereicht und durch die Fakultät für Chemie am 21.06.2016 angenommen.

# Zusammenfassung:

ATP-abhängige molekulare Motoren hydrolysieren ATP, um Nukleinsäuren oder Proteine zu bewegen. Die am häufigsten vertretene und vielseitigste Familie dieser Enyzme ist die der „phosphate-binding loop" (P-loop) ATPasen. Ihre Kennzeichen sind eine dreischichtige Rossman-Faltung und konservierte Walker A und B Motive. Die größte Unterfamilie der P-loop ATPasen stellt die „additional strand catalytic glutamate" (ASCE) Gruppe, die sich durch einen weiteren Strang im zentralen β-Faltblatt und einen zusätzlich hoch konservierten Glutamatrest auszeichnet. In dieser Arbeit untersuche ich die Struktur homohexamerer ASCE-ATPasen im archaealen DNA-End-prozessierungs-Komplex HerA-NurA und der molekularen Segregase p97 mittels modernste Instrumentation der Kryo-Elektronenmikroskopie (EM) und Einzelpartikelanalyse.

In Archaeen ist HerA-NurA als gekoppelter Helikase-Nuklease-Komplex für die Generierung einzelsträngiger DNA-Enden am 5´-3´-Ende eines DNA-Doppelstrangbruchs verantwortlich. Durch diese Aktivität wird die Reparatur durch homologe Rekombination ermöglicht. Mittels Einzelpartikelanalyse war es möglich, die Struktur des Enzymkomplexes bei einer Auflösung von 7.3 Å zu rekonstruieren. In das Volumen der erhaltenen Elektronendichte konnten wir sowohl die Kristallstruktur von NurA als auch ein auf verwandten Strukturen beruhendes Model großer Teile der HerA-Helikase eindeutig platzieren. NurA bindet mit seiner hydrophoben Seite auf die Oberfläche von HerA. Durch den gesamten Komplex zieht sich ein zentraler Kanal, der im Inneren von HerA bis zu drei Windungen doppelsträngige (ds)DNA aufnehmen kann. HerA arbeitet als Motor, der ein dsDNA-Substrat durch den zentralen Kanal in Richtung NurA schiebt. In NurA verengt sich der Kanal, was dazu führt, dass die dsDNA aufgeschmolzen wird. Ein derartiger Helikase-Nuklease-Mechanismus wurde bisher noch nicht beschrieben und unterscheidet sich stark von den Mechanismen vergleichbarer Enzyme in den unterschiedlichen Domänen des Lebens.

Die ATPase p97 vom Typ AAA+ spielt eine wichtige Rolle für den proteasomalen Abbau von Proteinen. Zusammen mit dem dimeren Kofaktor Npl4-Ufd1 segregiert p97 ubiquitinierte Substrate von der zellulären Umgebung, um sie dem 26S-Proteasom zuzuführen. Bisherige Versuche, die dafür notwendigen Zustände von p97 mittels Röntgenkristallografie zu

untersuchen, lieferten nur geringfügig unterschiedliche Strukturen, da die feste Kristallpackung die Bewegungen der einzelnen Domänen künstlich einschränkt. Mit Kryo-EM und Einzelpartikelanalyse dagegen kann man die verschiedenen Zustände von p97 in Lösung untersuchen. Unter Verwendung verschiedener Nukleotid-Analoga gelang es mir so, sechsfach symmetrische Strukturen zu bestimmen, die bis in den Sub-Nanometerbereich aufgelöst sind. In diesen Rekonstruktionen ist die Pore der D1-ATPase-Domäne von p97 immer geschlossen, während sich die Pore der D2-ATPase-Domäne nach ATP-γS-Bindung verengt. Des weiteren zeigen die Strukturen ein Schwingen der N-terminalen Domäne, die damit als Hebel dienen könnte, um ubiquitylierte Proteine aus der zellulären Umgebung zu lösen. Die D2-Pore dient vermutlich anderen Zwecken, wie dem Entfalten bereits gebundener Proteine.

Die vorliegende Arbeit zeigt Beispielhaft die vielfältigen Einsatzmöglichkeiten der Einzelpartikelanalyse mittels Kryo-EM. Im Falle von HerA-NurA ließ sich demonstrieren, dass atomare Modelle aus Kristallstrukturen in einen größeren molekularen Kontext eingefügt werden können. Zusätzlich zeigte die Studie von p97 das große Potenzial der eingesetzten Methode zur Untersuchung beweglicher Proteinkkomplexe in Lösung. Deshalb kann man davon ausgehen, dass für Struktur-Funktionsstudien komplexerer molekularer Maschinen, ähnlich denen in der vorliegenden Arbeit, die Einzelpartikelanalyse in Zukunft zunehmend an Bedeutung gewinnen wird.

# Summary:

Motor ATPases use the energy of ATP hydrolysis to move nucleic acids or polypeptides. The phosphate-binding loop (P-loop) ATPase superfamily is by far the most abundant and diverse nucleotide-binding fold. All P-loop ATPases share a three-layered Rossman-fold and conserved Walker A and B motifs. The largest subfamily herein is the additional strand catalytic glutamate clade (ASCE), characterized by an additional strand in the central sheet and by a conserved catalytic glutamate, positioned C-terminally of the Walker B motif. In this thesis I used state-of-the-art cryo-EM single particle analysis (SPA) to investigate the structure and mechanism of two homohexameric ASCE ATPases: the archaeal DNA end-resection complex HerA-NurA and the molecular segregase p97.

In archaea, the helicase-nuclease molecular machine HerA-NurA is responsible for processing the 5'-3' end of double-stranded (ds) DNA for repair by homologous recombination. Using cryo-EM single particle analysis we obtained a 7.3Å resolution structure of the HerA-NurA complex. The NurA crystal structure as well as a quasi-atomic model of HerA can unambiguously be fitted into the obtained density map. HerA contains a central channel capable of accommodating up to three turns of dsDNA. This cavity is aligned with the hydrophobic face of the NurA dimer, forming a continuous channel traversing the entire complex. Based on our structure, we proposed a model where the HerA motor threads a dsDNA substrate through its central channel towards the NurA nuclease for processing. The tight annulus of the NurA would act as a molecular ploughshare, splitting the dsDNA. This mechanism differs substantially from the end-resection mechanisms observed in the other kingdoms of life, revealing a novel mode of translocation-coupled DNA-end processing in archaea.

The AAA+ ATPase p97 is a major facilitator of proteasomal degradation. In conjunction with the heterodimeric Npl4-Ufd co-factor, p97 extracts ubiquitylated proteins from their cellular environment. Previous attempts to study the conformational landscape of p97 by X-ray crystallography were hampered by crystal packing forces, with the N-terminal domains most strongly affected. Thus, we determined sub-nanometer cryo-EM structures of p97 in the presence of nucleotide analogs, allowing the study of nucleotide-dependent

conformational changes. The structures revealed that p97 has six-fold symmetry in solution. In all structures, the pore formed by the D1 domains remains closed irrespective of the bound nucleotide, whereas the pore of the D2 domains shows an iris-like closing of its center upon binding ATP-γS. In addition, we observed nucleotide-dependent large-scale swinging motions of the N-terminal domains. This movement could provide the force necessary for the segregation activity of p97. The D2 cavity may contribute to additional enzymatic functions such as unfolding of bound substrate proteins.

My work exemplifies the power of modern SPA. The HerA-NurA study shows that it is straightforward to use SPA to place individual crystal structures into their molecular context. Furthermore, our analysis of the p97 conformational cycle demonstrates that cryo-EM methods can provide important insights into the functional mechanisms of dynamic protein complexes. Thus, we expect that SPA will enhance our understanding of a wide variety of protein assemblies.

# Table of Contents

# Preface

The work of my thesis resulted in two peer reviewed first-author articles. Since these two manuscripts represent the major part of the work undertaken in course of my doctoral project, this thesis is written in cumulative style. The first chapter features an introduction to the biological and methodological background and the current state of research. The second chapter includes the original manuscripts, which represents the classical "Results" and "Material and Methods" chapters. The final section features a thorough discussion that integrates the main aspects from both manuscripts.

First author articles:

**Schuller JM**, Beck F, Lossl P, Heck AJ, Forster F (2016) Nucleotide-dependent conformational changes of the AAA+ ATPase p97 revisited. *FEBS Lett* 590**:** 595-604

Byrne RT*, **Schuller JM***, Unverdorben P, Forster F, Hopfner KP (2014) Molecular architecture of the HerA-NurA DNA double-strand break resection complex. *FEBS Lett* 588**:** 4637-4644

* These authors contributed equally

Other contributions:

Butryn A, **Schuller JM**, Stoehr G, Runge-Wollmann P, Forster F, Auble DT, Hopfner KP (2015) Structural basis for recognition and remodeling of the TBP:DNA:NC2 complex by Mot1. *Elife* 4

Chen X, Chen Y, **Schuller JM**, Navab N, Forster F (2014) Automatic particle picking and multi-class classification in cryo-electron tomograms. IEEE 978: 838-841

Chen Y, Pfeffer S, Hrabe T, **Schuller JM,** Forster F (2013) Fast and accurate reference-free alignment of subtomograms. *J Struct Biol* 182**:** 235-245

Review:

Forster F**, Schuller JM**, Unverdorben P, Aufderheide A (2014) Emerging mechanistic insights into AAA complexes regulating proteasomal degradation. *Biomolecules* 4**:** 774-794

# Chapter I: Introduction

## 1.1 The structure and function of AAA and RecA-like machines

### 1.1.1 AAA and RecA machines are evolutionary related

A variety of cellular functions are performed by ATPases, which convert chemical energy into mechanical work. One general requirement within the cell is the physical motion of biopolymers, such as nucleic acids and polypeptides, relative to each other. ATPases undergo conformational changes as they hydrolyze nucleotides and can act as polymer-dependent motors to generate tension. Thus, they mediate the unfolding or folding of polypeptides, the dissociation of protein-protein or protein-nucleotide interactions, as well as the unidirectional movement along tracks (Vale, 2000). These processes require powerful molecular motors, which are built from one of four basic families: the P-loop NTPase (phosphate-binding loop), the GHL (gyrase, Hsp90, MutL), actin and chaperonin family (Thomsen and Berger, 2008).

The P-loop NTPases superfamily is by far the most abundant and diverse nucleotide-binding fold. Its members share a three-layered Rossman-fold that contains regularly recurring α-β units with the β-strands forming a central β-sheet (Leipe et al., 2003). The reaction catalyzed by this enzyme family is the hydrolysis of the beta-gamma phosphate bond of a bound nucleoside triphosphate (NTP, either GTP or ATP). In order to respond to the hydrolysis event the difference between the bound nucleotide and its hydrolysis product must be recognized. Either the γ-phosphate is directly sensed by a positively charged residue or indirectly, by sensing the $Mg^{2+}$ ion that binds to the γ-phosphate. The energy of the hydrolysis can thus be utilized to induce conformational changes within the molecule. Two conserved sequence motifs are present in all P-loop NTPases, the Walker A and B motif (Walker et al., 1982). The Walker A or phosphate-binding loop (P-loop) is located at the tip of the first beta strand (Figure 1) and has a sequence of G-x(4)-GK-T/S, where x is any residue. The Walker B motif is more variable: the consensus sequence is R/K-x(4)-G-x(4)-LhhhhD(E), where h denotes a hydrophobic residue. The residues in these motifs are responsible for nucleotide and phosphate binding. They also help to coordinate

the $Mg^{2+}$ ion bound to the γ-phosphate. Typically, the Walker B motif is located between helices α1 and α3 of the ATPase.
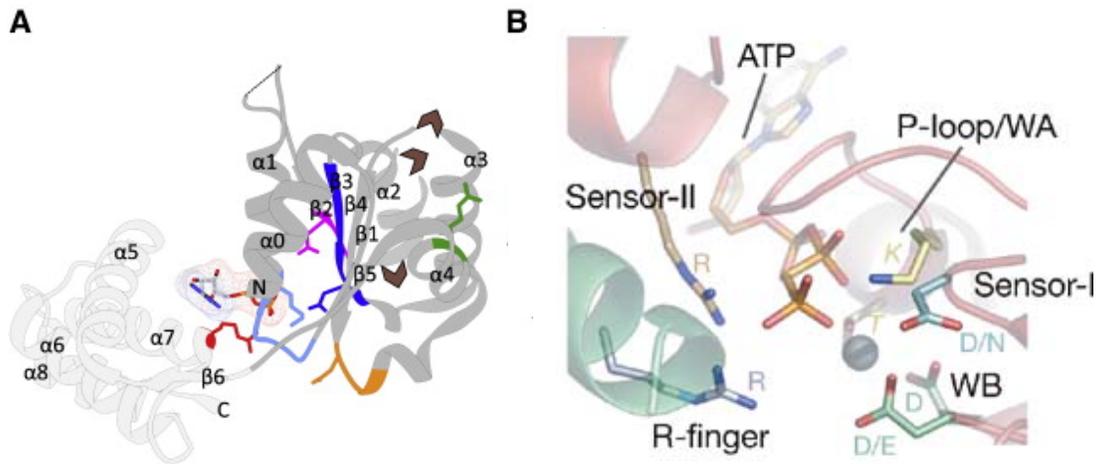


Figure 1: Structure of the ASCE ATPase. (A) Overview of a typical ASCE protomer. The ASCE domain is colored in dark grey and the C-terminal domain in light grey  (B) Detail of the P-loop active site of ATP-DnaA showing the position of nucleotide-interacting motifs and ATP. Figures modified from (Erzberger and Berger, 2006; Wendler et al., 2012).

The P-loop ATPases are separated in two large groups. The kinase-GTPase clade, to which kinases and GTPases belong, as well as the ASCE group group (additional strand, catalytic E), which is characterized by an additional strand in the central sheet, located between the two Walker motifs (Leipe et al., 2003). For ATP hydrolysis the ASCE ATPases utilize a conserved catalytic glutamate, just C-terminal of the Walker B motif, that primes a water molecule for a nucleophilic attack on the γ-phosphate group of ATP; thus the name ASCE clade.

Members of the ASCE group can be further classified based on additional signature motifs (Neuwald et al., 1999)(Figure 2). The two largest subfamilies thereof are the AAA and RecA-like ATPases proteins (Iyer et al., 2004a; Lyubimov et al., 2011).  The folds of RecA-like and AAA ATPases are related, but differ in their topology, conserved ATP-binding residues and quaternary assembly in oligomers. The RecA-fold always forms homo-oligomers, with dimers (SF1/2 family) and hexamers (classical RecA) being most common. The functional diversity of the RecA-fold is mediated by the attachment to or the insertion of functional domains within the RecA-like fold, such as the C-terminal helical bundles found in F1-ATPase or domains Ib and IIb found in SF1/SF2 helicases (Thomsen and Berger, 2008).
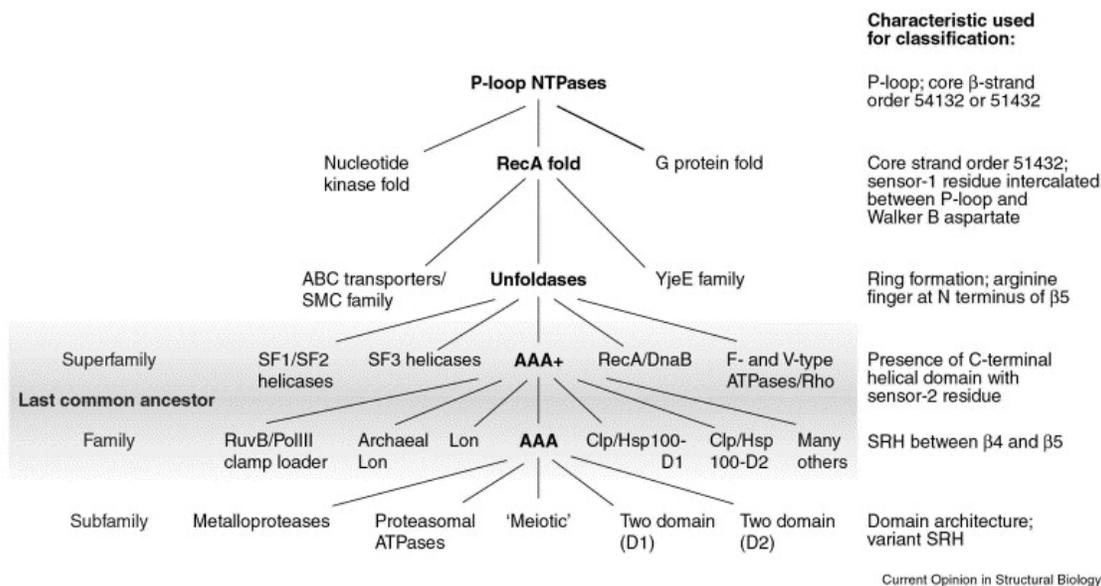
Figure 2: Phylogenetic classification of AAA domains within the P-loop NTPases. The criteria used for classification at each node are listed on the right. Figure adapted from (Lupas and Martin, 2002).

All RecA-like proteins contain an arginine finger (Hopfner and Tainer, 2003). This residue points towards the γ-phosphate of the bound ATP in the adjacent subunit. It is responsible for stabilization of the transition state during nucleotide hydrolysis. In addition to enhancing hydrolysis, this arginine-finger is critical for coupling nucleotide hydrolysis to large-scale conformational changes in adjacent RecA-like domains.

### 1.1.2 Structure and mechanism of RecA-type ATPases

The mechanism of dimeric SF1/2 ATPases is best understood (Lee and Yang, 2006). After The hydrolysis, the intra-domain conformational changes are comparably small, but they are- amplified by inter-domain interface rearrangements. In the apo conformation both RecA-folds splay apart, removing the arginine-finger from the adjacent fold. Upon ATP binding, the RecA-dimer rotates giving rise to a closed conformation and the arginine finger is inserted into the adjacent ATP-binding pocket, thereby triggering ATP hydrolysis (Figure 3). This ATP-dependent rearrangement of two RecA folds is utilized to mediate the association with nucleic acid substrate or move along nucleic acid tracks (Velankar et al., 1999).

The mechanism of ring-forming RecA-like proteins is less clear. However, it is commonly accepted that a series of inter-subunit conformational changes moves the respective

substrate through a pore in the center of the ring. The central pore is lined by substrate-engaging loops. The residues of these loops mediate the interaction with the substrate in a nucleotide dependent fashion. In theory, three distinct mechanisms are possible by which ATP turnover is coupled to substrate movement: a sequential, concerted or stochastic progression of nucleotide hydrolysis around the ring have been proposed (Lyubimov et al., 2011). Structures of fully symmetrized particles in which all subunits occupy similar conformations can serve as evidence for the concerted model. It predicts that all ATPase sites hydrolyze the bound nucleotide simultaneously (Gai et al., 2004; Wang et al., 2001). The rotary model in contrast proposes a sequential nucleotide hydrolysis from one subunit to the next, as observed in case of the F1 ATPase.
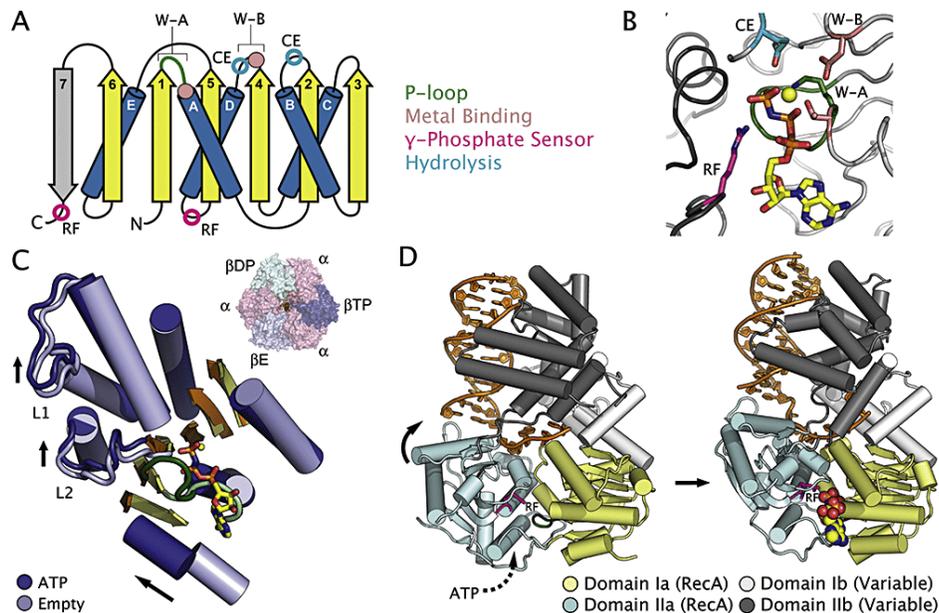


Figure 3: Nucleotide binding and conformational changes in RecA-like ATPases. (A) Topology diagram of the RecA fold. ( B) The active site of the *Saccharomyces cerevisiae* F1-ATPase β-subunit (PDB entry 2HLD) (C) Structural superposition of the conserved RecA fold of the ATP-bound (βTP) subunit (dark blue) and the empty (βE) subunit (light blue) of S. cerevisiae F1-ATPase (PDB entry 2HLD) (D) ATP-dependent conformational changes between two RecA-like folds shown by the structure of E. coli UvrD bound to DNA (PDB entry 2IS1 and 2IS4). Figure and text adapted from (Thomsen and Berger, 2008).

Support for the sequential model comes from the structure of the Rho transcription termination factor bound to a substrate single stranded (ss) DNA (Bogden et al., 1999). This structure reveals a radially asymmetric arrangement of ATPase states, in a sequence around the ring that indicates a full hydrolysis cycle.

## 1.1.3 Structure and mechanism of AAA/AAA+ proteins

The fold of the AAA protein is highly conserved. The defining feature of the AAA family is another highly conserved amino acid sequence within their ATPase domain that has been termed the second region of homology (SRH). The about 250 amino acid residue sequence that encompasses both Walker motifs and the SRH is termed the AAA module. Proteins that contain one or two copies of this module are part of the AAA protein family. The crystal structures of the δ′subunit of *Escherichia coli* DNA polymerase III (Pol III) and of the NSF-D2 hexamer reveal that a whole class of hexameric unfoldases has a high structural similarity with the AAA proteins, although having limited sequence conservation in the SRH (Guenther et al., 1997; Lenzen et al., 1998). They are termed AAA+ proteins and share the positions of the conserved residues of the P-loop catalytic core. The structural core of the AAA+ ATPases consists of two subdomains, the large N-terminal α/β ASCE-P-loop-fold and a smaller C-terminal three-helix bundle α-helical subdomain, that forms a lid on top of the large subunit (Figure 4).
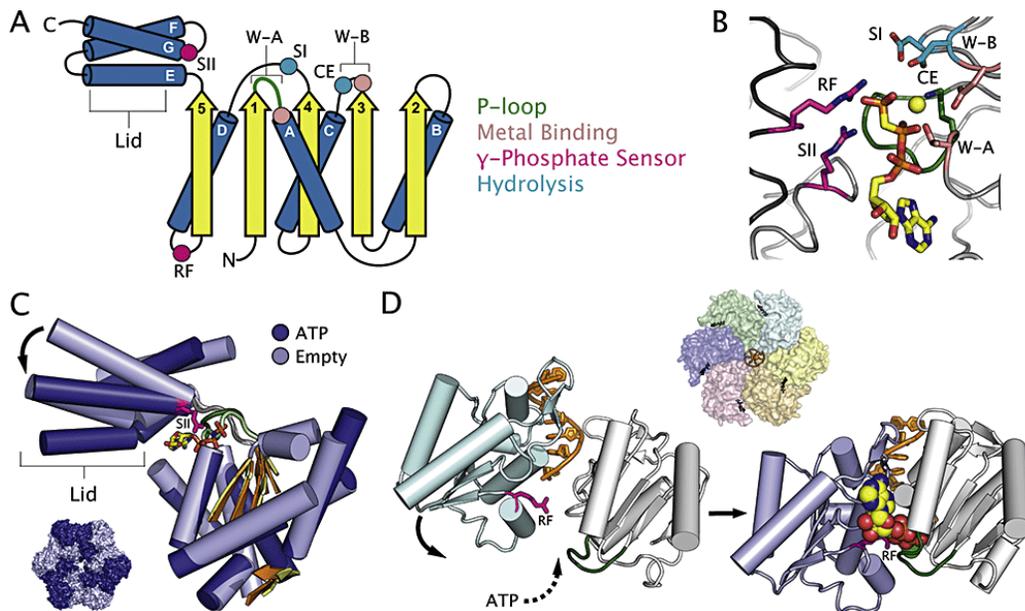


Figure 4: Nucleotide binding and conformational changes in AAA+ ATPases. (A) Topology diagram of the typical two-domain AAA+ fold. (B) The active site of *Aquifex aeolicus* DnaA (PDB entry 2HCB) ( C) Structural superposition of the conserved AAA+ fold of ATP-bound (dark blue) and empty (light blue) subunits of *Escherichia coli* HslU (PDB entry 1DO2) (D) ATP-dependent conformational changes between two AAA+ folds exemplified by the structure of the human papillomavirus E1 helicase bound to DNA (PDB entry 2GXA). Figure and text adapted from (Thomsen and Berger, 2008).

In the AAA+ fold the catalytic active glutamate is always located in or directly adjacent to the Walker-B motif. Two additional evolutionary highly conserved catalytic residues also characterize the AAA+ family, the sensor-I (SI) and sensor-II (SII) motifs. The sensor-I motif is a polar residue at the C-terminus of the β-4 strand, which orients the water molecule, attacking the γ- phosphate (Figure 1). The sensor 2 motif is a positively charged residue positioned at the N-terminus of the seventh alpha helix, contacting the bound nucleotide (Botos et al., 2004). This residue is responsible for coupling the nucleotide sensing to the intra-subunit conformational changes. The arginine finger and the two sensor motifs are not conserved in all AAA+ members, but are very common. If present, the arginine finger is always located at the C-terminus of the fourth α-helix. The crystal structure of the E1 helicase from papillomavirus bound to ssDNA has revealed the mechanism of AAA+ nucleotide translocases (Enemark and Joshua-Tor, 2006). Despite being a homohexamer this structure adopts an asymmetric conformation. The position of each protomer varies sequentially around the hexamer and is coupled directly to the position of their central pore-loops. This strongly suggests a sequential hydrolysis mechanism.

Further evidence for a sequential mechanism, comes from biochemical mutant doping experiments. In these experiments, the WT protein is mixed with increasing amounts of the catalytically inactive Walker A motif mutation. It was shown that the formation of heterologous rings renders most ring forming translocasses inactive, suggesting a coordinated series of ATP hydrolysis around the ring. Most RecA-type helicases are rendered inactive after introducing a single inactive protomer (Massey et al., 2006), whereas the AAA+ rings are only inactivated after the insertion of several defective protomers (Moreau et al., 2007). These findings suggest a rotary mechanism, which retains sufficient flexibility to allow alternative subunits to hydrolyze ATP if one or more corresponding protomers are inactivated. Thus, it seems that the RecA-like helicases act in a more 'strictly sequential' fashion than AAA+ helicases.

## 1.1.4 Comparison of RecA-type and AAA+ ring translocases

The RecA-type and AAA+ nucleotide translocases differ in their quaternary structures in the respective hexameric oligomers. They are oriented such that the central channel of the hexamer runs roughly perpendicular for the two types of hexameric rings (Wang, 2004). In

case of single-stranded nucleic acid translocases such different quaternary structure is known to determine the directionality of substrate threading, which can be either 5'to 3' (RecA-like proteins) or 3' to 5' (AAA+ machines) oriented.

In contrast to RecA-like proteins, which exclusively interact with nucleic acids, AAA+ proteins also act on polypeptides. The respective AAA+ proteins are involved in the unfolding and disassembly of proteins and protein complexes(Nyquist and Martin, 2014). The best-studied system for the mechanism for such protein unfoldases is the protease system of the AAA+ ring ClpX together with the self-compartmentalized ClpP protease (Nyquist and Martin, 2014) as well as the 26S proteasome (Forster et al., 2014).

## 1.2 DSB repair and DNA end-resection in the three kingdoms of life

### 1.2.1 DSB damage and repair in the three kingdoms of life

Most proteins in the archaeal kingdom of life display a clear homology with their eukaryotic counterparts. This is especially true for proteins involved in fundamental genomic processes, such as DNA replication, transcription and recombination. These molecular machines have most likely a common evolutionary ancestor (Kelman and White, 2005). The integrity of the genome is a major feature of all living organisms. Moreover, compounds such as reactive oxygen species, generated during normal cellular metabolism, and environment hazards, such as ionizing radiation or chemical mutagens, threaten the genomic integrity of cells (Hoeijmakers, 2001).

The most cytotoxic form of DNA damage is the double-strand break (DSB), when both strands of the DNA duplex are simultaneously broken. Free DNA ends may lead to illegitimate recombination and could cause chromosomal rearrangements (Hoeijmakers, 2001). The two predominant DSB-repair mechanisms are non-homologous end-joining (NHEJ) and homologous recombination (HR) (Harper and Elledge, 2007). NHEJ is based on the rejoining and ligation of two broken DNA ends and is active throughout the entire cell. The HR pathway is dependent on the existence of an intact sister chromatid as template (Koch et al., 2004), and is therefore active almost exclusively during cell division. In archaea, HR exclusively repairs the DSB, whereas no archaeal homologues of the NHEJ machinery have been identified.

In the first stage of HR, the DSBs are detected by the Mre11:Rad50 (MR) complex, which senses the break and initiates the repair process (Williams et al., 2010). It is responsible for nucleolytic processing of DNA ends. The next step during HR is the extensive 5′–3′ end resection of the DSB, which is mediated by the HerA-NurA complex in archaea. This process generates 3′ single-strand tails. The free ssDNA is then coated by the single-strand-binding protein (SSB), which protects it from further erosion and is necessary for binding of the RadA recombinase (RecA in bacteria and Rad51 in eukaryotes) to form the nucleoprotein filament that mediates HR. The DNA end-resection step is an evolutionary conserved process, observed in all domains of life, and is dependent on specialized RecA-like helicases and nucleases (Blackwood et al., 2013).

In the next step, the homology search and strand invasion between homologous DNA sequences is induced by the RadA recombinase. The RadA-DNA nucleoprotein filament invades the homologous strand with the help of accessory proteins, such as the translocase Rad54, which induces supercoiling in dsDNA. The homologous sequence binds to the template strand by correct base-pairing, forming the so-called displacement loop (D-loop) (Sung and Klein, 2006). A DNA polymerase extends the invading single strand that the initial lesion can be repaired. Next, the D-Loop needs to be resolved (Sung and Klein, 2006). This is carried out in a way that the 3′ overhang not involved in D-loop formation gets annealed to the extended D-loop, therefore forming a structure designated as Holliday junction. Dependent on the enzyme, which resolves the DSB, a crossover or non-crossover product is generated (Heyer, 2004).

The archaeal system represents a simplified version of the complex eukaryotic HR, but carries some unique features. Especially the end-resection step shares little similarity with the bacterial or eukaryotic system.

### 1.2.2 DNA end-resection in the three kingdoms of life

In Gram-negative bacteria, the tripartite RecBCD helicase–nuclease complex is responsible for DSB resection (Figure 5A) (Wigley, 2013). It has a strong affinity towards DNA ends, which loads it on the DSB (Dillingham and Kowalczykowski, 2008). RecB and RecD are helicases that translocate with different polarity leading to a net translocation in the same direction away from the DSB (Wu et al., 2010). The entering DNA duplex is split by the RecC

subunit and the resulting strands are passed through two different channels. The magnesium-dependent nuclease domain of RecB then degrades the emerging 3′- strand and also occasionally cleaves the 5′- strand emerging from the RecD motor (Singleton et al., 2004). The activity of RecC is modulated by the highly conserved Chi sequences that are dispersed throughout the whole bacterial genome (Spies et al., 2003). The Chi site tightly binds to the RecC subunit and restrains a loop that covers an opening in the complex. The sequestration of this loop causes an opening of an alternative exit channel that allows extrusion of a loop of single-stranded DNA, to which the RecA protein can bind (Churchill and Kowalczykowski, 2000). In Gram-positive bacterial species, the AddAB helicase–nuclease complexes, replaces the RecBCD complex and employs a very similar mechanism, switching from a dsDNA digestion mode to a single-strand resection mode at Chi sequences (Chedin et al., 2006).



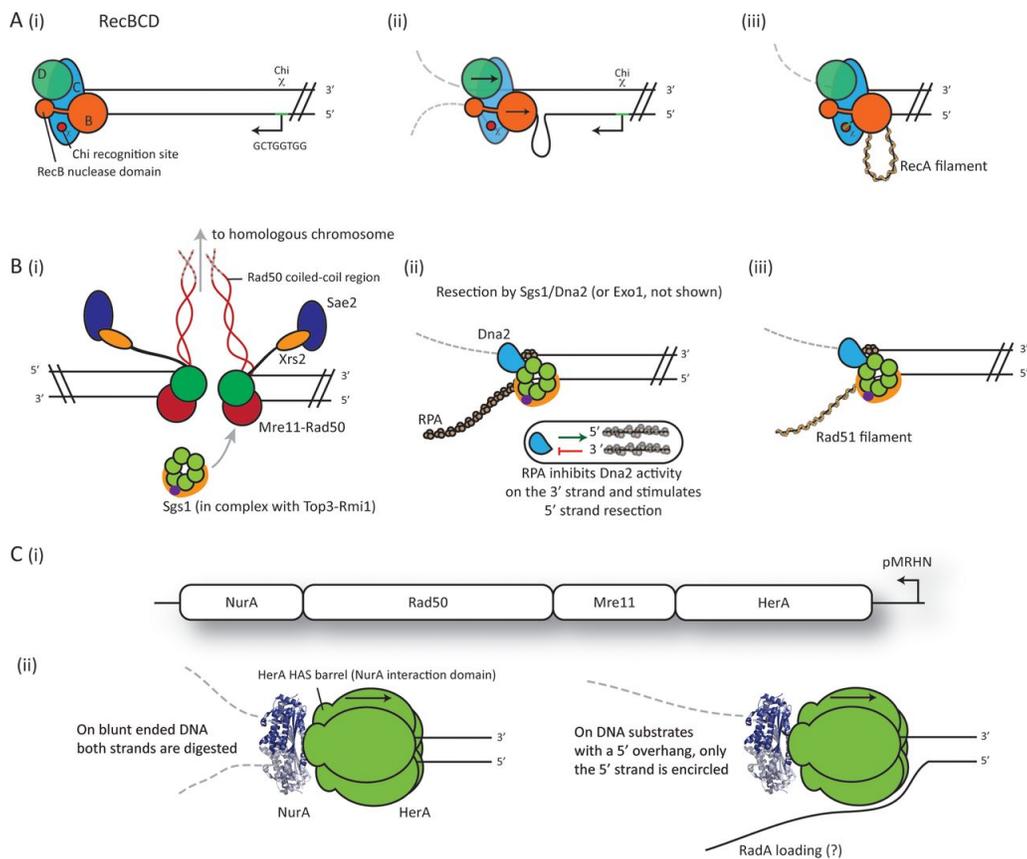Figure 5: End-resection of DSBs in the three domains of life (A) The bacterial RecBCD pathway. (B) The eukaryotic Dna2–Sgs1 complex pathway. (C) The archaeal HerA–NurA helicase–nuclease pathway. Either one or two strands of the duplex can pass through the core of the complex, resulting in either the single-stranded resection (right), or wholesale destruction of both strands (left). Figure adapted from (Blackwood et al., 2013).

In the eukaryotic domain of life, the DNA end-resection process is more complex and involves redundant combinations of nucleases and helicases (Figure 5B). Additionally, DNA end-resection is blocked by histones. To overcome this barrier chromatin remodeling complexes, such as Fun30 (Chen et al., 2012) or Ino80 (Gerhold et al., 2015), are required.

After the initial processing of the DSB is carried out by the combined action of MRN and Sae2 endonuclease, the long-range resection is mediated by the ExoI 5′–3′ exonuclease and the coupled helicase-nuclease Dna2, both activated by the association with Sgs1 helicase complex (Mimitou and Symington, 2008; Nimonkar et al., 2011; Zhu et al., 2008). The replication protein A (RPA) modulates the nuclease activity of Dna2. It stimulates the digestion of RPA-coated 5′ strands, but inhibits 3′ strand digestion (Nimonkar et al., 2011).

In contrast to the complex machinery of eukaryots and bacteria the minimal end-resection machinery of thermophilic archaea consists of four genes, encoded by a conserved DNA-repair-associated operon (Constantinesco et al., 2004)(Figure 5C): (i/ii) the archaeal Mre11–Rad50 complex, which generates short 3′ tails by limited endonucleolytic digestion of the 5′ strand (Hopkins and Paull, 2008), (iii) a helicase of the HerA/FtSK superfamily HerA, which is characterized by the presence of a central RecA-like domain containing a α-helical insert (Iyer et al., 2004b), and (iv) NurA, a 5′–3′ exonuclease (Constantinesco et al., 2002).

## 1.2.3 Initial functional and structural studies of the HerA-NurA complex

The crystal structure of NurA led to the first structural and mechanistic insights into end-resection in archaea (Blackwood et al., 2012; Chae et al., 2012). NurA forms an obligate dimer of active RNAse-H like domains (Figure 6). The NurA dimer harbors a central cavity with a diameter not large enough to pass a dsDNA helix. However, both single strands of unwound dsDNA could reach into the separate active sites located within the channel. Thus, the DNA duplex has to be unwound by the HerA helicase.  HerA and NurA form a complex with 6:2 stoichiometry (Blackwood et al., 2012). The flat surface of the NurA dimer most likely interacts with the N-terminal HerA-ATP Synthase barrel (HAS) domain of HerA (Blackwood et al., 2012). Furthermore, the nature of the DNA end influences the outcome of the HerA-NurA processing (Figure 5C).
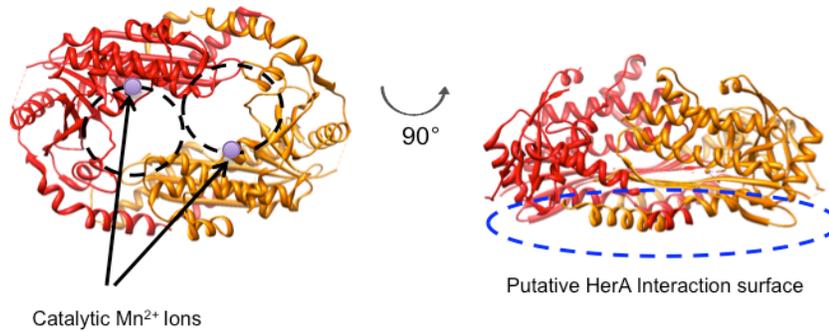
Figure 6: Crystal structure of NurA (pdb). Dimeric arrangement of NurA chains, as observed in the crystal. In the dimer one protomer is colored red the other is colored orange. Two views are shown, related by a 90° rotation. A blue circle highlights the putative HerA interaction side. The dashed circles indicate a potential ssDNA exit site.

The complex digests both strands if the type of DNA end presented to the HerA–NurA complex is a blunt end dsDNA or has only a short overhang (less than five bases). However, longer overhangs lead to degradation of only one strand of the DNA duplex, required for generation of the ssDNA for DNA repair by HR (Blackwood et al., 2012). Possibly, the nature of the DNA end presented to the HerA–NurA complex could modulate a specific biological function of the complex (Figure 5C). The dual-strand degradation mode of the HerA–NurA complex might be used as a defense mechanism against foreign DNA. Thus, the HerA–NurA complex may protect the genome from foreign DNA, as has been already observed for the RecBCD complex in bacteria (Myers and Stahl, 1994). In archaea, the limited endonucleolytic digestion of the DNA ends by Mre11 could be sufficient to generate a 5′ overhang that shifts the activity of the complex towards single-strand resection (Quaiser et al., 2008). To date, it is not understood how the enzymatic activities are coordinated within the complex in order to unwind the dsDNA and trigger the digestion of the 5′ strands, but inhibit 3′ strand degradation.

## 1.3 Structure and function of the p97 segregase

### 1.3.1 The function of Cdc48/p97 in cellular protein quality control

In concert with its cofactors, the type II AAA+ protein p97/Cdc48 (mammalian/yeast) participates in a diverse range of cellular processes that involve ubiquitylation (Baek et al., 2013). p97 is highly conserved throughout evolution and essential for cell viability. Single amino acid mutations in human p97 (also called VCP) have been linked to neurodegenerative disorders, such as inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia (IBMPFD) (Kimonis et al., 2008).

The p97 complex plays a central role as a facilitator of proteasomal degradation (Figure 7) (Richly et al., 2005). In conjunction with the hetero-dimeric substrate recruiting cofactor Npl4-Ufd (NU), p97 hydrolyzes ATP to segregate ubiquitylated proteins from their environment (Braun et al., 2002). This environment may be other subunits of a complex, a tightly associated E3 ligase or an organelle membrane. Prominent examples involving p97 are the disassembly of the replisome after chromosome replication (Maric et al., 2014), the Ltn1 quality control system in the Non-stop-decay pathway (Defenouillere et al., 2013), the degradation of stalled RNAPII (Verma et al., 2011) and the disassembly of centromeres after cell division (Merai et al., 2014).The best-studied system for p97/Cdc48's segregase function is the endoplasmic reticulum associated degradation (ERAD) pathway in the yeast *Saccharomyces cerevisiae* (Stein et al., 2014).
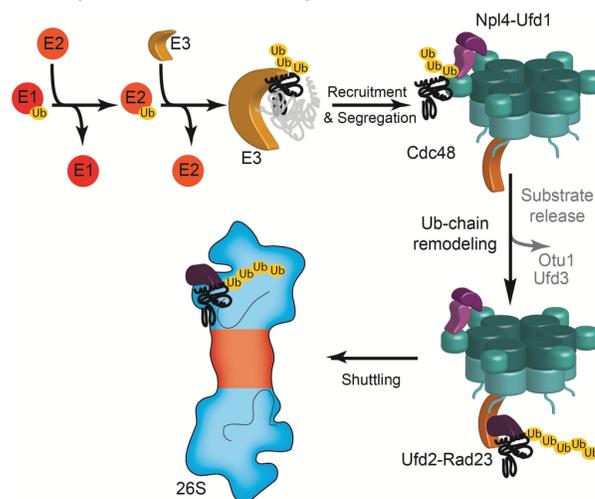


Figure 7: The p97/Cdc48 centered escort function for substrates of the UPS. Adapted from (Forster et al., 2014).

The degradation of many ERAD substrates requires the action of the AAA-ATPase Cdc48 and the NU complex (Wolf and Stolz, 2012). A likely mode of action is that Cdc48 contributes to the retrotranslocation process, by segregating the ERAD substrate from its corresponding E3 ligase and extracting it from the retro-translocon channel. In addition, the degradation of many ERAD substrates involves the polyubiquination factor Ufd2 (E4-enzyme) and the proteasome escort factor Rad23 (Medicherla et al., 2004). Suggesting a coordinated delivery of misfolded ER proteins to the 26S proteasome in order to prevent the exposure of toxic, aggregation-prone polypeptides to the cytosol. This delivery pathway may be facilitated by direct interaction between the membrane-bound E3 ligase complex and the 26S proteasome. Indeed, it has been shown in yeast *Saccharomyces cerevisiae* that large amounts of 26S proteasome reside close to the ER (Enenkel et al., 1998).

## 1.3.2 Structural studies of p97

The quaternary structure of p97 has been studied by X-ray crystallography (Davies et al., 2008). It consists of three domains (Figure 8A), a pseudo-symmetrical N-terminal (N) domain adopting a double-psi-beta barrel fold and two ring-forming AAA+ domains, D1 and D2. The AAA-rings are stacked in head-to-tail orientation and the N-domain tightly-associates in plane with the outer surface of the D1 domain (Figure 8B). The pore of the D1 ring is narrow with a diameter of 7.5 Å (Figure 8C). Additionally, p97, in contrast to its archaeal homologs, lacks critical aromatic residues in the D1 pore that are required for unfoldase activity in ATP-dependent proteases (Rothballer et al., 2007). Thus, it is unlikely that p97 threads substrates through its central channel. Full-length p97 has been crystallized in the presence of different nucleotides. In the wild-type protein, ADP is always bound to the D1 domain, while D2 was either bound to ADP, AMP-PNP (pre-activated state), or ADP-AlFx (transition state) (Davies et al., 2008). After changing the crystal-lattice environment by mutating the crystal contact sites (Figure 9A), structures of p97 were also solved in the apo-state with no nucleotide bound as well as in the pre-activated state with ATP-γS bound to the D1 and D2 domain, respectively (Hanzelmann and Schindelin, 2016b). ATP-γS binding to the D2 domain led to a rigid body-like conformational change. A superposition of the structures in the apo and ATP-γS bound state revealed a rotation of about 14° of the D2 with respect to the ND1 domains, mediated by the D1-D2 linker (Figure 9B).
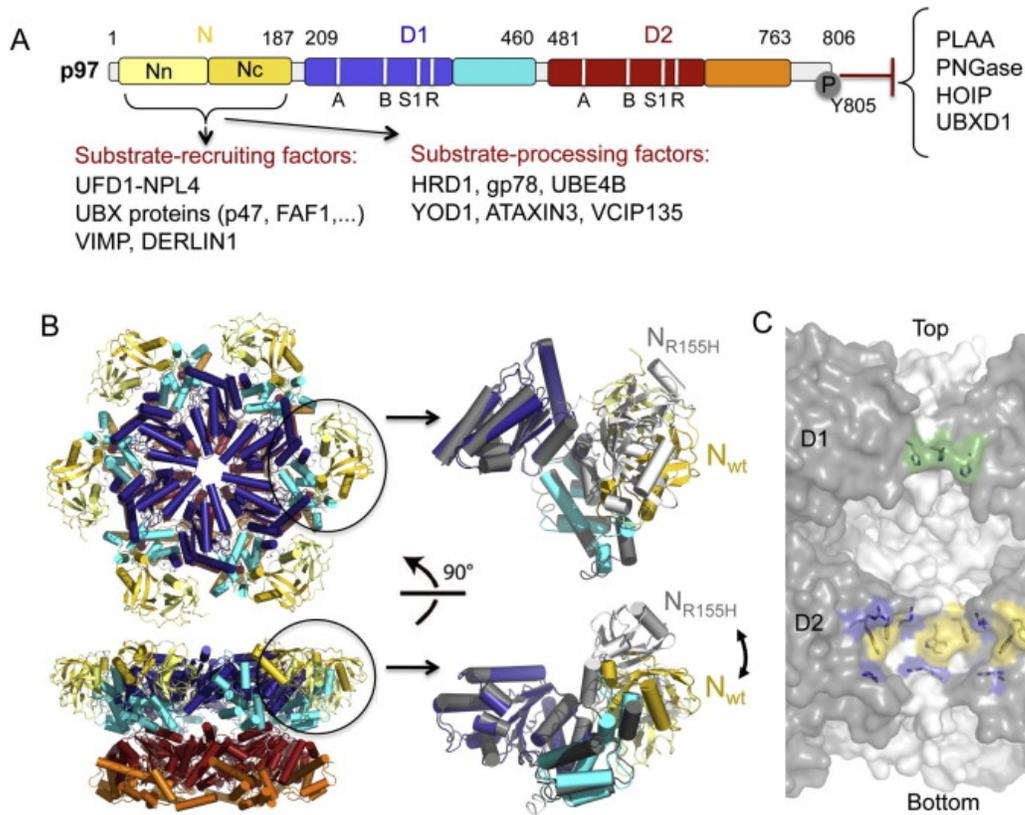
Figure 8: Structural characteristics of p97. (A) Domain architecture of p97 (B) Top and side view of the full-length p97 structure and side view of N domain conformational changes in the R155H mutant. (C) Molecular surface representation of the p97 axial channel viewed from the side. Positively charged (Arg586, Arg599) residues are highlighted in blue and hydrophobic residues in yellow (Trp551, Phe552). Figure adapted from (Buchberger et al., 2015).

However, the quaternary arrangements of these crystals structures are likely to be non-physiological and probably induced by crystal packing forces (Davies et al., 2008). In particular, the positions of the N-domains seem to be strongly influenced by the crystal packing (Figure 9A). Evidence comes from solution studies by small-angle X-ray scattering (SAXS) (Davies et al., 2005) and low-resolution cryo-EM (Rouiller et al., 2002).

 The radii of gyration obtained from SAXS exceed those obtained from the crystal structures for the different nucleotide states and the cryo-EM structures suggest a dynamic behavior of the central channel, as well as flexibility within the positioning of the N-domains. Furthermore, a crystal structure of an N-D1 fragment of p97 carrying the IBMPFD point mutations (R155H or A232E) in the presence of ATP-γS reveals an unusual N-domain conformation, in which the N-domain resides significantly above the D1-ring (Figure 8B) (Tang et al., 2010).
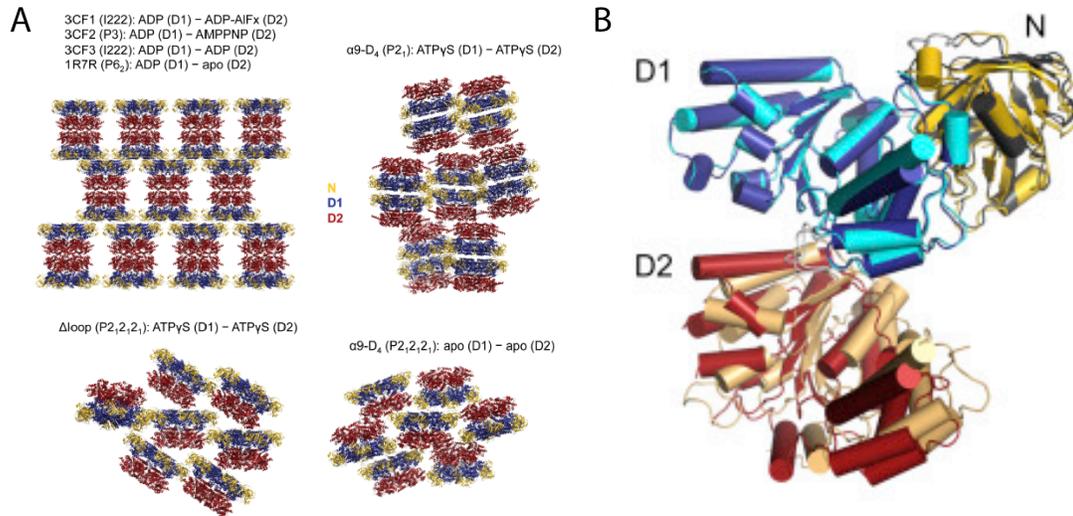
Figure 9: X-ray structures of p97 are confined by crystal packing forces. (A) Ribbon representation of the crystal packing arrangement in the different p97 constructs. (B) Ribbon representation of domain-wise superpositions of the Δloop-ATPγS structure (N in gold, D1 in blue, D2 in red) and the apo α9-D4 structure (N in gray, D1 in cyan, D2 in light orange). Figures adapted from (Hanzelmann and Schindelin, 2016b).

### 1.3.3 Cofactor interaction and specificity of the p97 segregase

The specificity of p97 is conferred by its substrate recruiting cofactors (Buchberger et al., 2015). The most important substrate-recruiting cofactors are the NU complex and p47. The NU complex is involved in poly-ubiquitin binding and mediates proteasomal degradation (Braun et al., 2002). In contrast p47 binds to mono-ubiquitylated substrates and is mostly involved in homotypic vesicle fusion (Meyer et al., 1998). Most p97 cofactors contain more than one p97 binding site, enabling them to interact with p97 in a bipartite manner (Figure 10).



Figure 10: Model for the binding of mutually exclusive p97 cofactor assemblies. Figure adapted from (Buchberger et al., 2015).

They often bind in a hierarchical and competitive manner. The NU and p47 cofactors even bind mutually exclusively to the N-domain (Hanzelmann et al., 2011). The ability to interact with a large number of cofactors in a highly regulated manner is critical for p97 function. It has been shown that co-factorbinding to the N-domain is regulated by the D1-ATPase nucleotide binding state (Chia et al., 2012). Moreover, it has been demonstrated that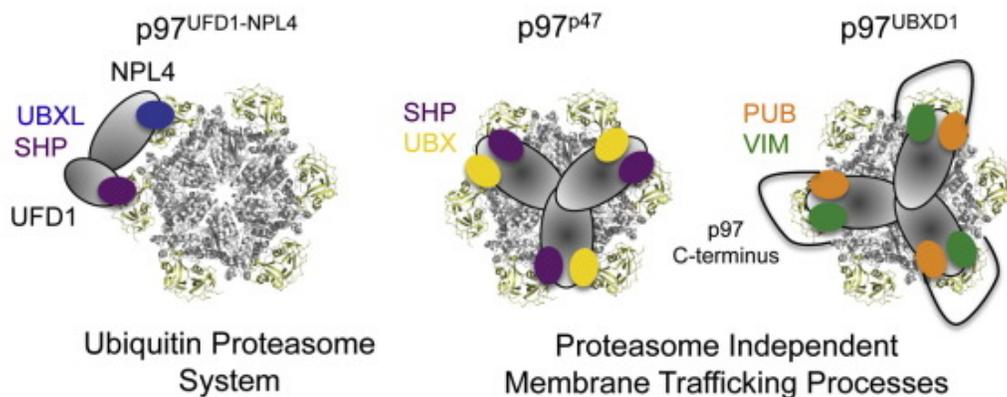 cofactors can affect the ATPase activity of p97 (Meyer et al., 1998; Zhang et al., 2015). Further studies are required to determine how nucleotide binding is reflected by conformational rearrangements of p97 and how its conformational landscape influences co-factor function.

## 1.4    Single-Particle Analysis

## 1.4.1 The principle of cryo-EM single particle analysis

In single particle analysis (SPA) the 3D reconstruction of a macromolecule is obtained from 2D projection images by averaging a large number of identical copies of the same structure at different orientations. The biological samples have to be imaged in vacuum, which imposes certain requirements on sample preparation (Dubochet, 2016). Near-to native physiological sample fixation is achieved by rapid freezing, preserving the macromolecules in a layer of amorphous, vitreous ice (Dubochet et al., 1982). Rapid cooling prevents ice crystal formation, which would disrupt the sample. Successful vitrification occurs only with a very high cooling rate, as provided by liquid ethane or ethane-propane mixtures. Once vitrified, the samples always have to be kept at liquid nitrogen temperature to avoid devitrification and ice crystal formation.
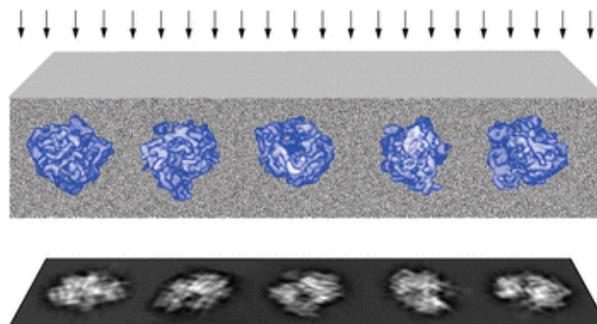


Figure 11: Generation of 2D projections of a 3D molecule. Arrows symbolize the incident parallel electron beam. 2D projections of randomly oriented copies of a 3D molecule in a thin layer of vitreous ice are generated. Figure adapted from (Frank, 2002).

A TEM image of a macromolecule corresponds to its 2D projection (Figure 11). The generation of a 3D reconstruction from 2D projections roots in the projection slice theorem: the Fourier transform of each 2D projection corresponds to a central slice of the imaged object in 3D Fourier space (De Rosier and Klug, 1968). By adding up the signal of different individual particles the object can thus be retrieved in 3D. To obtain a 3D structure of the molecule, many images displaying copies of the molecule of interest need to be analyzed to (i) obtain all required views for 3D reconstruction and (ii) to increase the signal-to-noise ratio (SNR) and hence resolution by averaging.

### 1.4.2 Low-dose data acquisition

Radiation damage is on of the key concerns to be taken into account while designing a single particle cryo-EM experiment. Due to the radiation sensitivity of biological samples the resolution of the image is fundamentally limited by the applicable electron dose, determining the signal-to-noise ratio (Baker and Rubinstein, 2010; Glaeser, 1971). Therefore, single-particle data acquisition requires strict low-dose conditions, which are most easily achieved by automation. The required automation consists of three basic steps: tracking, autofocus and exposure. Commonly, regularly spaced holes with a fixed diameter, Quantifoil grids, are used for high throughput data collection. Due, to the identical geometry of the imaging areas, the tracking step is applied to center the current Foil Hole, based on a pre-defined reference image. Auto focusing is usually carried out on the carbon support foil, 2 micrometers away from the exposure center. For auto-focusing the beam-tilt method is applied , with a beam tilt of 4mrad (Koster et al., 1992). In the exposure state, the image itself is recorded within the center of the Foil Hole. For all the work presented here the *TOM* acquisition software was employed for the above outlined automated data acquisition (Korinek et al., 2011).

### 1.4.3 Direct detector cameras and counting

In modern electron microscope systems the image is recorded on a radiation-hardened CMOS-based direct electron detection camera. These direct detectors enable direct recording of electrons on the chip, significantly increasing the SNR compared to CCDs, which detect electrons only indirectly via various statistical conversion events (McMullan et

al., 2014). The fast readout of CMOS chips allows saving the individual exposure images as a stack of frames, allowing for computational correction of image blurring and thus preserving detailed structural features (Li et al., 2013). Image blurring is caused by either instability of the sample stage or motion induced by the illuminating electron beam. Alignment of subframes allows for the majority of motion to be corrected (Brilot et al., 2012).
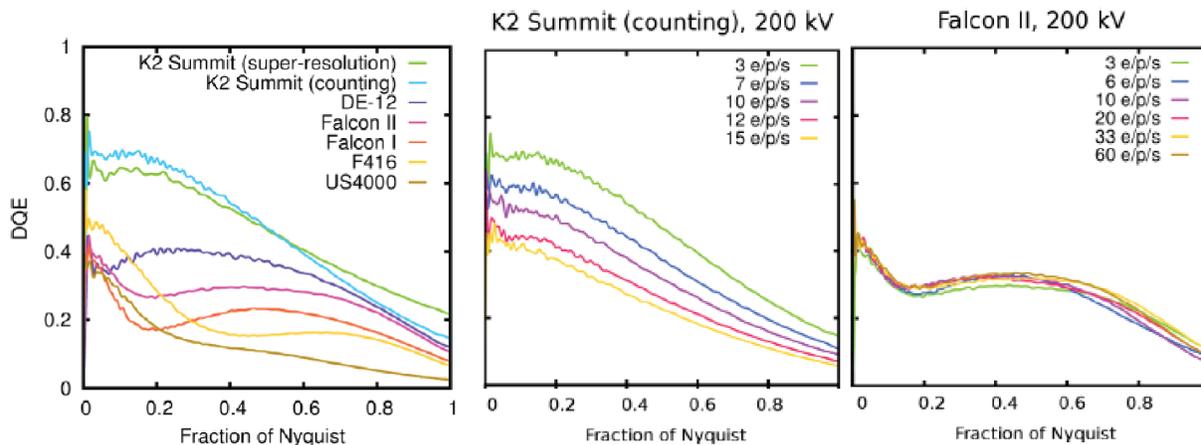


Figure 12: DQE of direct electron detectors. Left panel: DQE of detectors at 200 kV. Middel panel: Response of DEDs to increasing dose rate. The K2 Summit (counting mode) DQE decreases with dose rate. Right panel: The Falcon II DQE does not change with dose rate. Figure modified from (Ruskin et al., 2013).

The two most popular camera systems are the K2 Summit camera (Gatan) and the Falcon direct electron camera (FEI). The internal readout of the K2 Summit camera is much faster than the frame-writing speed (max. 40 Hz). This internal readout of 400 frames per second is essential for operation in counting mode: individual electron events are spatially localized when they reach the detector, which allows binary dose counting and consequently circumvents various sources of noise CCD cameras suffer from. To count efficiently, the camera must operate at a low-dose rate such that individual electron events can be separated (Ruskin et al., 2013). The benefit of counting is that it suppresses readout noise and as a consequence the detector's quantum efficiency (ratio of input and output SNR, DQE) increases across all spatial frequencies.

In counting mode, the DQE is greatly affected by the dose rate because temporally and spatially adjacent electron impacts cannot be distinguished. Thus, it is important to use the camera with a sufficiently low dose rate to benefit from its full potential (Figure 12). The K2

Summit camera requires a dose-rate lower than 5 electrons/pixel/s for optimal performance (Ruskin et al., 2013) (Figure 12B). The DQE at a target resolution (typically measured in 1/Å) is dependent on the magnification. Increasing the magnification for data acquisition is a means to increase the DQE for a certain physical frequency. Thus, high magnifications are favorable when signal transfer is limiting despite the less efficient data acquisition. This holds in particular for small complexes that exhibit little signal for particle alignment (Bartesaghi et al., 2015).

### 1.4.4 CTF correction

The contrast transfer function (CTF) is a Fourier-based description on how the image of an electron microscope is affected by the lens systems of the electron microscope as well as by the imaging parameters. The CTF is determined by adjustable parameters, such as defocus, astigmatism and the acceleration voltage, as well as fixed properties of the electron microscope, such as lens aberrations and beam coherence (David B. Williams, 2009). The CTF oscillates around zero contrast, modulating the amplitude and also reversing the phase for some frequency intervals. Moreover, the CTF has an envelope, which leads to an attenuation of signal transfer with increasing spatial frequency (Zhu et al., 1997). To determine a high-resolution cryo-EM structure the image modulation by the CTF must be corrected for.

Prior to CTF correction its parameters must be approximated accurately. In this work, the applied defocus and beam astigmatism were determined by using the *ctffind4* program (Rohou and Grigorieff, 2015). The CTF parameters are determined by computing the maximum cross correlation between the averaged power spectrum of the respective image and of a set of theoretical CTF spectra.

Currently, CTF correction is most commonly done as an integral part of the 3D reconstruction. This is also true for the *RELION* workflow used in this work (section 1.4.5). In this approach the images are subjected to a Wiener Filter prior to merging the 2D images in the 3D volume, which is a common means to avoid enhancing noise near CTF zeros (Penczek, 2010). The filter depends on the SNR of the data. In *RELION*, optimal image restoration is achieved by determination of the frequency-dependent SNR using a Bayesian approach (Scheres, 2012).

## 1.4.5 Single particle analysis processing workflow

In this work image processing is typically carried out using the *RELION* software suite (Scheres, 2012) using the processing workflow depicted in Figure 13. After selecting the individual particles, they are subsequently classified in 2D according to their shape and structural features (van Heel and Stoffler-Meilicke, 1985). The defining feature of 2D classification is that no efforts are made to explain the 2D classes by 3D model(s). The underlying assumption of this approach is that noise has a random distribution and the features of the particles have identical signal. The classification starts with random seeds, so the entire data set is divided into a number of classes, which are randomly averaged. Next, all particles are classified to the generated classes and subsequently aligned. Thus, individual particles
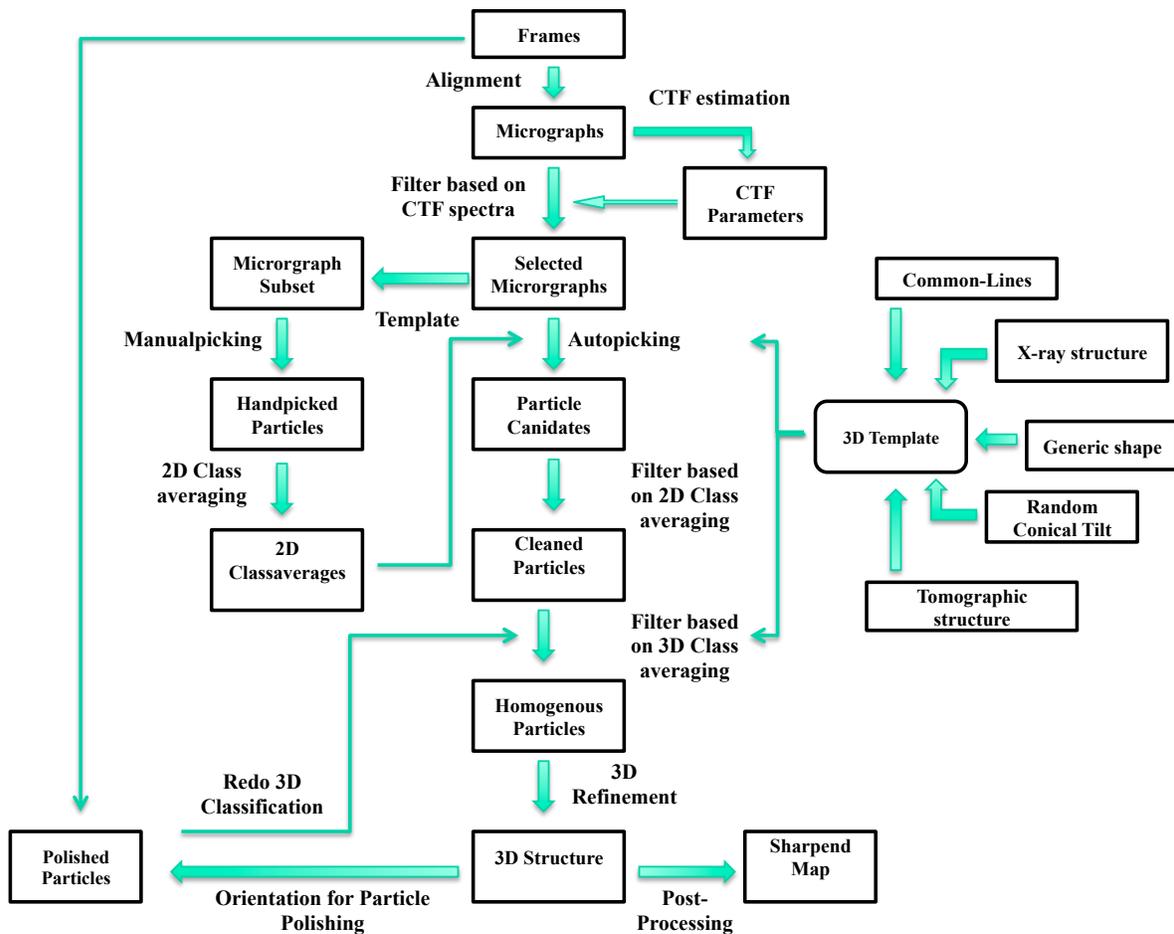


Figure 13: Single particle analysis processing workflow. It is recommended to repeat the 3D Classification and Refinement steps after the Particle Polishing routine.

21

showing the same orientation and conformation can be aligned to increase the SNR of the individual classes (Yang et al., 2012). This procedure operates in an iterative fashion: the generated output of the previous iteration is used as a model for the current iteration. The 2D classification is useful to clean the data set from non-particle candidates, e.g. ice contamination or broken particles. It is also useful to remove particles, which depict only low-resolution information due to blurring arising from drift or beam-induced charging. The accordingly cleaned particle data set is subsequently used for 3D alignment and reconstruction (see also section 1.4.1).

The relative orientations of the particles in the projections are unknown and need to be determined. In the projection-matching technique the particles are compared with computationally-generated projections of a 3D structure in order to assign 3D orientations to them (Penczek et al., 1994). This is done in an iterative fashion, by increasing the local sampling rate in order to obtain more accurate angular assignments and thus higher-resolution reconstructions.

Macromolecules often co-exist in distinct conformations in solution. In order to obtain 3D reconstructions of the different conformations, the 3D Classification step generates, similar to the 2D one, a defined number of random seeds, based on the initial model (Scheres et al., 2007). Additionally, prior to the alignment step every particle is compared to every individual reconstruction and assigned to a class. The *RELION* software package implements classification of single-particle cryo-EM data in 3D based on Maximum Likelihood (Scheres, 2010).

The individual 3D classes are subsequently refined to a higher resolution by the auto-refine procedure. In *RELION* this procedure employs the so-called gold-standard Fourier Shell Correlation (FSC) calculations for resolution estimation to avoid over-fitting (Scheres and Chen, 2012). The FSC is a function of consistency. It measures over shells of equal spatial frequencies in Fourier space the normalized cross-correlation coefficients between the Fourier transform of two 3D volumes (van Heel and Schatz, 2005). The two 3D volumes that are compared are obtained from splitting the dataset into two halves and refining them completely separately ('gold standard').

The final refined density is further filtered at the correct resolution with a map that masks away the solvent in the post-processing procedure of *RELION* (Chen et al., 2013). Additionally, the density is sharpened according to an automatically estimated B-factor (Rosenthal and Henderson, 2003). The B-factor or temperature factor describes the attenuation of structure factors by thermal mobility of the molecule atoms.

## 1.4.6 Particle polishing

Although large movements on the level of the entire micrograph are corrected by drift-correction (see also section 1.4.3), the individual particles often still show some individual motions. In the Particle Polishing algorithm of *RELION* the tracks of the individual particle motion trajectories are estimated and fitted to a linear trajectory (Scheres, 2014). Moreover, to account for accumulating radiation damage during acquisition, per-frame B-factors and linear intensity factors are estimated by comparing the reconstructed half-maps from individual frames to the full-frame half maps (Bai et al., 2013). In this way a new set of "shiny" particles are reconstructed with the spatial frequency contribution from each frame weighted according to their radiation damage (Figure 13). The shiny particles have a higher signal to noise ratio, which improves the alignment accuracy and typically yields a higher final resolution.

## 1.5    Scope of the thesis

The aim of this doctoral project was to use cryo-EM single particle analysis to study the structure and function of two ring-forming ASCE ATPases.

The study of the archaeal DNA end-resection complex HerA-NurA was carried out in cooperation with Robert Thomas Byrne from the Hopfner group (Ludwig Maximilians University Munich). The aim was to understand how the enzymatic activities are coordinated within the complex in order to unwind the dsDNA and stimulate the digestion of the 5′ strands, but inhibit 3′ strand degradation. The major question was how the HerA-NurA complex is organized and if the diameter of the central HerA channel allows for the translocation of ssDNA or dsDNA.

The work on the AAA+ protein p97 encompassed the major part of my doctoral project. Although intensely studied, the molecular mechanism underlying the segregase activity of p97 is poorly understood. It is generally accepted that AAA+ proteins undergo major conformational changes during their nucleotide cycle. The conformational landscape of p97 has been probed by various biochemical and biophysical methods, but the segregase activity of p97 remains poorly understood due to the technical limitations of the respective approaches used. To determine the conformational changes underlying the force-generation in the segregase activity of p97 in solution I obtained an ensemble of high-resolution cryo-EM single particle structures of p97 in presence of different nucleotides.

# Chapter II: Results

## 2.1      Part I: Molecular architecture of the HerA-NurA DNA double-strand break resection complex

This article was published in 2014 in FEBS Letters (Volume 588, Issue 24, Pages 4637–4644). The supplemental material is attached at the end of the article.

**Summary:**

In this study we applied cryo-EM single particle analysis to clarify the quaternary structure of *Sulfolobus solfataricus* HerA-NurA complex. In the archaeal kingdom of life, this helicase-nuclease is responsible for 5'-3' resection of double-stranded DNA (dsDNA) ends required for repair by homologous recombination. We obtained a 7.3 Å resolution density of the macromolecule after applying the appropriate $C_6$ symmetry. This EM density clearly shows how the HerA helicase, a RecA-like ATPase, is capped by the NurA nuclease dimer. We could unambiguously fit the NurA crystal structure and a comparative model comprising the HerA RecA-like and helical-insert domains into the EM map, whereas positioning of the N-terminal HAS domain is only approximate due its low sequence identity to available atomic structures. Based on the HerA-NurA structure we propose a model, in which dsDNA is translocated through a large tunnel formed in the HurA hexamer, feeding the DNA ends to NurA for processing. This mechanism differs substantially from the end-resection mechanisms observed in the other kingdom of lives, revealing a novel mode of translocation-coupled DNA-end processing in archaea.

**Contribution:**

For this article, I carried out specimen preparation for electron microscopy, acquisition of cryo-EM data and the complete workflow of single particle analysis and map interpretation. With the help of Pia Unverdorben, I generated the HerA homology model and applied flexible fitting using VMD. I wrote major parts of the manuscript and I prepared all Figures (Figure 2 and 3) and Supplement Figures that dealt with the structural analysis of the HerA-NurA complex. I was also involved in conceptualising the mechanistic model depicted in Figure 4.

**FEBS**
*Letters*

ELSEVIER

# Molecular architecture of the HerA–NurA DNA double-strand break resection complex

Robert Thomas Byrne [a,1], Jan Michael Schuller [b,1], Pia Unverdorben [b], Friedrich Förster [b,*], Karl-Peter Hopfner [a,c,*]

[a] Gene Center and Department of Biochemistry, Ludwig-Maximilians-University Munich, Feodor-Lynen-Straße 25, D-81377 Munich, Germany
[b] Department of Molecular Structural Biology, Max-Planck Institute of Biochemistry, D-82152 Martinsried, Germany
[c] Center for Integrated Protein Sciences, Munich, Germany

ABSTRACT

DNA double-strand breaks can be repaired by homologous recombination, during which the DNA ends are long-range resected by helicase–nuclease systems to generate 3′ single strand tails. In archaea, this requires the Mre11–Rad50 complex and the ATP-dependent helicase–nuclease complex HerA–NurA. We report the cryo-EM structure of *Sulfolobus solfataricus* HerA–NurA at 7.4 Å resolution and present the pseudo-atomic model of the complex. HerA forms an ASCE hexamer that tightly interacts with a NurA dimer, with each NurA protomer binding three adjacent HerA HAS domains. Entry to NurA's nuclease active sites requires dsDNA to pass through a 23 Å wide channel in the HerA hexamer. The structure suggests that HerA is a dsDNA translocase that feeds DNA into the NurA nuclease sites.

© 2014 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

## 1. Introduction

Double-strand DNA breaks (DSBs) are among the most deleterious forms of DNA damage. They can arise from exposure to ionising radiation, oxygen radicals, but also from chromosomal replication. All three kingdoms possess repair pathways that are capable of repairing DSBs through either homologous recombination repair (HRR) or end joining. The HRR pathway restores the damaged chromosome to its original state without any loss or change of genetic information through the use of the undamaged chromosome as a template [1]. In contrast, the template-independent end joining pathways – essentially a ligation reaction – involve limited processing of the DNA and may result in the alteration of genetic information [2].

HRR requires long-range resection of the 5′ strand of the DNA ends. This process is catalyzed in a kingdom-specific manner by a variety of nucleases, helicases and other factors [1,3,4]. The newly generated 3′ ssDNA overhang is a substrate for the RecA/Rad51 recombinases that mediate the strand exchange with the homologous region in the undamaged chromosome. The paired 3′ tail serves as primer for templated repair synthesis, accurately restoring the disrupted genetic information.

In bacteria, the DSB resection is typically carried out by the large helicase–nuclease complexes RecBCD or AddAB, although other pathways exist [5,6]. In eukaryotes, long-range resection after commitment to HRR involves the nuclease Exo1, the nuclease-helicase Dna2 and the helicase Sgs1 in addition to other factors [7–9]. The factors responsible for long-range resection in archaea were identified through observation that the genes *mre11* and *rad50* form a conserved operon with *nurA* and *herA* [10–12]. NurA is a nuclease with predominant 5′-3′ exonuclease (dsDNA and ssDNA) and ssDNA endonuclease activities [10,13,14]. Crystal structures of *Pyrococcus furiosus* and *Sulfolobus solfataricus* NurA revealed that it forms a dimer with structural similarity to argonaute, endonuclease V and RNaseH [13,15]. A striking feature of the structure is an elongated central channel through which both unwound strands of dsDNA (but not duplex dsDNA) could conceiv-

* Corresponding authors at: Max-Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany. Fax: +49 89 8578 2641 (F. Förster). Gene Center, Feodor-Lynen-Str. 25, 81377 Munich, Germany. Fax: +49 89 2180 76999 (K.-P. Hopfner).
*E-mail addresses:* foerster@biochem.mpg.de (F. Förster), hopfner@genzentrum.lmu.de (K.-P. Hopfner).
[1] These authors contributed equally.

ably be fed into the separate active sites located within the channel. HerA is a member of the HerA/FtsK superfamily (itself a member of the Additional Strand Catalytic glutamatE (ASCE) group of ATPases) that is characterized by the presence of a central RecA-like domain containing an α-helical insert [16]. It has been denoted a bipolar helicase on the basis that it can in isolation unwind blunt dsDNA substrates in addition to dsDNA substrates with either 3′ or 5′ ssDNA overhangs [11,17].

To coordinate the nuclease and helicase activities necessary for efficient resection, NurA and HerA form a complex with 6:2 stoichiometry through the flat surface on one face of the NurA dimer and the N-terminal HAS (HerA-ATP Synthase barrel) domain of HerA [13–15]. Reconstitution of the HRR pathway from *P. furiosus* revealed that the nuclease activity of NurA within the context of a complex is modulated such that only the 5′ strand is significantly resected [14]. Similarly, the helicase activity of *S. solfataricus* HerA in complex with NurA is significantly enhanced compared to its activity in isolation [15].

The current lack of structural information on the entire HerA–NurA complex has limited our understanding of how the individual enzymatic activities are coordinated within the complex in order to bring about specific long-range resection of 5′ DNA. Using cryo-EM single particle analysis we have determined the structure of a reconstituted HerA–NurA complex from *S. solfataricus* at a resolution of 7.4 Å and created a pseudo-atomic model by fitting previously determined crystal structures and homology models followed by flexible fitting. The structure reveals the basis for the 6:2 architecture of the HerA–NurA complex, and shows that the entry to NurA's active sites requires DNA to pass through HerA. We propose that HerA uses a dsDNA translocase activity to feed DNA into the active site of NurA for processive resection.

## 2. Results

### 2.1. ATP-dependent nuclease activity

Mixing separately purified wild-type NurA and HerA (without fusion tags on either protein) produced a stable complex that was purified by gel filtration (Fig. 1a and b) and used for cryo-electron microscopic analysis. To confirm the stoichiometry of the complex we analyzed samples by size-exclusion chromatography coupled right-angle light scattering (SEC-RALS) in addition to densitometry of SDS–PAGE gels (Fig. S1). Both the molecular weight of 419kDa determined by SEC-RALS and the 3:1 HerA:NurA ratio determined by densitometry are consistent with the previously observed HerA–NurA complex with 6:2 stoichiometry (calculated molecular weight of 416kDa) [14,15].

To confirm that our proteins are active, we first performed nuclease assays with either linearized or supercoiled ΦX174 RFI DNA. Resection of linear dsDNA is only observed in the presence of NurA, HerA and ATP (Fig. 1c), confirming that resection occurs in an ATP-dependent manner and requires the activities of both NurA and HerA. In contrast, we observe no resection of supercoiled circular DNA (Fig. 1d) indicating that resection can only be initiated from a DNA end.

### 2.2. Cryo-electron microscopic analysis of the HerA–NurA complex

To elucidate the molecular organisation of the HerA–NurA complex we subjected the reconstituted complex to structure determination by single particle cryo-electron microscopy (cryo-EM). The micrographs depicted monodisperse, easily identifiable "bullet-shaped" particles, embedded in vitrified ice (Fig. 2a). A total of
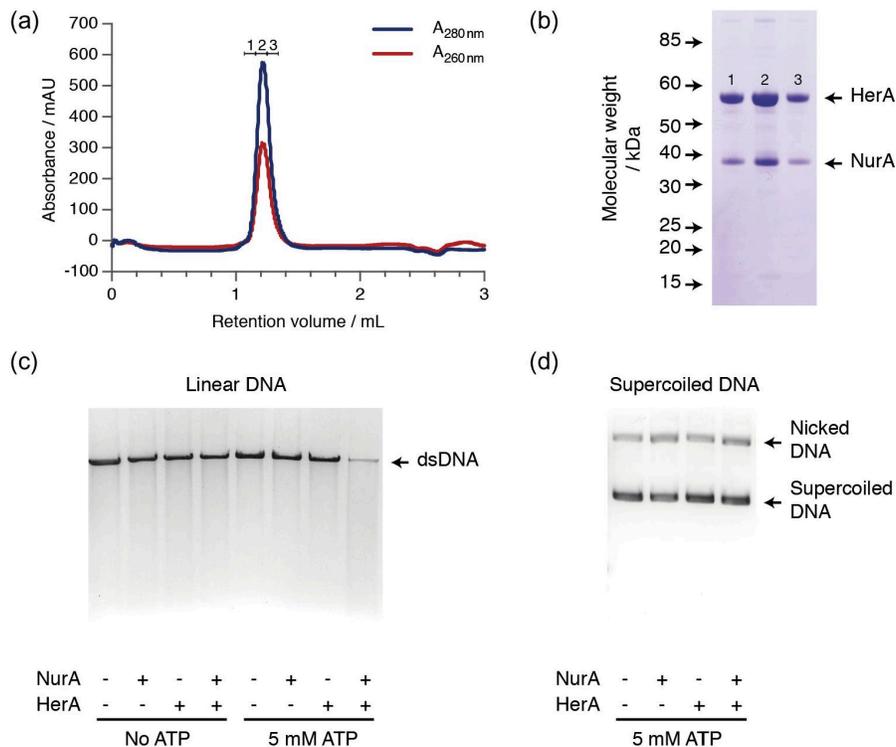


**Fig. 1.** Reconstitution of HerA–NurA complex. (a) Size-exclusion chromatography and (b) SDS–PAGE confirm the presence of the HerA–NurA complex in the fractions used for cryo-EM analysis. (c) Resection assays performed with recombinantly expressed and purified NurA and HerA confirm that both proteins are necessary for ATP-dependent resection of linearised ΦX174 RFI DNA. (d) In contrast, no resection of supercoiled ΦX174 RFI DNA is observed.
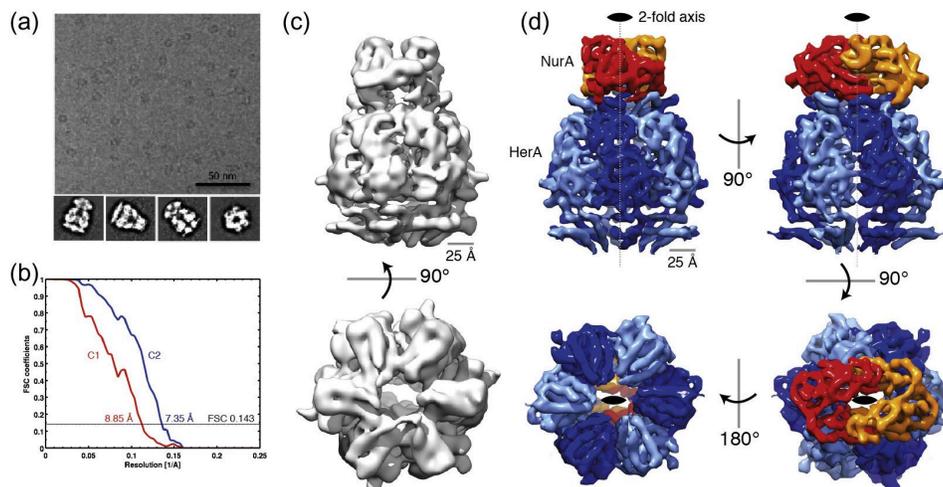
**Fig. 2.** Cryo-EM single particle analysis of HerA–NurA. (a) Part of a cryo-EM micrograph showing HerA–NurA particles. Approximately threefold magnified 2D class averages from the data are shown in the lower panel. (b) Fourier shell correlation (FSC) for two independently refined halves of the data. Shown are the curves for reconstructions with either $C_1$ (red) or $C_2$ (blue) symmetry imposed during refinement and the resolutions at which FSC = 0.143. (c) Non-symmetrized 3D reconstruction filtered to a resolution of 8.9 Å. Also shown is a 25 Å scale bar. (d) $C_2$-symmetrized reconstruction filtered to a resolution of 7.4 Å. Approximately 2-fold and 6-fold symmetrical segments are colored in orange/red and blue/light blue, respectively. Also shown are the twofold rotational symmetry axes and a 25 Å scale bar.

863 micrographs were acquired, yielding approximately 250 000 particles that were localized semi-automatically. Two-dimensional (2D) classification of the extracted particles resulted in well-defined class averages, indicating a high degree of homogeneity of the sample. To remove particles of insufficient quality (e.g. due to thicker ice or motion) for accurate alignment from the dataset, we subjected the particles to 2D classification and removed those that contributed to featureless class averages from the dataset. A total of 100 000 particles were retained for projection-matching refinement in RELION [18]. As the initial model, the density of the TrwB hexamer (PDB code 1gki) filtered to 60 Å was used because of its homology with HerA. To rule out model bias, the refinement was also performed using a generic cylinder as an initial model and this yielded an essentially identical reconstruction (Fig. S2). Refinement of an ab initio model obtained from negative stain data using the software package *SIMPLE* [19] also resulted in a similar reconstruction (Fig. S2). The three-dimensional (3D) refinement procedure yielded a reconstruction of the HerA–NurA complex at a resolution of 8.9 Å as determined by Fourier Shell Correlation (FSC) of two independently refined halves of the data (Fig. 2b). The density clearly displayed $C_2$ symmetry (Fig. 2c) and we therefore applied $C_2$ symmetry for a final round of refinement, which improved the resolution to 7.4 Å (Fig. 2d). The resulting density shows features expected for the indicated resolution, such as α-helices and β-sheets. In addition to refinement against a single 3D density, we also attempted to sort the data according to different conformers (3D classification, see Section 4) but this was not indicative of major additional conformers beyond the one we obtained from single-reference refinement.

### 2.3. Molecular architecture of the HerA–NurA complex

The HerA–NurA complex has an overall length of 140 Å and diameter of 95 Å, and harbors a central channel with a diameter of approximately 23 Å at its narrowest point (Fig. 2d). The upper part of the density clearly displays $C_2$ symmetry whereas the cylinder-shaped lower sub-complex adopts approximate $C_6$ symmetry (Fig. 2d). The overall organization of the density therefore suggests that it consists of dimeric and hexameric sub-complexes,

consistent with both our characterization of the complex and previously published data [13–15].

### 2.4. The NurA dimer

To see whether the dimeric sub-complex corresponds to two NurA molecules, we made use of the available crystal structure of *S. solfataricus* NurA (PDB code 2ygk) [20] in which it forms a dimer. We individually fitted both protomers of the NurA dimer into the EM-density for dimeric ($C_2$ symmetry) sub-complex (Fig. 3a); the secondary structure elements of the atomic model and the EM density clearly co-localize over the entire length of the model (Fig. 3b). Of note, since the crystal structure of NurA was not used as a reference in the 3D reconstruction, the excellent agreement of the NurA crystal structure and the EM density of dimeric sub-complex validates the EM analysis. Both fitted NurA protomers are arranged in the EM density in a highly similar manner to those in the crystal structure, including a similar interaction surface. Compared to the crystal structure, however, the two NurA protomers are rotated as rigid bodies by approximately 23° with respect to each other (Fig. 3c). The Fourier cross resolution (FCR) of the resulting atomic model and the corresponding segment of the EM density is 7.2 Å, consistent with the resolution of the EM map indicated by FSC. Thus, this independent validation indicates that the EM map can be interpreted to at least the 7.4 Å resolution indicated by FSC.

### 2.5. The HerA hexamer

The $C_6$ symmetric density likely corresponds to the HerA's hexamer, as previously seen in low-resolution negative stain EM images of HerA alone [12]. Since a crystal structure of HerA is not available at this time, we analyzed the domain structure of HerA using *HHPred* (Fig. 3d) [21]. HerA contains three domains: an N-terminal HerA-ATP synthase (HAS) barrel domain as observed in the N-terminus of the α-subunit of *Escherichia coli* $F_1$-ATP Synthase (PDB code 3oaa); a central RecA-like domain (residues 120–192 and 301–500); and an α-helical insert within the RecA-like domain (residues 197–290).
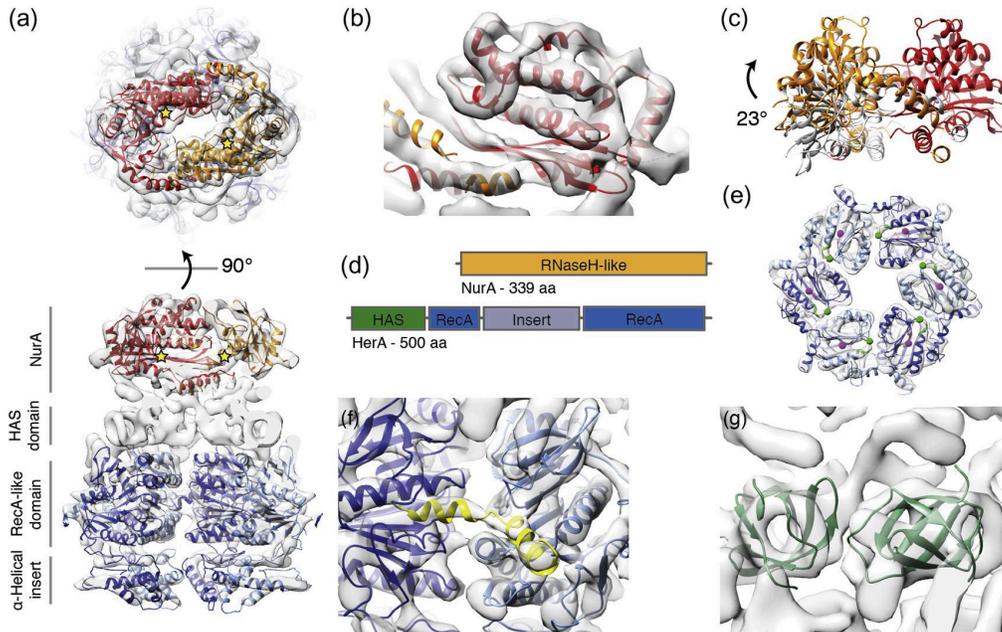
**Fig. 3.** Molecular model of HerA–NurA. (a) Top and cut-open side view of the $C_2$-symmetrized map with the fitted pseudo-atomic models. The two NurA monomers are depicted in orange and red while the the RecA-like domains and α-helical inserts of the six HerA monomers are shown in blue and light blue. The approximate locations of the active sites in NurA are indicated with yellow stars. (b) The fit of NurA reveals an excellent fit to the secondary structure elements resolved in the map. (c) In the NurA homodimer the relative orientations of the monomers differ by 23° with respect to the X-ray crystal structure (white). (d) Domain prediction indicates that HerA comprises an N-terminal HAS-domain and an RecA-like domain with an α-helical insert. (e) Fitted RecA-like domains of HerA with Walker A and Walker B motifs colored in green and pink, respectively. (f) The extreme C-terminal residues of each HerA monomer form an α-helix (yellow), protruding from the dark blue adjacent monomer. (g) Two HAS-barrels of the *E. coli* $F_1$-ATP-synthase (PDB code 3oaa) are fitted into the segment of the HerA density that corresponds to the HAS domain.

The closest structural homolog of HerA is the VirB4 ATPase (PDB code 4ag6), which pumps dsDNA as part of the Type IV secretion system [22]. The homology between VirB4 and HerA comprises the RecA-like domain as well as the α-helical insertion (sequence identity approximately 17%). A homology model of HerA based on VirB4 can be rigidly fitted into the EM density. The fit suggested that the overall fold is preserved, although the α-helices within the α-helical insert are repositioned. We subjected the VirB4-based homology model to molecular dynamics flexible fitting (MDFF) [23]. In the converged atomic model α-helices and β-sheets clearly overlap with corresponding density elements (Fig. 3a and e). The FCR of the atomic model and the corresponding region of the EM map is 7.5 Å, again consistent with the determined resolution of the EM map, indicating that over-fitting is not a major concern.

Consistent with the robust ATPase activity [12,15], HerA has a canonical RecA-like domain, with solvent accessible Walker A (residues 150–154) and Walker B motifs (353–356) at the interfaces of adjacent RecA-like domains (Fig. 3e). This location suggests that ATP binding and hydrolysis translocate DNA on the basis of conformational changes between the RecA-like domains [22,24,25]. HerA's C-terminal helix mediates interaction of adjacent RecA-like domains in the holocomplex (Fig. 3f), while the α-helical inserts are located at the NurA-distal site of the HerA hexamer. Their arrangement in the hexamer generates the central pore with a diameter of 23 Å (calculated from the map). This is the entry to the channel that spans the entire complex until it reaches the NurA dimer.

### 2.6. HerA–NurA assembly

The remaining EM density between the RecA-like domain and NurA can be assigned to the N-terminal HAS domain of HerA, consistent with biochemical data that the flat surface between the NurA dimer and the N-terminal HAS (HerA-ATP Synthase barrel) domain of HerA are critical for complex formation [13–15]. In this density, a segment that is similar to the HAS barrel within the α-subunit of *E. coli* $F_1$-ATP Synthase is clearly identifiable (PDB code 3oaa; Fig. 3g). Due to the low sequence identity of approximately 10% between the query and the template, however, an unambiguous positioning of a homology model is not possible at this point and we hence refrained from incorporating an atomic model of this domain into our model for the HerA–NurA complex.

The nature of the HerA–NurA interface is significant for three reasons. Firstly, each NurA protomer forms intimate contacts with each of the three adjacent HAS domain regions, explaining the basis for the 6:2 interaction. Secondly, the ring of HAS domains connects the NurA dimer with the RecA-like hexamer without an obvious exit channel, suggesting that NurA's active sites can only be accessed by DNA that has first passed through the entry pore and then the central channel of HerA. The positioning of the NurA's active sites additionally prevents non-specific degradation of DNA that has not passed through HerA. Thirdly, the central channel that runs from the RecA-like domain of HerA to the HAS-NurA interface varies in diameter from 23 and 36 Å (calculated from the map) and possesses no obvious DNA splitting element.

## 3. Discussion

We report here the cryo-electron microscopic structure of the HerA–NurA DNA end resection helicase–nuclease, providing a structural framework for the resection machine that generates recombinogenic 3′ DNA ends in archaea. The activity of nucleases such as those implicated in the resection of DNA ends needs to be carefully controlled, but the resection nuclease must also once activated proceed for the considerable distance of hundreds to thousands of bases. This long-range activity requires on one hand

a tight regulation of initiation, yet it also requires mechanisms to efficiently displace proteins from DNA that would otherwise interfere with the exonuclease activities. While the HRR resection machineries in all kingdoms always contain ATP-dependent helicase/translocase activities as well as nuclease activities, the underlying proteins are quite different and apparently not evolutionary related.

Long-range resection in bacteria is performed by RecBCD- and AddAB-like complexes, which are found in gram-negative and a niche of gram-positive bacteria, respectively [26,27]. The RecBCD heterotrimer is formed from the 3′–5′ helicase and bipolar nuclease RecB, the scaffold RecC, and the 5′–3′ helicase RecD, while the AddAB heterodimer is formed from the 3′–5′ helicase and nuclease AddA, and the 5′–3′ nuclease AddB. Despite their differing subunit compositions, the two enzymes utilise a common architecture to bring about DNA resection in an equivalent manner (Fig. 4c and d): the incoming DNA is split by a 'pin' so that the separated strands can enter their cognate tunnels, where they are engaged by the appropriate motor domains (RecB/AddA or RecD) before being passed towards the nuclease domain (RecB or AddA/AddB) [28,29]. In the post-χ recognition state of the bacterial complexes, during which only the 5′ strand is degraded, protection of the 3′ strand is achieved through use of an alternate exit tunnel that is distant from the RecB/AddA nuclease domains [28,30].

Considerably less is known about the molecular architecture of eukaryotic resection proteins or complexes. Analysis of resection in *Saccharomyces cerevisiae* and biochemical studies with purified proteins showed that a minimal complex capable of resecting DNA in vitro is formed by the helicase/nuclease Dna2, the RecQ family helicase Sgs1 and the single-strand binding protein RPA [31,32]. The precise functional or structural architecture of this machinery needs to be determined, but it is interesting to note that also in eukaryotes, multiple helicase/motor domains along with nucleases are implicated in DNA end resection.

The HerA–NurA complex as determined here differs significantly from both eukaryotic and bacterial resection machineries both mechanistically and structurally (Fig. 4a and d). The nuclease active sites in the HerA–NurA complex are to a large extent self-compartmentalized through the interaction of NurA with HerA, preventing uncontrolled degradation. The nuclease active sites are only accessible for DNA that passes through the entry pore of HerA, explaining why HerA–NurA only degrades linear DNA and is inactive on circular DNA. The HerA channel diameter of approximately 23 Å suggests that the incoming dsDNA passes through the entry pore and continues to be translocated as dsDNA until it encounters the interface between NurA and HerA, where it will be separated owing to the smaller diameter of the channel in NurA (Fig. 4Aand d). Thus, while the motors in RecBCD, AddAB and Dna2–Sgs1–RPA translocate on single-strand of the unwound duplex, the diameter of the channel in the HerA hexamer and the similarity of the HerA hexamer to the double-strand translocases FtsK and VirB4 argues that HerA translocates on dsDNA instead of ssDNA. This activity would elegantly explain why HerA (in the absence of NurA) was initially denoted a bipolar helicase that can
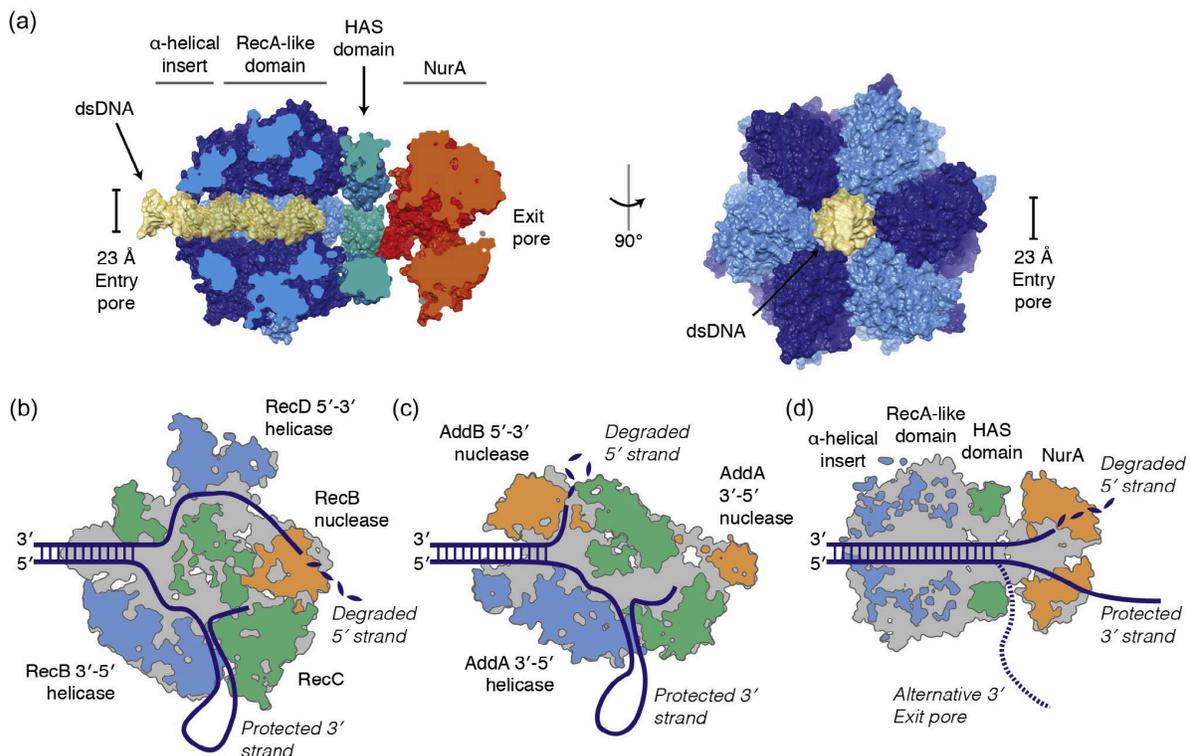


**Fig. 4.** Comparison of HerA–NurA with bacterial RecBCD and AddAB complexes. (a) Orthogonal views of HerA–NurA show that the 23 Å wide central channel is capable of accommodating the modeled dsDNA (depicted as light yellow surface). The side view is shown as a cross-section in order to visualize the central channel. The channel is 36 Å wide at the interface between NurA and the HAS-barrels and only becomes narrower within NurA. (b–d) Cross-sections showing the different architectures and mechanisms of the bacterial (b) RecBCD and (c) AddAB, and archaeal (d) HerA–NurA complexes. Motor domains are shown in light blue, nuclease domains are shown in orange, structural domains are shown in green and channels are shown in gray. In contrast to RecBCD and AddAB, the structure indicates that HerA–NurA translocates along dsDNA and that there are no obvious exit channels for the protected 3′ strand.

unwind DNA duplexes with both 3′ and 5′ tails. The proposed model of translocation is furthermore consistent with both the absence of a structural feature within HerA that separates the strands of dsDNA, and the apparently weak helicase of HerA when it is not in complex with NurA.

When DNA exits the central channel of HerA the two strands must presumably be separated before they can enter the significantly smaller channel that leads through NurA towards the exit pore. Although we observe a moderate pivot change between the NurA dimer in the crystal structure and the dimer as bound to HerA, the overall structure of the NurA dimer is remarkably preserved. It is therefore unlikely that the NurA dimer is substantially remodeled to allow passage of dsDNA. The precise way in which DNA is separated requires the structural analysis of DNA-HerA–NurA complexes and needs to be addressed in future studies (Fig. 4a and d).

Given NurA's intrinsic endonuclease activity on ssDNA it is unclear why the 5′ strand is processed but the 3′ strand is not cleaved. One option is that the 3′ strand exits the complex before passing the NurA active sites (alternative 3′ exit pore in Fig. 4d). In contrast to the bacterial AddAB and RecBCD enzymes, however, the HerA–NurA complex lacks an obvious exit between the HerA entrance and NurA exit pores that would allow escape of the 3′ strand. An alternative option is that both strands pass through the channel in NurA (Fig. 4d). The apparent symmetry of the two channels in the present reconstruction does not support the idea that there is a '5′ strand degradation' channel with a functional active site, and a '3′ strand exit' channel with an inhibited active site. It is therefore possible that the differing fates of the 5′ and 3′ strands are simply a result of their different orientations with respect to the active sites. We cannot exclude the possibility, however, that the presence of DNA may generate an asymmetry in the complex that either creates an additional exit pore for the 3′ ssDNA or inhibits one of the active sites in NurA. In addition, it will be important to visualize HerA–NurA in the presence of DNA and different nucleotides (ATP or ADP), as nucleotide binding will lead to substantial conformational changes that are linked to translocation and unwinding. The current reconstruction most likely represents an apo state although we cannot exclude that HerA retained some bound nucleotides (ATP or ADP) through purification.

Both archaeal and eukaryotic resection machineries cooperate with the ATP regulated endo/exonuclease Mre11–Rad50. In archaea, the *mre11* and *rad50* genes often form an operon with *herA* and *nurA*, suggesting that the corresponding four proteins constitute one functional machinery [10–12]. Indeed, *P. furiosus* Mre11 and Rad50 stimulate 5′ resection by HerA–NurA and these four proteins together with the Rad51 homolog RadA can promote strand exchange in vitro [14]. In a similar fashion, the activity of yeast Dna2–Sgs1–RPA is stimulated in vitro by – among others – the Mre11–Rad50–Xrs2 (MRX) complex [31,32]. Rad50–Mre11 and HerA–NurA, as well as NurA and SSB can be co-immunoprecipitated [33,34], suggesting they either directly interact or at least co-localize on DNA. The architecture of the HerA–NurA complex provides now an excellent basis to clarify the interaction architecture of HerA–NurA with DNA and other DSB repair proteins.

## 4. Material and methods

### 4.1. Molecular biology

The coding regions of *herA* and *nurA* were amplified from *S. solfataricus P2* genomic DNA with Phusion polymerase (NEB) and the resulting PCR products were gel purified. These were individually cloned using an In-Fusion HD Cloning kit (Takara) into *pETDuet-1* (Merck Millipore) linearised with the enzymes NcoI and NotI and

the reactions were used to transform XL1-Blue cells (Agilent). The plasmids were named *pETDuet-1::SsHerA* and *pETDuet-1::SsNurA* and their sequences were confirmed by Sanger sequencing (GATC Biotech).

Rosetta pLysS (DE3) cells (Merck Millipore) were transformed with either the vector *pETDuet1::SsHerA* or *pETDuet1::SsNurA* according to standard procedures. 3 L of LB media containing ampicillin (100 μg/ml) and chloramphenicol (50 μg/ml) was inoculated with an overnight starter culture at a 1:100 v:v ratio. Flasks were incubated at 37 °C until $A_{600nm} \sim 0.6$, at which point IPTG was added to a final concentration of 1 mM and the flasks were incubated at 18 °C for 18 h. Cells were subsequently harvested by centrifugation and stored at $-20$ °C.

### 4.2. Protein purification

Cells were re-suspended in lysis buffer (300 mM NaCl, 50 mM HEPES–NaOH pH 8.0, 5 mM DTT and 5% v/v glycerol) supplemented with an EDTA-Free SigmaFast Protease Inhibitor Cocktail Tablet (Sigma–Aldrich). Cells were lysed by sonication and the lysate was clarified by centrifugation. The supernatant was placed in a 70 °C water bath for 20 min and clarified by centrifugation to remove denatured protein. The supernatant was then loaded onto a 5 ml Heparin HP column (GE Healthcare) equilibrated with lysis buffer. Bound protein was eluted with a linear gradient from 100 mM to 1 M NaCl. Fractions containing the protein of interest were pooled and dialysed against 100 mM NaCl, 50 mM HEPES–NaOH pH 8.0, 5 mM DTT and 5% v/v glycerol for 18 h. The dialysed protein was loaded onto a 5 ml HiTrap Q HP column (GE Healthcare) equilibrated with lysis buffer. Bound protein was eluted with a linear gradient from 100 mM to 1 M NaCl. Fractions containing the protein of interest were pooled and dialysed against 300 mM NaCl, 20 mM HEPES–NaOH pH 8.0, 5 mM DTT and 5% v/v glycerol for 18 h. The protein was concentrated with Amicon Ultra-15 centrifugal filter units (EMD Millipore) and aliquots were flash-frozen in liquid nitrogen and stored at $-80$ °C. Protein concentrations were estimated from absorbance at 280 nm using extinction coefficients calculated from the sequence with *ProtParam*. The hexameric and dimeric states of HerA and NurA respectively were taken into account when calculating the molar concentrations of the proteins.

### 4.3. Nuclease assays

ΦX174 RFI DNA (NEB) was either used as provided or linearised with Eco1471 then purified with a NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). 10 μl nuclease assay reactions were formed on ice and contained 200 ng (5 nM) of supercoiled or linear ΦX174 RFI DNA, 80 nM NurA (dimer) and 80 nM HerA (hexamer) in a buffer containing 50 mM NaCl, 20 mM HEPES–NaOH pH 8.0, 5 mM MnCl₂, 2 mM MgCl₂, 5 mM ATP, 1 mM DTT, 100 μg ml⁻¹ BSA and 5% v/v glycerol. NurA, HerA or ATP were omitted as indicated within the figure. Reactions were briefly vortexed and centrifuged then incubated at 60 °C for 1 h. Reactions were stopped by the addition of 5 μl of 150 mM EDTA pH 8.0, 1.5% w/v SDS and 20% v/v glycerol and 1 μl of 1 mg ml⁻¹ proteinase K and incubated at 37 °C for 30 min. Samples were loaded directly onto 1% agarose gels containing GelRed (Biotium) and separated by electrophoresis in 1× TAE buffer. DNA was visualised with a UV transilluminator and the resulting images were inverted for clarity.

### 4.4. Complex formation

The HerA–NurA complex was formed by mixing 20 μM of HerA hexamer and 30 μM of NurA dimer in 300 mM NaCl, 20 mM HEPES–NaOH pH 8.0, 5 mM DTT and 5% v/v glycerol in a total vol-

ume of 150 μl. The sample was heated at 60 °C for 20 min then centrifuged. 50 μl of supernatant was loaded onto a Superdex 200 5/150 gel filtration column (GE Healthcare) equilibrated with gel filtration buffer (100 mM NaCl, 20 mM HEPES–NaOH pH 8.0) in order to separate the complex from free NurA.

### 4.5. Analysis of the complex stoichiometry

Right-angle light scattering data were collected using an ÄKTA-micro chromatography system (GE Healthcare) with inline 270 RALS and VE3580 RI detectors (Viscotek). Samples were separated with a Superdex 200 Increase 10/300 GL (GE Healthcare) gel filtration column equilibrated with 100 mM NaCl and 20 mM HEPES pH 8.0. The HerA–NurA complex was formed as described above and a volume of 100 μl was loaded for analysis. Data were analyzed with *OmniSEC* (Malvern) using 4 mg/ml BSA as a standard.

The stoichiometry of the complex was analyzed by running size-exclusion chromatography fractions on an SDS–PAGE gel. The gel was first stained in SYPRO-Orange (Life Technologies) and scanned with a Typhoon FLA 9500 (GE Healthcare), then subsequently stained with Coomassie Blue and re-scanned. Bands were quantified with *ImageJ* [35] as described previously [36].

### 4.6. Cryo-electron microscopy

Purified protein at a concentration of 0.05 mg/ml was applied to 2:1 holey carbon grids (Quantifoil Micro Tools, Germany), vitrified, and imaged at 300 keV using a Titan Krios transmission electron microscope (FEI, Netherlands) using a K2 Summit Camera (Gatan, USA) in super resolution mode. The pixel size at the micrograph level was 1.77 Å. Data were automatically acquired at defocus values ranging from 1.5 to 3.0 μm using TOM$^2$ [37]. A total dose of 25 e/A$^2$ was applied in a movie of 20 images. The images of each movie were aligned to compensate for drift and beam-induced motions using an in-house developed implementation of the algorithm in [38]. The resulting integrated images were corrected for the contrast transfer function using *ctffind3* [39]. Particles were selected semi-automatically using *e2boxer* [40]. Particles that were assigned to well-resolved 2D classes were retained for further refinement. Determination of the most probable 3D map given the observed particle images was performed in *RELION* [18] using a single class. Alternative runs of *RELION* using multiple references resulted in classification according to orientations, indicating that conformational heterogeneity of the complex was too subtle to be captured. The refinement in *RELION* converged in a map of 8.9 Å resolution according to FSC of two independently refined halves of the data. As the resulting map displayed twofold symmetry, a final refinement run employing C$_2$ symmetry converged in a map of 7.4 Å resolution according to FSC. The map has been deposited in the EMDB with the accession code 2808.

A comparative model for HerA was built based on the crystal structure of VirB4 (PDB code 4ag6) using *HHpred* [41] and *MODEL-LER* [42,43]. The N-terminal segment without structural templates was removed from the model. For the C-terminus, secondary structure analysis predicted two helices, where also density was observed. Thus these helices were additionally modelled. Initially, one monomer was rigidly fitted into the EM map with *Chimera* [44,45] followed by flexible fitting with *MDFF* [23] using a simulated annealing protocol and implicit solvent. Additional distance restraints for β-strands were introduced to keep large β-sheets together. Finally, the hexamer of HerA was built and fitted into the density. The accuracy of the model was assessed by local cross-correlation of map and model [46]. The same procedure was applied to NurA, where the crystal structure (PDB code 2ygk) was used.

### 4.7. Negative-stain electron microscopy

Purified complex at a concentration of 0.05 mg/ml was applied to R3/3 grids with 2 nm carbon support (Quantifoil Micro Tools, Germany), stained with 2% w/v uranyl acetate, and imaged at 160 kV using a CM200 transmission electron microscope (Phillips) using a TVIPS 4kx4k CCD camera. The pixel size at the micrograph level was 2.16 Å. Data were acquired at defocus values ranging from 1.1 to 2.5 μm.

The resulting images were imported into *eman2* and downsampled to give a final pixel size of 4.32 Å. 4461 particles from 10 images were picked semi-automatically with *e2boxer* and corrected for the contrast transfer function. An ab initio model was created using *SIMPLE* [19] from the resulting particle stack with *simple_rndrec* and this was refined with *simple_refine*. The refinement procedure converged after 12 cycles and the resulting model was used as an initial model for refinement against the cryo-EM data in *RELION* as described above.

### Acknowledgements

### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2014.10.035.

### References

[1] Symington, L.S. (2014) End resection at double-strand breaks: mechanism and regulation. Cold Spring Harb. Perspect. Biol. 6.
[2] Chiruvella, K.K., Liang, Z. and Wilson, T.E. (2013) Repair of double-strand breaks by end joining. Cold Spring Harb. Perspect. Biol. 5, a012757.
[3] Schiller, C.B., Seifert, F.U., Linke-Winnebeck, C. and Hopfner, K.P. (2014) Structural studies of DNA end detection and resection in homologous recombination. Cold Spring Harb. Perspect. Biol. 6, 1–25.
[4] Hopfner, K.P., Putnam, C.D. and Tainer, J.A. (2002) DNA double-strand break repair from head to tail. Curr. Opin. Struct. Biol. 12, 115–122.
[5] Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D. and Rehrauer, W.M. (1994) Biochemistry of homologous recombination in *Escherichia coli*. Microbiol. Rev. 58, 401–465.
[6] Alonso, J.C., Cardenas, P.P., Sanchez, H., Hejna, J., Suzuki, Y. and Takeyasu, K. (2013) Early steps of double-strand break repair in *Bacillus subtilis*. DNA Repair (Amst.) 12, 162–176.
[7] Gravel, S., Chapman, J.R., Magill, C. and Jackson, S.P. (2008) DNA helicases Sgs1 and BLM promote DNA double-strand break resection. Genes Dev. 22, 2767–2772.
[8] Mimitou, E.P. and Symington, L.S. (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature 455, 770–774.
[9] Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E. and Ira, G. (2008) Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell 134, 981–994.
[10] Constantinesco, F., Forterre, P. and Elie, C. (2002) NurA, a novel 5′–3′ nuclease gene linked to rad50 and mre11 homologs of thermophilic archaea. EMBO Rep. 3, 537–542.
[11] Constantinesco, F., Forterre, P., Koonin, E.V., Aravind, L. and Elie, C. (2004) A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea. Nucleic Acids Res. 32, 1439–1447.
[12] Manzan, A., Pfeiffer, G., Hefferin, M.L., Lang, C.E., Carney, J.P. and Hopfner, K.P. (2004) MlaA, a hexameric ATPase linked to the Mre11 complex in archaeal genomes. EMBO Rep. 5, 54–59.
[13] Chae, J., Kim, Y.C. and Cho, Y. (2012) Crystal structure of the NurA-dAMP-Mn2+ complex. Nucleic Acids Res. 40, 2258–2270.
[14] Hopkins, B.B. and Paull, T.T. (2008) The *P. furiosus* mre11/rad50 complex promotes 5′ strand resection at a DNA double-strand break. Cell 135, 250–260.

[15] Blackwood, J.K., Rzechorzek, N.J., Abrams, A.S., Maman, J.D., Pellegrini, L. and Robinson, N.P. (2012) Structural and functional insights into DNA-end processing by the archaeal HerA helicase–NurA nuclease complex. Nucleic Acids Res. 40, 3183–3196.

[16] Iyer, L.M., Makarova, K.S., Koonin, E.V. and Aravind, L. (2004) Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. Nucleic Acids Res. 32, 5260–5279.

[17] Zhang, S., Wei, T., Hou, G., Zhang, C., Liang, P., Ni, J., Sheng, D. and Shen, Y. (2008) Archaeal DNA helicase HerA interacts with Mre11 homologue and unwinds blunt-ended double-stranded DNA and recombination intermediates. DNA Repair (Amst.) 7, 380–391.

[18] Scheres, S.H. (2012) A Bayesian view on cryo-EM structure determination. J. Mol. Biol. 415, 406–418.

[19] Elmlund, H., Elmlund, D. and Bengio, S. (2013) PRIME: probabilistic initial 3D model generation for single-particle cryo-electron microscopy. Structure 21, 1299–1306.

[20] Blackwood, J.K., Rzechorzek, N.J., Bray, S.M., Maman, J.D., Pellegrini, L. and Robinson, N.P. (2013) End-resection at DNA double-strand breaks in the three domains of life. Biochem. Soc. Trans. 41, 314–320.

[21] Soding, J., Biegert, A. and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33, W244–W248.

[22] Wallden, K., Williams, R., Yan, J., Lian, P.W., Wang, L., Thalassinos, K., Orlova, E.V. and Waksman, G. (2012) Structure of the VirB4 ATPase, alone and bound to the core complex of a type IV secretion system. Proc. Natl. Acad. Sci. U.S.A. 109, 11348–11353.

[23] Trabuco, L.G., Villa, E., Mitra, K., Frank, J. and Schulten, K. (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. Structure 16, 673–683.

[24] Gomis-Ruth, F.X., Moncalian, G., Perez-Luque, R., Gonzalez, A., Cabezon, E., de la Cruz, F. and Coll, M. (2001) The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. Nature 409, 637–641.

[25] Lyubimov, A.Y., Strycharska, M. and Berger, J.M. (2011) The nuts and bolts of ring-translocase structure and mechanism. Curr. Opin. Struct. Biol. 21, 240–248.

[26] Dillingham, M.S. and Kowalczykowski, S.C. (2008) RecBCD enzyme and the repair of double-stranded DNA breaks. Microbiol. Mol. Biol. Rev. 72, 642–671 (Table of Contents).

[27] Wigley, D.B. (2013) Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. Nat. Rev. Microbiol. 11, 9–13.

[28] Saikrishnan, K., Yeeles, J.T., Gilhooly, N.S., Krajewski, W.W., Dillingham, M.S. and Wigley, D.B. (2012) Insights into Chi recognition from the structure of an AddAB-type helicase–nuclease complex. EMBO J. 31, 1568–1578.

[29] Singleton, M.R., Dillingham, M.S., Gaudier, M., Kowalczykowski, S.C. and Wigley, D.B. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. Nature 432, 187–193.

[30] Yang, L., Handa, N., Liu, B., Dillingham, M.S., Wigley, D.B. and Kowalczykowski, S.C. (2012) Alteration of chi recognition by RecBCD reveals a regulated molecular latch and suggests a channel-bypass mechanism for biological control. Proc. Natl. Acad. Sci. U.S.A. 109, 8907–8912.

[31] Cejka, P., Cannavo, E., Polaczek, P., Masuda-Sasa, T., Pokharel, S., Campbell, J.L. and Kowalczykowski, S.C. (2010) DNA end resection by Dna2–Sgs1–RPA and its stimulation by Top3-Rmi1 and Mre11–Rad50–Xrs2. Nature 467, 112–116.

[32] Niu, H., Chung, W.H., Zhu, Z., Kwon, Y., Zhao, W., Chi, P., Prakash, R., Seong, C., Liu, D., Lu, L., Ira, G. and Sung, P. (2010) Mechanism of the ATP-dependent DNA end-resection machinery from *Saccharomyces cerevisiae*. Nature 467, 108–111.

[33] Quaiser, A., Constantinesco, F., White, M.F., Forterre, P. and Elie, C. (2008) The Mre11 protein interacts with both Rad50 and the HerA bipolar helicase and is recruited to DNA following gamma irradiation in the archaeon *Sulfolobus acidocaldarius*. BMC Mol. Biol. 9, 25.

[34] Wei, T., Zhang, S., Zhu, S., Sheng, D., Ni, J. and Shen, Y. (2008) Physical and functional interaction between archaeal single-stranded DNA-binding protein and the 5′–3′ nuclease NurA. Biochem. Biophys. Res. Commun. 367, 523–529.

[35] Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675.

[36] Peterson, T. (2010) Densitometric Analysis using NIH Image. North American Vascular Biology Organisation (NAVBO) Newsletter, 16. http://www.navbo.info/DensitometricAnalysys-NIHimage.pdf.

[37] Korinek, A., Beck, F., Baumeister, W., Nickell, S. and Plitzko, J.M. (2011) Computer controlled cryo-electron microscopy–TOM(2) a software package for high-throughput applications. J. Struct. Biol. 175, 394–405.

[38] Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A. and Cheng, Y. (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat. Methods 10, 584–590.

[39] Mindell, J.A. and Grigorieff, N. (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol. 142, 334–347.

[40] Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I. and Ludtke, S.J. (2007) EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. 157, 38–46.

[41] Remmert, M., Biegert, A., Hauser, A. and Soding, J. (2012) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat. Methods 9, 173–175.

[42] Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U. and Sali, A. (2006) Comparative protein structure modeling using Modeller. Curr. Protoc. Bioinformatics (Chapter 5: Unit 5.6).

[43] Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815.

[44] Goddard, T.D. and Ferrin, T.E. (2007) Visualization software for molecular assemblies. Curr. Opin. Struct. Biol. 17, 587–595.

[45] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera – a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

[46] Beck, F., Unverdorben, P., Bohn, S., Schweitzer, A., Pfeifer, G., Sakata, E., Nickell, S., Plitzko, J.M., Villa, E., Baumeister, W. and Forster, F. (2012) Near-atomic resolution structural model of the yeast 26S proteasome. Proc. Natl. Acad. Sci. U.S.A. 109, 14870–14875.

# Supplementary Information

## Molecular architecture of the HerA–NurA DNA double-strand break resection complex

Robert Thomas Byrne [a,1], Jan Michael Schuller [b,1], Pia Unverdorben [b], Friedrich Förster [b*],

Karl-Peter Hopfner [a,c,*]

[a] Gene Center and Department of Biochemistry, Ludwig-Maximilians-University Munich, Feodor-Lynen-Straße 25, D-81377 Munich, Germany

[b] Department of Molecular Structural Biology, Max-Planck Institute of Biochemistry, D-82152 Martinsried, Germany

[c] Center for Integrated Protein Sciences, Munich, Germany
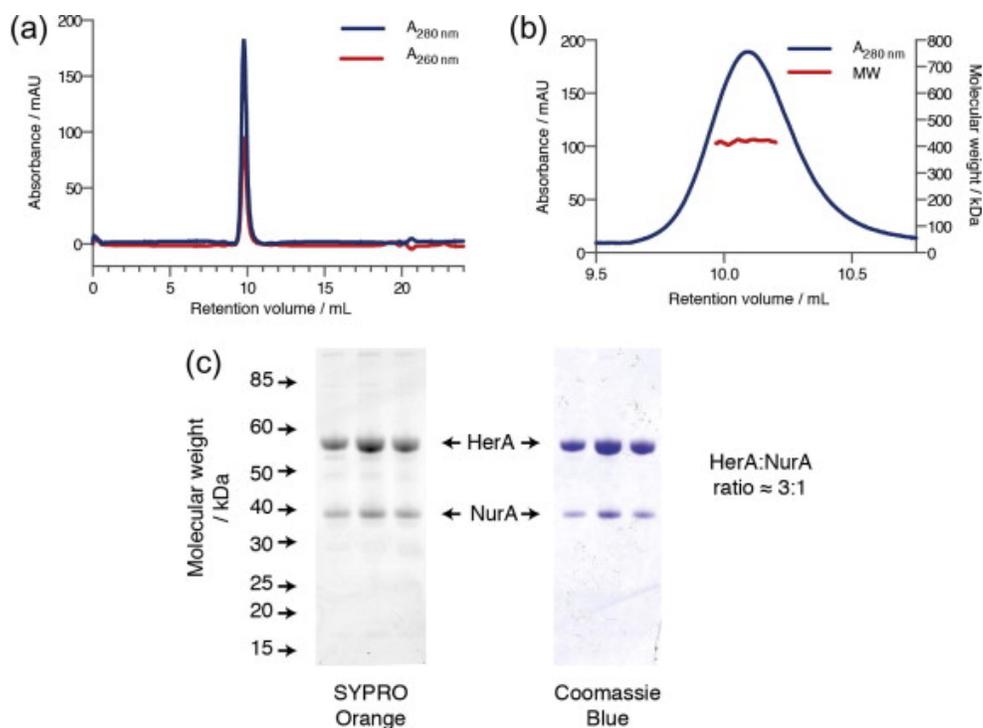
[*] Corresponding authors at: Max-Planck Institute of Biochemistry, Am Klopfer-spitz 18, 82152 Martinsried, Germany. Fax: +49 89 8578 2641 (F. Förster). Gene Center, Feodor-Lynen-Str. 25, 81377 Munich, Germany. Fax: +49 89 2180 76999 (K.-P. Hopfner).

E-mail addresses:

foerster@biochem.mpg.de (F. Förster), hopfner@genzentrum.lmu.de

(K.-P. Hopfner).

[1] These authors contributed equally.

**Supplementary Fig. 1.** HerA–NurA has 6:2 stoichiometry. (a) Size-exclusion chromatogram of the HerA–NurA complex and (b) right-angle light scattering (RALS) data for the peak shown in (a) indicates that the species has a molecular weight of 419kDa. (c) Densitometry of peak fractions reveals that the peak fractions have average HerA:NurA ratios of 3.8:1 (Coomassie Blue) and 2.9:1 (SYPRO Orange). The data are therefore most consistent with a HerA–NurA complex with 6:2 stoichiometry (calculated molecular weight of 416kDa).
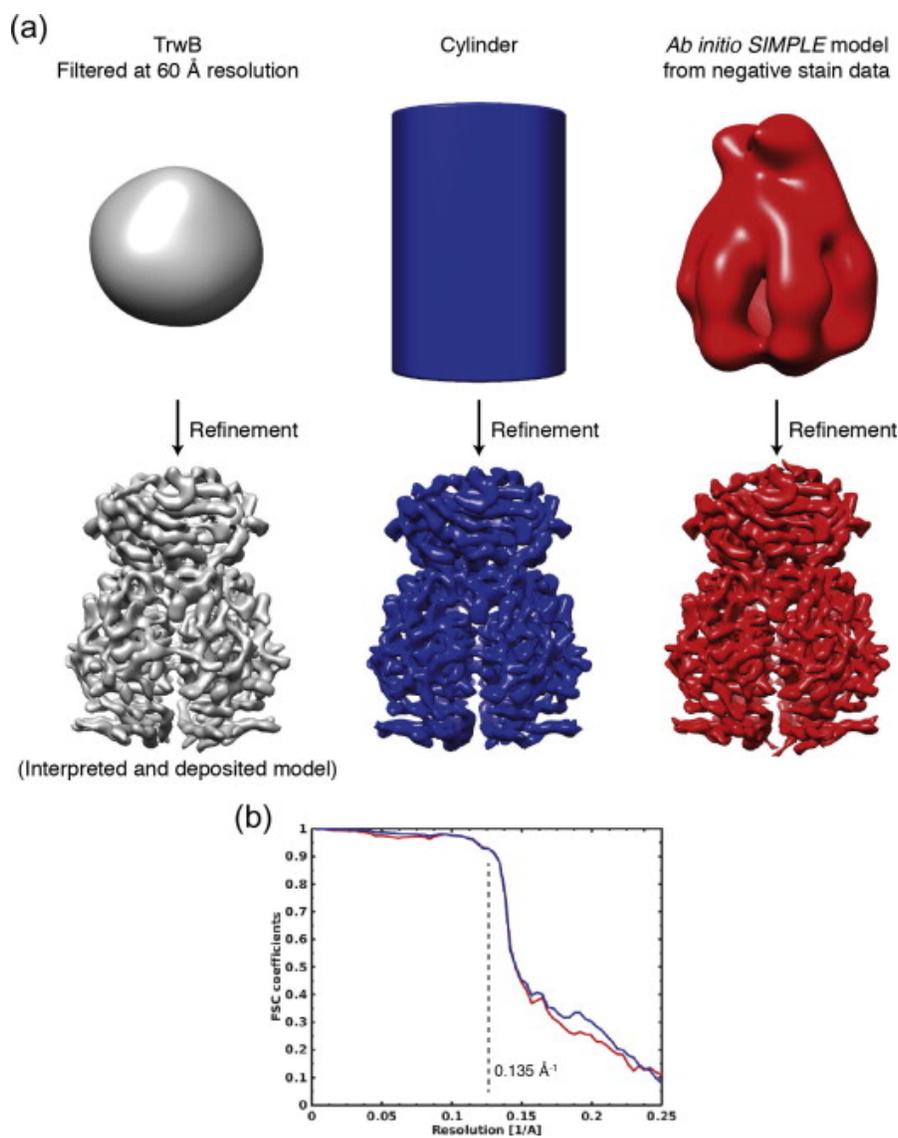
**Supplementary Fig. 2.** Influence of initial model on refined density. (a) Gallery of initial densities and results of respective refinement in RELION. The initial models are the density of the TrwB hexamer (PDB code 1gki) filtered to 60 Å resolution (left, gray), a cylinder with the approximate dimensions of the HerA–NurA complex (center, blue), and a reference-free reconstruction from negative stain data obtained by SIMPLE (right, red). (b) Fourier Shell Correlation (FSC) between either the cylinder-based refined structure (blue curve) or the final SIMPLE-based structure (red curve) and the TrwB-based structure interpreted within this manuscript and deposited in the EMDB. The vertical line indicates the resolution (7.35 Å) reported in the manuscript, indicating that all densities are essentially identical to this frequency.

## 2.2 Part II: Nucleotide-dependent conformational changes of the AAA+ ATPase p97 revisited

This article was published in 2016 in FEBS Letters (Volume 590, Issue 5, Pages 595-604. The supplemental material is attached at the end of the article.

**Summary:**

In this article, we use single-particle cryo-EM analysis to investigate nucleotide-binding induced conformational changes of the AAA+ protein p97. It is a major facilitator of proteasomal degradation, by extracting polyubiquitylated proteins from their cellular environment. Previous attempts to study p97 by X-ray crystallography revealed a homohexameric, six-fold symmetrical arrangement, and only minor structural differences between the various nucleotide-binding states. Crystallographic analyses are in conflict with lower resolution cryo-EM experiments of p97 and related AAA+ enzymes, which suggest large conformational changes. In the crystal packing forces likely bias the positioning of the N-terminal domains. We determined the solution structure of p97 in the presence of ADP, AMP-PNP and ATP-γS at resolutions of 6.9, 7.5 Å and 6.1 Å respectively. In all structures, p97 is essentially six-fold symmetrical and the pore of the D1 ring remains closed irrespective of the bound nucleotide. In contrast, the D2 ring shows an iris-like closing of its center upon binding of ATP-γS. The largest conformational changes within the p97 molecule are large-scale swinging motions of the N-domains. We suggest that the nucleotide hydrolysis induced N-domain motion together with substrate-recruiting cofactors may enable segregation of ubiquitinated substrates from their environment, whereas the D2-cavity may contribute to further processing.

**Contribution:**

For this article, I carried out grid preparation for electron microscopy, acquisition of cryo-EM data and the complete workflow of single particle analysis and map interpretation. I wrote major parts of the manuscript and prepared all figures.

# Nucleotide-dependent conformational changes of the AAA+ ATPase p97 revisited

Jan M. Schuller[1], Florian Beck[1], Philip Lössl[2], Albert J. R. Heck[2] and Friedrich Förster[1]

1 Department of Molecular Structural Biology, Max-Planck Institute of Biochemistry, Martinsried, Germany
2 Biomolecular Mass Spectrometry and Proteomics and Netherlands Proteomics Center, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, The Netherlands

The ubiquitous AAA-ATPase p97 segregates ubiquitylated proteins from their molecular environment. Previous studies of the nucleotide-dependent conformational changes of p97 were inconclusive. Here, we determined its structure in the presence of ADP, AMP-PNP, or ATP-γS at 6.1–7.4 Å resolution using single particle cryo-electron microscopy. Both AAA domains, D1 and D2, assemble into essentially six-fold symmetrical rings. The pore of the D1-ring remains essentially closed under all nucleotide conditions, whereas the D2-ring shows an iris-like opening for ADP. The largest conformational changes of p97 are 'swinging motions' of the N-terminal domains, which may enable segregation of ubiquitylated substrates from their environment.

**Keywords**: AAA-ATPase; Cdc48; Cryo-EM; ERAD; protein quality control

The type II AAA+ protein p97/Cdc48 (mammalian/yeast) is a chaperone-like molecular machine that utilizes the chemical energy released by ATP hydrolysis for segregation of ubiquitylated proteins from their environment, such as protein complexes and membranes [1,2]. P97 copies assemble into homohexameric complexes, which acquire substrate and processing specificity through a large number of transiently associating cofactors [1,3,4]. Single amino acid mutations in human p97 (also referred to as VCP) have been linked to neurodegenerative disorders, such as inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia (IBMPFD) [5].

P97 consists of an N-terminal effector binding domain (N-domain) and two AAA domains (D1 and D2). X-ray crystallographic studies revealed the architecture of these domains in the homohexameric complex [6–8]. D1 and D2 both form six-fold

homohexameric rings, which are stacked on top of each other in a 'head-to-tail' arrangement. This symmetrical arrangement is in contrast to some other AAA-ATPases, including the closely related N-ethylmaleimide sensitive factor (NSF) [9], which adopt staircase-like topologies that undergo changes during their nucleotide cycles. In the full-length X-ray crystallographic structures, the structural differences of the AAA domains of p97 are comparatively small between the different nucleotide states [7,8]. In all full-length structures, the centers of mass of the N-domains are arranged essentially coplanar with those of D1. However, the positioning of the N-domain differs in the X-ray structure of the N-D1 fragment carrying the IBMPFD associated point mutations (R155H or A232E) in the presence of ATP-γS, in which the centers of mass of the N-domains resides significantly above the plane of the centers of mass of D1 [10,11].

While the present X-ray crystallographic structures provide accurate structural information about the single domain structures, the quaternary structure observed in the crystals might not be physiological. Low-resolution small-angle X-ray scattering (SAXS) [12], atomic force microscopy [13] and cryo-electron microscopy (EM) studies [14,15] all suggest much larger conformational differences between the different nucleotide states than observed in the full-length X-ray crystallographic structures. A likely source for the discrepancy between the scales of motion observed in the crystallographic and solution data are the extensive contacts between the D2 and N-domains of adjacent particles in the crystal lattices [7,8]. In particular, the positions of the N-domains seem to be strongly biased by the crystal packing. However, due to insufficient resolution, solution studies have so far failed to conclusively address the precise domain positions and their internal rearrangements for the different nucleotide states.

In this study, we revisit the conformational landscape of p97 in the presence of three nucleotides, AMP-PNP, ATP-γS, and ADP, using cryo-EM single particle analysis and recent advances in electron detector technology [16,17].

## Material and methods

### Sample preparation

Full-length p97 was purified as described previously [7]. About 4 μL of purified protein was applied to a lacey carbon copper grid (Quantifoil, Großlöbichau, Germany), incubated for 30 s, manually blotted with filter paper and washed two times with 4 μL buffer (containing the respective nucleotide) and 0.05% NP40, before final blotting and vitrification in liquid ethane.

### Mass spectrometry

P97 was buffer exchanged to 150 mM $CH_3CO_2NH_4$ pH 7.5 (for native MS) or 0.1% ($v/v$) HCOOH (for MS under denaturing conditions) using Micro Bio-Spin P30 columns (Bio-Rad, Hercules, CA, USA). To study nucleotide binding by native MS, p97 was incubated with 1 mM of the respective nucleotide and 1 mM $MgCl_2$ for 30 min at room temperature prior to buffer exchange. The buffer exchanged samples were loaded in gold-coated borosilicate capillaries and were analyzed by nano-electrospray ionization-MS at a protein concentration of 4–5 μM using a modified Orbitrap Exactive Plus EMR mass spectrometer (Thermo Scientific, Waltham, MA, USA), externally calibrated with CsI clusters [18]. The following instrument settings were used for native MS: capillary voltage = 1.3–1.4 kV, source fragmen-

tation voltage = 20 V, S-lens extraction voltage = 150–200 V, HCD energy = 50 V, $N_2$ gas pressure in HCD cell = $7 \times 10^{-10}$ bar, mass resolution at 200 $m/z$ = 6000. For MS under denaturing condition, the energy and gas pressure in the HCD cell were set to minimum. The mass spectra were analyzed with XCALIBUR v2.2 and the mass of denatured p97 was determined using PROTEIN DECONVOLUTION v3.0 (both Thermo Scientific).

### Electron microscopy

All data were collected on a 300 keV FEI Titan Krios electron microscope equipped with a K2 Summit detector (Gatan, Pleasanton, CA, USA) operated in counting mode (10 $e^-$ pixel$^{-1} \cdot s^{-1}$, 30 frames with a total dose of 40 $e^-/\text{Å}^{2-}$). The data sets of p97-AMP-PNP and p97-ADP were collected at a magnification of 81 000× (1.74 Å·pixel$^{-1}$), whereas the p97-ATP-γS data were collected at 105 000× magnification (1.35 Å·pixel$^{-1}$). TOM$^2$ automated acquisition software [19] was used to collect a total of 2464 (p97-AMP-PNP), 3276 (p97-ADP) and 1728 (p97-ATP-γS) micrographs.

### Image analysis

From micrographs that did not show astigmatism or strong drift particles were selected automatically [20]. CTF parameter determination was carried out on the whole micrograph level using *ctffind3* [21]. All subsequent processing of the particles was carried out in *RELION*. To evaluate conformational heterogeneity 3D classification was carried out using 50 classes. For subsequent analysis we aligned the protomers of the 50 classes based on the pseudo six-fold rotational symmetries (Fig. S2), resulting in 300 volumes. The densities inside the mask encapsulating the N-domain of one protomer (approximately 4 nm in diameter) were then analyzed statistically by principal component analysis of the correlation matrix and subsequent hierarchical clustering according to the first eigenvector in TOM [22]. For the AMP-PNP, ADP, and ATP-γS datasets, resolutions of 9.3, 9, and 8.1 Å were determined, respectively. The nonsymmetrized maps were subjected to rotational correlation analysis in TOM [22]. Subsequently, six-fold rotational symmetry was imposed for a final refinement round increasing the resolution to 7.4, 6.8, and 6.1 Å for AMP-PNP, ADP, and ATP-γS datasets, respectively. The local resolution was computed using Bsoft [23], and the maps were filtered accordingly. The EM densities have been deposited in the EMDataBank for the ADP state (EMDB-3323, EMDB-3326), AMP-PNP state (EMD-3324, EMDB-3328), and ATP-γS state (EMD-3325, EMDB-3327). Molecular graphics and analyses were performed with the UCSF CHIMERA package [24]. The consistency of all fits was analyzed by Fourier cross-resolution of the reconstruction and the density simulated from the atomic model in TOM [22].

39

# Results

## Characterization and data acquisition of pre- and posthydrolysis states

To correlate our data with prior structural data and biochemical characterization we used incubated purified p97 with three extensively studied nucleotides: (a) The posthydrolysis state was induced by adding excess ADP to the buffer ('ADP dataset') [7,14,15]. Biochemical studies indicate that all 12 nucleotide binding sites of the p97 hexamer have ADP molecules bound under the conditions of the experiment [12]. (b) To mimic an ATP-bound state the nonhydrolysable ATP-analog AMP-PNP has often been employed [7,12,14]. However, the affinity of AMP-PNP for p97 is much lower than that of ATP and ADP [25]. For our experimental conditions, approximately five ADP and five AMP-PNP molecules have been detected per p97 particle [12]. (c) The slowly hydrolysable analog ATP-γS is an alternative means to lock a prehydrolysis state. The affinity of ATP-γS is comparable to that of ATP and biochemical studies indicate that 9–10 ATP-γS nucleotides are bound per p97 molecule under the conditions of the sample [26].

Prior to structural analysis, we characterized the nucleotide states of our samples after purification and in the presence of ADP and ATP-γS by native mass spectrometry (MS) [27,28]. AMP-PNP was omitted in this experiment due to the anticipated strong heterogeneity of nucleotide states, which is not to be easily resolved by native MS. To determine the precise mass of apo p97 its monomeric mass was first measured after denaturation (Fig. S1A). The most abundant peak among four species is a form that likely corresponds to p97, which has been acetylated after removal of the N-terminal Met-Gly. Because each of the four monomeric species may be incorporated in the p97 hexamer, their weighted mass average was used to calculate the averaged expected mass of the hexameric apo p97 (548 491 Da).

We then aimed to analyze to which extent nucleotides remain prebound to p97 after recombinant expression and purification in nucleotide-free buffer. The major population in the recorded spectra corresponds to hexameric p97 with 10 ADP molecules bound (Fig. S1B). Upon incubation with ADP the spectrum indicates that the major mode becomes loaded with 11 ADP molecules (Fig. 1A). In both cases, only distinct ADP loading states are present, suggesting cooperative nucleotide binding.

When purified p97 was incubated with ATP-γS the spectra indicate exchange of the prebound ADP by ATP-γS. As a result, the major subspecies is shifted to higher molecular weights, indicating that the prebound ADP has been fully or partially replaced by ATP-γS (Fig. 1C). Assuming complete exchange of ADP by ATP-γS, the most abundant species can be assigned to p97 with 10 $Mg^{2+}$-ATP-γS bound, again indicating cooperative nucleotide exchange and binding (Fig. 1B).

## Conformational variability at different nucleotide conditions

Because statistical distribution of different monomer conformers within the homohexameric ring can give rise to a large number of different overall configurations, we classified each dataset into 50 different 3D classes. To identify the major monomer conformers we subsequently superposed the protomers of the ring-shaped class reconstructions according to their respective rotations around their six-fold pseudosymmetry axis in the center of the molecule and subjected them to hierarchical clustering focusing on an area encompassing their N-domains (Fig. S2A).

For the ADP dataset, this analysis indicates that the computed reconstructions from the ADP dataset are highly similar to each other and hence conformational heterogeneity is negligible (Fig. S2B). In all classes, the centers of mass of all N-domains are in plane with those of the D1 domain ('down' state in the following) akin to the X-ray crystallographic structures of full-length p97 [7].

In contrast, the analysis of the N-domain positioning indicates heterogeneity for the AMP-PNP dataset. The monomers adopt two preferred conformations that mostly differ in the positioning of the N-domains (Fig. S2C): their centers of mass position either above the plane of the centers of mass of the D1 domains ('up' state in the following) or in the 'down' state. In addition to these two major conformers, the classification reveals another less populated and less-defined class that resembles the up-state. This class corresponding to a 'background' of heterogeneous conformations ('flexible' state) is not considered in the further analysis. In contrast to the positions of the N-domains, the AAA domains do not show notable differences. For the ATP-γS dataset the classification reveals that only very few monomers adopt a defined up-conformation (Fig. S2D). The most populated classes show the N-domain also in a 'flexible' upward position, but much more washed-out, or the N-domain is not resolved at all. Thus, the conformational space of the N-domain is considerably larger in the presence of ATP-γS than for AMP-PNP and ADP and it primarily

40
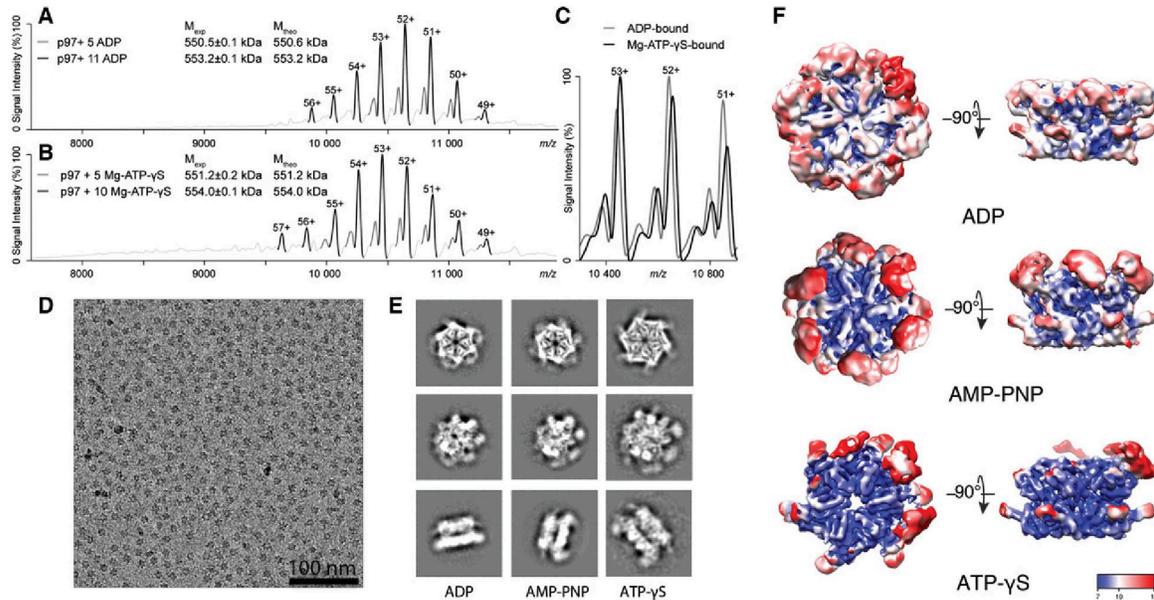
**Fig. 1.** Native MS and Cryo-EM single particle analysis of p97 after incubation with different nucleotides. (A) Broad-range native mass spectrum of ADP-bound p97. The peaks are labeled with their respective charge states. Listed are the experimentally determined masses $\pm$ standard deviation ($M_{exp}$) and the theoretical masses ($M_{theo}$) of the identified species, which were calculated based on the mass of apo-p97 (Fig. S1A. (B) Broad-range native mass spectrum of Mg-ATP-γS-bound p97. (C) Overlay of three charge states representing ADP- and Mg-ATP-γS-bound p97. All peaks are shifted to higher *m/z* values upon incubation with Mg-ATP-γS. (D) Part of a cryo-EM micrograph of p97 in presence of ATP-γS. The bar corresponds to 100 nm. (E) Reference-free 2D class averages from three data sets with p97 bound to different nucleotides. (F) Side- and top-views of the ADP-, AMP-PNP-, and ATP-γS-bound structures, which are colored according to their local resolution.

contains upward facing states. In contrast to the ADP and AMP-PNP datasets, the down-conformer is entirely absent in the ATP-γS dataset.

## Subnanometer resolution reconstructions with and without imposed six-fold symmetry

As the ADP data were found to be structurally homogeneous, we merged the particles from the best-resolved classes, yielding an 8.4 Å resolution reconstruction (Fig. 1F, Figs S3 and S5). To investigate the symmetry of the AAA core, we masked out the N-domains and computed the rotational autocorrelation function, which showed pronounced peaks at multiples of 60° (Fig. S4). This analysis indicates that the AAA core is essentially identical upon six-fold rotation to a resolution of 9 Å coinciding with the FSC estimate for the reconstruction. Subtle differences in the densities of the N-domains of the individual protomers are not significant at the lower local resolution, which does not reach subnanometer level (Fig. S3). Thus, at the resolution determined here no deviations from six-fold symmetry can be discerned, which does not exclude

the possibility of subtle deviations at higher resolutions. Hence, we realigned the particles with imposed six-fold rotational symmetry ($C_6$) resulting in a 6.9 Å resolution 'ADP density' (Fig. 2 and Fig. S5).

Among the classes representing the AMP-PNP dataset, a class with 'all-up' N-domains clearly yields the highest resolution indicating that the corresponding particles have the highest degree of homogeneity. From the ~ 20 000 particles contributing to the all-up class we obtained a reconstruction of 9.3 Å resolution (Fig. 1F, Figs S3 and S5). The symmetry analysis of the AAA core and N-domain again indicates six-fold symmetry at the achieved resolution (Fig. 1F and Fig. S4). Accordingly, we imposed six-fold symmetry for final alignment of these particles, which improved the resolution to 7.5 Å (Fig. 2 and Fig. S5).

As the N-domains in the ATP-γS data were determined to be structurally heterogeneous, we subsequently focused the alignment on the AAA core for this dataset. Classification did not indicate structural heterogeneity in the AAA core and the ~ 40 000 particles from the best-resolved classes yielded a reconstruction of 8.1 Å resolution (Fig. 1F, Figs S3 and S5). The

41
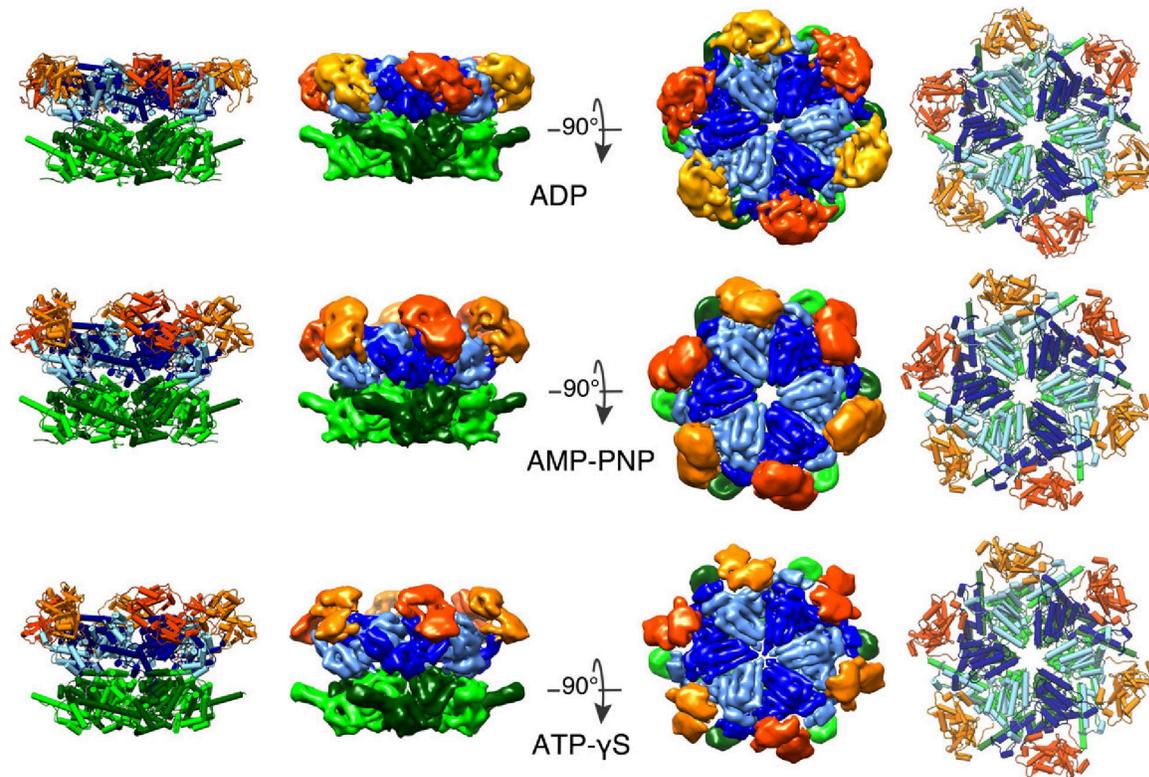
**Fig. 2.** p97 densities refined with imposed six-fold symmetry and corresponding pseudoatomic models. Center columns: the maps were segmented and colored according to their domains in alternating dark and light color tones. (orange: N-domain, blue: D1 domain, green: D2 domain). Left and right columns: the fitted atomic models are depicted in the same color code.

subsequent symmetry analysis of the AAA core showed six-fold rotational symmetry (Fig. 2, Fig. S4) and the resolution of the refined symmetrized reconstruction improved to 6.1 Å (Fig. 2 and Fig. S5).

### Conformational changes of the N-D1 segment

To interpret the structural rearrangements of p97 for the different nucleotides we sought to explain the observed EM density by the existing crystallographic structures. For the ADP densities, the crystal structure of full-length p97 bound to ADP (PDB: 3cf3) yielded the best fit. The Fourier cross-resolution (FCR) of this atomic model with the EM data (6.8 Å) shows that it explains the ADP density to the reported resolution (6.9 Å) (Figs S5 and S6A).

In contrast, the AMP-PNP density and the X-ray crystallographic model of p97 bound to AMP-PNP (PDB: 3cf2) differ substantially, most notably in the positioning of the N-domain. The structures of the IBMPFD associated p97 N-D1 mutants (R155H or

A232E) bound to ATP-γS (PDB: 4ko8/4kln) [10], which both have the N-domain positioned at an elevated position, fit the N-D1 segment of the AMP-PNP density best yielding a FCR (7.4 Å), consistent with the resolution estimate (7.5 Å) (Figs S5 and S6B). The crystal structures of the ATP-γS-bound IBMPFD p97 N-D1 mutants with the elevated N-domain also explain the ATP-γS density best, markedly better than the full-length ATP-γS-bound Δ709–728 p97 crystal structure [8] (Fig. 2, Fig. S6C).

The largest structural difference between the N-D1 segments of the cryo-EM ADP- and AMP-PNP/ATP-γS structures is the repositioning of the N-domains: they translate upwards from the D1 ring by 12.5 Å (center of gravity of the domain) and rotate by 90°, resulting in a 'swinging' motion (Fig. 3A). The conformational differences between the D1 domains in the AMP-PNP/ATP-γS and the ADP structures are comparatively small with displacements of the Cα-atoms below 5 Å throughout. The largest changes involve the helices H12 and H13 of the small subdomain of D1,

42

which form a major interface with the N-domain in the ADP state. A notable feature of all D1 conformations observed here is that the D1 pore is extremely narrow, in agreement with the fitted crystal structures.

Limited proteolysis experiments on the IBMPFD p97 mutants have implicated ordering of the N-terminal residues upon ATP binding [29]. Indeed, the AMP-PNP density shows a well-defined density segment (Fig. 3B), for which the fitted atomic model suggests that it corresponds to the very N-terminal residues (residues 1–20). These N-terminal 20 residues are absent from all p97 crystal structures but one protomer of the ATP-γS-bound IBMPFD mutant A232E, which includes residues 12–20 [10]. These eight residues coincide with part of the segment. Tracing it further suggests that the very N-terminus of p97 is positioned between the helices H1 and H2 of the D2 domain and the peptide further

traverses adjacent to the loop connecting H12 and H13 of the D1 domain. Thus, our cryo-EM structure confirms that the ATP-induced ordering of the N-termini is a genuine feature of p97.

## Conformational changes of the D2 domain

For the ADP and the AMP-PNP D2 ring, the best fitting atomic model is the p97-ADP crystal structure (PDB: 3cf3). The FCRs for this model and the ADP and AMP-PNP maps (6.8 and 7.4 Å, respectively) both match the determined resolution (Fig. S6). The best fit to the ATP-γS D2 cryo-EM density is the ATP-γS-bound Δ709–728 p97 crystal structure (PDB: 5C18, Fig. S6) [8], yielding the FCR for this atomic model and the ATP-γS map (7 Å) is in the range of the determined resolution, indicating that the model
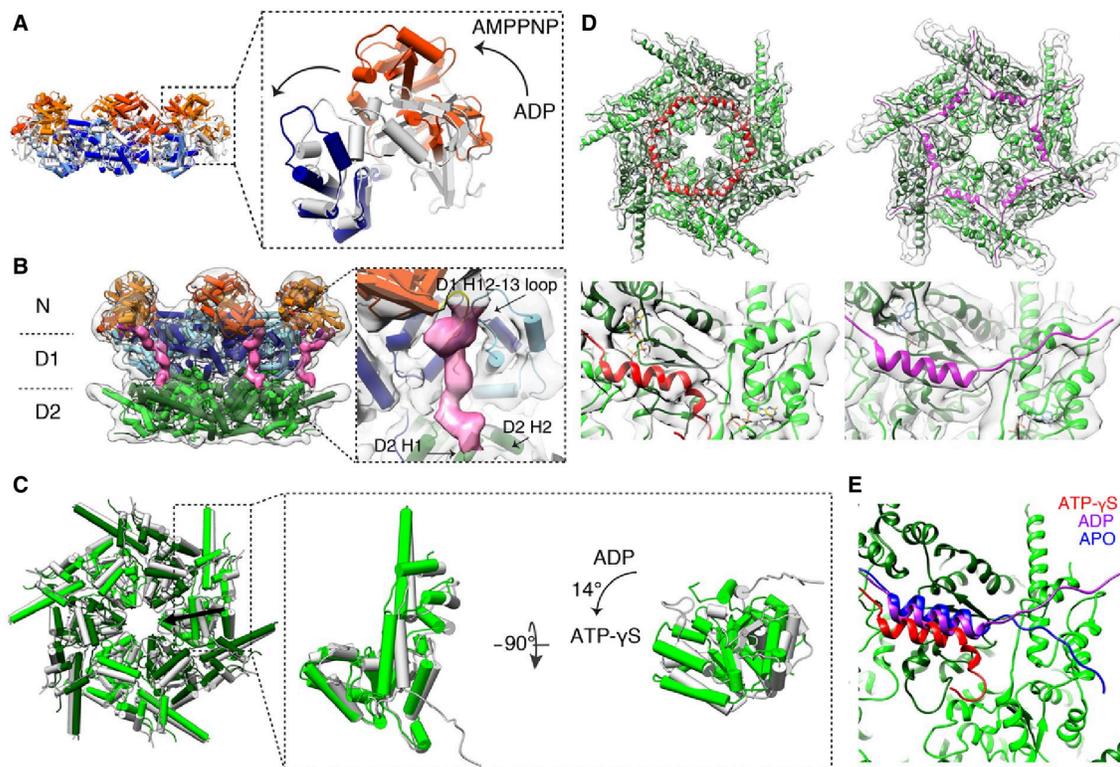


**Fig. 3.** Structural changes in the domain segments. (A) Structural changes of the AMP-PNP-model (colored) compared to the ADP-model (white). The swinging motion of the N-domain is indicated by an arrow. (B) Visualization of the most N-terminal residues in the AMP-PNP structure (pink). (C) The change of the D2 ring in the presence of ATP-γS compared to ADP can be explained by a rigid body transformation. The green model shows the ATP-γS state, whereas the gray-model displays the ADP state. An arrow for one protomer indicates the tilt axis of its 14° rotation. (D) Visualization of the C-termini in the ATP-γS (red) and ADP (purple) states. (E) Comparison of the conformations of the C-termini in different nucleotide states. In the ATP-γS-bound p97 (red) the C-terminal helix is kinked and points toward p97's central channel. In the ADP bound state (purple) the C-terminus binds to the side of the adjacent D2 small subdomain. In the apo-state (from crystal structure 5C19), the C-terminus binds to the top of the small D2 subdomain.

43

explains the density well (Fig. S6). Compared to the ADP/AMP-PNP structure, the D2 domain is rotated by approximately 14° along an axis perpendicular to the symmetry axis, narrowing the D2 pore in an aperture-like fashion.

The C-terminal helix of p97 is resolved in all D2 maps. It binds to the small subdomains of the adjacent monomer, similar as in several RecA-type ATPases [30–32]. In the ATP-γS density, the C-terminal helix is kinked and binds in the cleft between the large and small subunits of the neighboring ATPase domain, consistent with the ATP-γS-bound Δ709–728 p97 crystal structure. In addition, to the C-terminal helix, 11 residues are resolved that continue along the bottom of the small D2 subdomain of the adjacent monomer (Fig. 3D). Compared to the ATP-γS state the C-terminus binds at the opposite face of the small D2 subdomain of the neighboring protomer in the ADP/AMP-PNP structure (Fig. 3E).

## Discussion

To capture p97 in different steps of its hydrolysis cycle, we analyzed its structure in the presence of in the ATP analogs AMP-PNP and ATP-γS, as well as in the ADP ground state. High-resolution native mass spectrometry was employed to thoroughly characterize the p97 nucleotide loading state in the presence of ADP or ATP-γS, demonstrating that the particles studied by EM were predominantly loaded with either 11 ADP or 10 ATP-γS molecules. Moreover, the high mass resolving power allowed us to observe distinct p97 nucleotide loading states, indicating that nucleotide binding takes place in discrete cooperative steps of 5–6 nucleotides. In agreement with this notion, we could not detect significant deviations from the six-fold symmetry for the AAA core for any of the three nucleotides tested, which clearly distinguishes p97 from the related NSF [9]. Nevertheless, we cannot exclude the possibility of subtle deviations from six-fold symmetry at higher resolutions than those obtained in our study.

As the studied nucleotide states have been investigated previously, we can clarify the partially contradicting structural data on p97. Extensive classification revealed that it is structurally homogeneous in the presence of ADP and its six-fold symmetrical, compact structure is essentially identical to that of full-length p97 determined by X-ray crystallography [7] and SAXS [12]. A relative rotation of the D1 and D2 domains, as described in AFM experiments in the presence of ATP or ADP [13], were not seen in any of the cryo-EM maps determined here.

In the presence of AMP-PNP, the homohexamers adopt heterogeneous conformations. A variety of different hexameric combinations occur that can be attributed to the down protomer state observed in the ADP density and an up-state, in which the centers of mass of the N-domains are positioned above those of the D1-domain. This finding differs from SAXS analysis of the same sample, which suggested that the N-domains are positioned rather further below the D1 domain and move laterally [12]. This discrepancy may be due to the SAXS interpretation inherently assuming homogeneous conformations, whereas our cryo-EM single particle analysis accounts for conformational heterogeneity, which can most likely be attributed to differing occupancies of the low-affinity AMP-PNP and ADP in the particles. The finding that a truncated N-D1 construct binds exclusively AMP-PNP, whereas full-length p97 has similar amounts of AMP-PNP and ADP bound [12] supports the interpretation that N-D1 protomers in the up-state have AMP-PNP bound, whereas the down conformers probably have prebound ADP or no nucleotide bound.

The nonsymmetrized and symmetrized all-up reconstructions are clearly different from the crystal structure of full-length p97 bound to AMP-PNP [7]: the compact X-ray crystallographic conformation is most likely induced by crystal contacts, as also anticipated by the authors [7]. However, the all-up N-D1 conformation in the AMP-PNP map is essentially identical to that in the crystal structures of N-D1 fragments with R155H or A232E mutations bound to ATP-γS [10,11], indicating that the 'up' N-D1 configuration is a generic conformation of p97 and not specific for the IBMPFD mutants. Thus, the IBMPFD mutations rather change the equilibrium of the naturally occurring p97 conformations, in line with previous hypothesis [10,11]. A notable feature of the AMP-PNP structure is the binding of the p97 N-termini to the D2 domains, which apparently stabilizes this configuration.

In the ATP-γS dataset, the down-conformer of the N-domain was not detected, supporting the notion that prebound ADP induces this conformer in the presence of the lower affinity AMP-PNP. Due to high variability, the N-domain is poorly resolved, but it clearly preferably adopts an up-conformation. Consistent with this finding, the D1 domain adopts the same conformation as in the AMP-PNP density. A possible reason for the more variable N-domain positioning is the different conformation of D2, which correlates with the absence of the p97 N-terminus at the D2 domain.

44
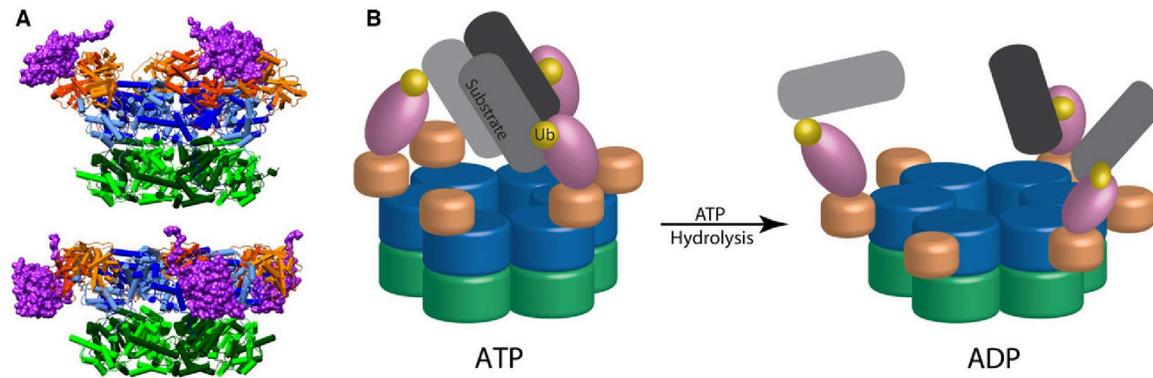
**Fig. 4.** Model of p97-mediated segregation. (A) The crystal structure of the C-terminal UBX domain of p47 (PDB: 1s3s, purple surface representation) is superimposed onto the p97 AMP-PNP and ADP models. (B) The cartoon indicates the possible positioning of a recruited substrate (gray), such as trimeric SNARE complexes for p47, in the ATP and ADP states. Via their ubiquitin tag (yellow), the substrates bind to the N-terminal domain of p47, which would position them near the pseudosymmetry axis of the D1-domain for the ATP state. The swinging motion of the N-domain upon hydrolysis may then segregate the SNARE monomers from each other and enable further cofactor-mediated substrate processing.

Three different mechanistic models for p97's segregation activity are currently considered in the field [1]: (i) threading of substrates through the central pore of p97, (ii) substrates enter and exit the pore at the D2 end, and (iii) pronounced domain movements of p97 rather than direct involvement of the pore.

The finding that the archaeal p97 homolog VAT can act as an 20S proteasome activator strongly suggests that VAT can thread substrates akin to ATPases of other ATP-dependent proteases [33,34]. In contrast to its archaeal homologs, p97 lacks critical aromatic residues in the D1 pore loop that are required for unfoldase activity in ATP-dependent proteases [35,36]. Contrary to early low-resolution cryo-EM reconstructions [14], the subnanometer p97 structures presented here indicate that the D1 pore remains too narrow to accommodate a peptide during the hydrolysis cycle. Thus, our structures do not corroborate the threading model (i).

The structural differences of the D2 domain observed in the crystal structures of p97 bound to ADP, AMP-PNP, and ATP-$\gamma$S support the notion that substrates are processed in the interior of the D2 ring [8,37]. Our cryo-EM structures confirm that these conformational changes occur in solution. Hence, our structural data are consistent with substrate processing in the D2 pore (ii).

Our structural data also agree with model (iii): the repositioning of the N-domain as a result of nucleotide hydrolysis may provide the necessary force to disassemble targeted complexes. To illustrate a possible segregation mechanism, we superposed the cocrystal structure of the p97 N-domain with the

C-terminal ubiquitin regulatory X domain (UBX) of the substrate-recruiting cofactor p47 [38] onto the atomic ATP-$\gamma$S and ADP models (Fig. 4A). These superpositions suggest that the substrate recruited by the N-terminal Ubiquitin-Associated domain (UBA) of p47 would be positioned near the D1 cylinder end in the ATP state, which would be analogous to the NSF/SNAP/SNARE supercomplex [9]. The ~ 90° rotation and concomitant ~ 12 Å translation induces a strongly leveraged repositioning of the substrate (Fig. 4B). Thus, the nucleotide-dependent conformational changes observed in this study may explain how substrates are segregated from their binding partners.

## Acknowledgements

## Author contributions

J.S. carried out cryo-EM experiments; J.S., F.B., and F.F. processed and analyzed the EM data; P.L carried

45

out the native MS experiments; all authors interpreted the results; J.S. and F.F. wrote the paper.

## References

1 Buchberger A (2013) Roles of cdc48 in regulated protein degradation in yeast. *Subcell Biochem* **66**, 195–222.

2 Braun S, Matuschewski K, Rape M, Thoms S and Jentsch S (2002) Role of the ubiquitin-selective CDC48 (UFD1/NPL4)chaperone (segregase) in ERAD of OLE1 and other substrates. *EMBO J* **21**, 615–621.

3 Jentsch S and Rumpf S (2007) Cdc48 (p97): a "molecular gearbox" in the ubiquitin pathway? *Trends Biochem Sci* **32**, 6–11.

4 Forster F, Schuller JM, Unverdorben P and Aufderheide A (2014) Emerging mechanistic insights into AAA complexes regulating proteasomal degradation. *Biomolecules* **4**, 774–794.

5 Watts GD, Wymer J, Kovach MJ, Mehta SG, Mumm S, Darvish D, Pestronk A, Whyte MP and Kimonis VE (2004) Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet* **36**, 377–381.

6 DeLaBarre B and Brunger AT (2003) Complete structure of p97/valosin-containing protein reveals communication between nucleotide domains. *Nat Struct Biol* **10**, 856–863.

7 Davies JM, Brunger AT and Weis WI (2008) Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. *Structure* **16**, 715–726.

8 Hanzelmann P and Schindelin H (2016) Structural basis of ATP hydrolysis and intersubunit signaling in the AAA+ ATPase p97. *Structure* **24**, 127–139.

9 Zhao M, Wu S, Zhou Q, Vivona S, Cipriano DJ, Cheng Y and Brunger AT (2015) Mechanistic insights into the recycling machine of the SNARE complex. *Nature* **518**, 61–67.

10 Tang WK, Li D, Li CC, Esser L, Dai R, Guo L and Xia D (2010) A novel ATP-dependent conformation in p97 N-D1 fragment revealed by crystal structures of disease-related mutants. *EMBO J* **29**, 2217–2229.

11 Tang WK and Xia D (2013) Altered intersubunit communication is the molecular basis for functional defects of pathogenic p97 mutants. *J Biol Chem* **288**, 36624–36635.

12 Davies JM, Tsuruta H, May AP and Weis WI (2005) Conformational changes of p97 during nucleotide hydrolysis determined by small-angle X-Ray scattering. *Structure* **13**, 183–195.

13 Noi K, Yamamoto D, Nishikori S, Arita-Morioka K, Kato T, Ando T and Ogura T (2013) High-speed atomic force microscopic observation of ATP-dependent rotation of the AAA+ chaperone p97. *Structure* **21**, 1992–2002.

14 Rouiller I, DeLaBarre B, May AP, Weis WI, Brunger AT, Milligan RA and Wilson-Kubalek EM (2002) Conformational changes of the multifunction p97 AAA ATPase during its ATPase cycle. *Nat Struct Biol* **9**, 950–957.

15 Yeung HO, Forster A, Bebeacua C, Niwa H, Ewens C, McKeown C, Zhang X and Freemont PS (2014) Inter-ring rotations of AAA ATPase p97 revealed by electron cryomicroscopy. *Open Biol* **4**, 130142.

16 Bai XC, Fernandez IS, McMullan G and Scheres SH (2013) Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2**, e00461.

17 Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA and Cheng Y (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* **10**, 584–590.

18 Rose RJ, Damoc E, Denisov E, Makarov A and Heck AJ (2012) High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat Methods* **9**, 1084–1086.

19 Korinek A, Beck F, Baumeister W, Nickell S and Plitzko JM (2011) Computer controlled cryo-electron microscopy–TOM(2) a software package for high-throughput applications. *J Struct Biol* **175**, 394–405.

20 Hrabe T, Beck F and Nickell S (2012) Automated particle picking based on correlation peak shape analysis and iterative classification. *Int J Med Biol Sci* **6**, 1–7.

21 Mindell JA and Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**, 334–347.

22 Nickell S, Förster F, Linaroudis A, Net WD, Beck F, Hegerl R, Baumeister W and Plitzko JM (2005) TOM software toolbox: acquisition and analysis for electron tomography. *J Struct Biol* **149**, 227–234.

23 Cardone G, Heymann JB and Steven AC (2013) One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions. *J Struct Biol* **184**, 226–236.

24 Goddard TD, Huang CC and Ferrin TE (2007) Visualizing density maps with UCSF chimera. *J Struct Biol* **157**, 281–287.

25 Zalk R and Shoshan-Barmatz V (2003) ATP-binding sites in brain p97/VCP (valosin-containing protein), a multifunctional AAA ATPase. *Biochem J* **374**, 473–480.

26 Briggs LC, Baldwin GS, Miyata N, Kondo H, Zhang X and Freemont PS (2008) Analysis of nucleotide binding to P97 reveals the properties of a tandem AAA hexameric ATPase. *J Biol Chem* **283**, 13745–13752.

27 Snijder J and Heck AJ (2014) Analytical approaches for size and mass analysis of large protein assemblies. *Ann Rev Anal Chem* **7**, 43–64.

46

28 Heck AJ (2008) Native mass spectrometry: a bridge between interactomics and structural biology. *Nat Methods* **5**, 927–933.

29 Fernandez-Saiz V and Buchberger A (2010) Imbalances in p97 co-factor interactions in human proteinopathy. *EMBO Rep* **11**, 479–485.

30 Whelan F, Stead JA, Shkumatov AV, Svergun DI, Sanders CM and Antson AA (2012) A flexible brace maintains the assembly of a hexameric replicative helicase during DNA unwinding. *Nucleic Acids Res* **40**, 2271–2283.

31 Rzechorzek NJ, Blackwood JK, Bray SM, Maman JD, Pellegrini L and Robinson NP (2014) Structure of the hexameric HerA ATPase reveals a mechanism of translocation-coupled DNA-end processing in archaea. *Nat Commun* **5**, 5506.

32 Byrne RT, Schuller JM, Unverdorben P, Forster F and Hopfner KP (2014) Molecular architecture of the HerA-NurA DNA double-strand break resection complex. *FEBS Lett* **588**, 4637–4644.

33 Barthelme D and Sauer RT (2012) Identification of the Cdc48*20S proteasome as an ancient AAA+ proteolytic machine. *Science* **337**, 843–846.

34 Forouzan D, Ammelburg M, Hobel CF, Stroh LJ, Sessler N, Martin J and Lupas AN (2012) The archaeal proteasome is regulated by a network of AAA ATPases. *J Biol Chem* **287**, 39254–39262.

35 Rothballer A, Tzvetkov N and Zwickl P (2007) Mutations in p97/VCP induce unfolding activity. *FEBS Lett* **581**, 1197–1201.

36 Barthelme D, Chen JZ, Grabenstatter J, Baker TA and Sauer RT (2014) Architecture and assembly of the archaeal Cdc48*20S proteasome. *Proc Natl Acad Sci USA* **111**, 1687–1694.

37 DeLaBarre B, Christianson JC, Kopito RR and Brunger AT (2006) Central pore residues mediate the p97/VCP activity required for ERAD. *Mol Cell* **22**, 451–462.

38 Dreveny I, Kondo H, Uchiyama K, Shaw A, Zhang X and Freemont PS (2004) Structural basis of the interaction between the AAA ATPase p97/VCP and its adaptor protein p47. *EMBO J* **23**, 1030–1039.

## Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

**Fig. S1.** MS analysis of p97 after recombinant expression and purification in nucleotide-free buffer.
**Fig. S2.** Hierarchical classification focused on the N-domains of protomers.
**Fig. S3.** Non-symmetrized ADP, AMP-PNP and ATP-γSdensities.
**Fig. S4.** Rotational symmetry analysis of the reconstructions.
**Fig. S5.** Resolution assessment of reconstructions.
**Fig. S6.** Domain specific FCR analysis of fitted atomic models, related to Figure 3.

# Supplementary Information

# Nucleotide-dependent conformational changes of the AAA+ ATPase p97 revisited

Jan M. Schuller[1], Florian Beck[1], Philip Lössl[2], Albert J.R. Heck[2], Friedrich Förster[1]*

[1] Max-Planck Institute of Biochemistry, Department of Molecular Structural Biology, D-82152 Martinsried, Germany

[2] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Centre for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH, Utrecht, The Netherlands

* Corresponding author.

Friedrich Förster: Fax: +49 89 8578 2641, E-mail address: foerster@biochem.mpg.de
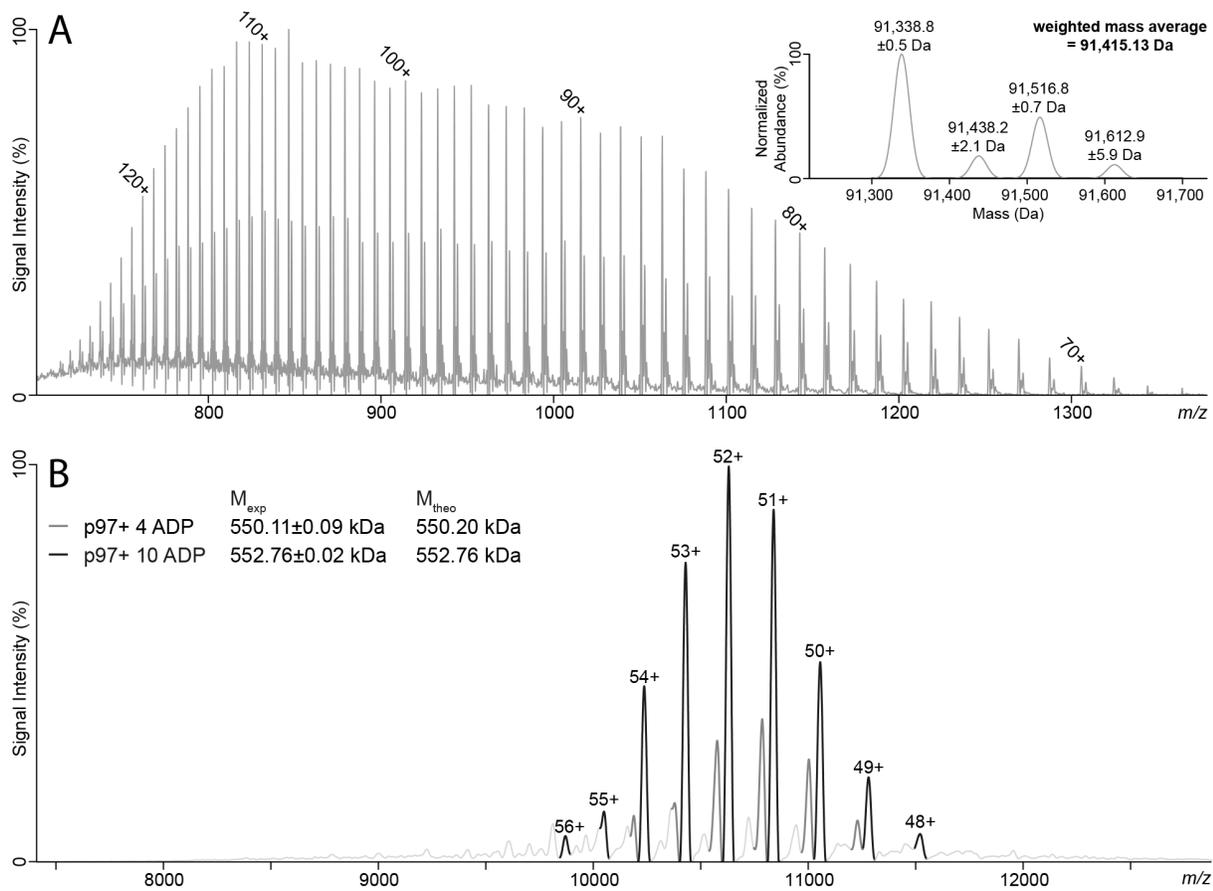
**Figure S1. MS analysis of p97 after recombinant expression and purification in nucleotide-free buffer.** (A) Mass spectrum of denatured p97. The charge states of the p97 monomer ions are shown above the peaks. Inset: mass spectrum after deconvolution to charge state 0. The molecular weights of the detected species and the weighted mass average deduced from these weights are listed. Based on this weighted mass average, the molecular weight of the apo-p97 hexamer was calculated to be 548490.8 Da. (B) Native mass spectrum of p97. The theoretical masses ($M_{theo}$) were calculated using the mass of apo -97 and then compared to the experimentally determined masses ($M_{exp}$). The charge states of the detected species are indicated.
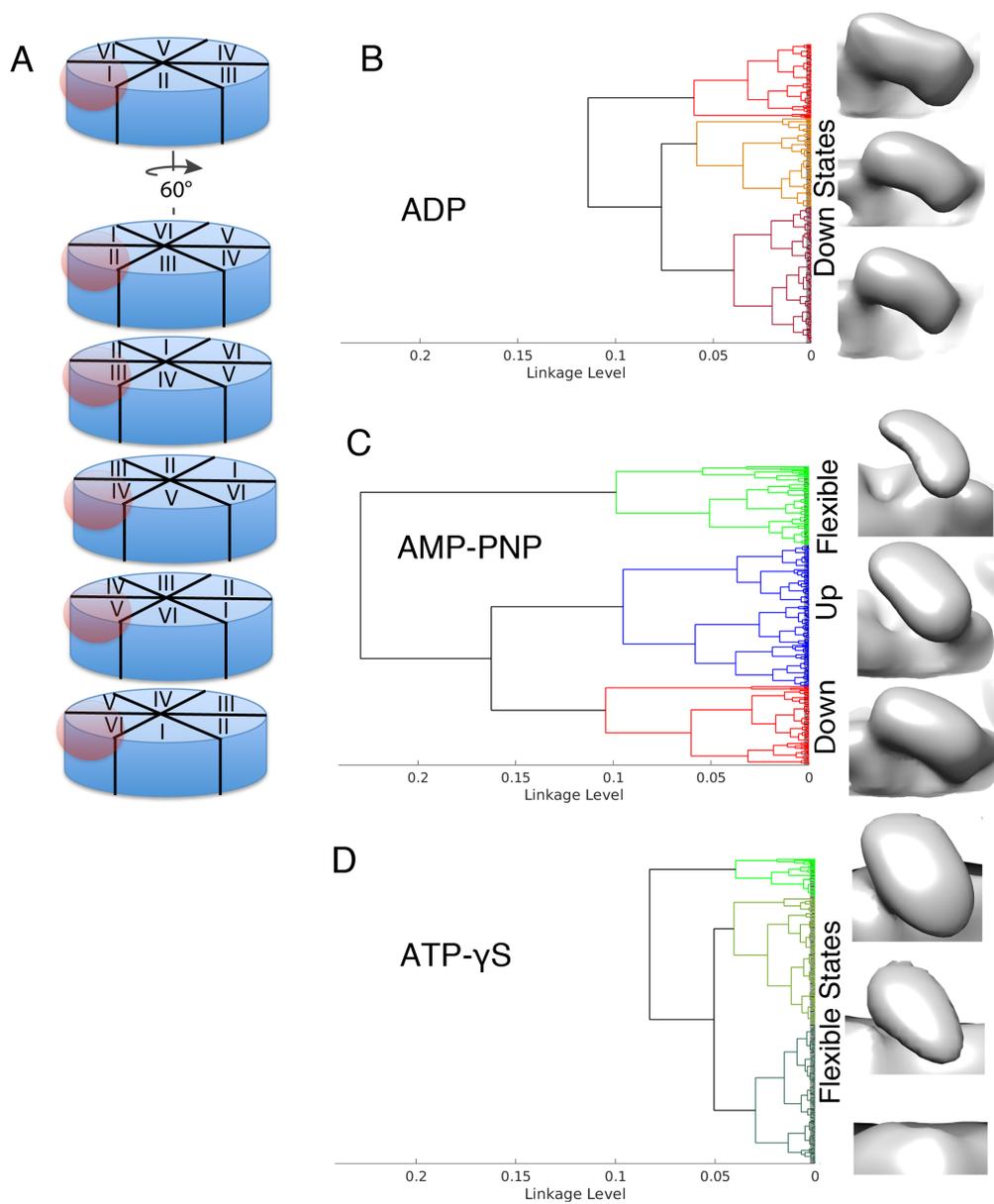
**Figure S2. Hierarchical classification focused on the N-domains of protomers.** (A) Processing scheme for N-domain classification. (B) Separation of the hierarchical tree into three bins reveals no up state protomer in p97-ADP. (C) Separation of the p97-AMP-PNP the hierarchical tree into three bins reveals the up (blue) and down (red) states and a more flexible state (green). (D) Separation of the p97 ATP-γS hierarchical tree into three bins reveals a sparsely populated up-state class and two more flexible states without a defined N-domain.
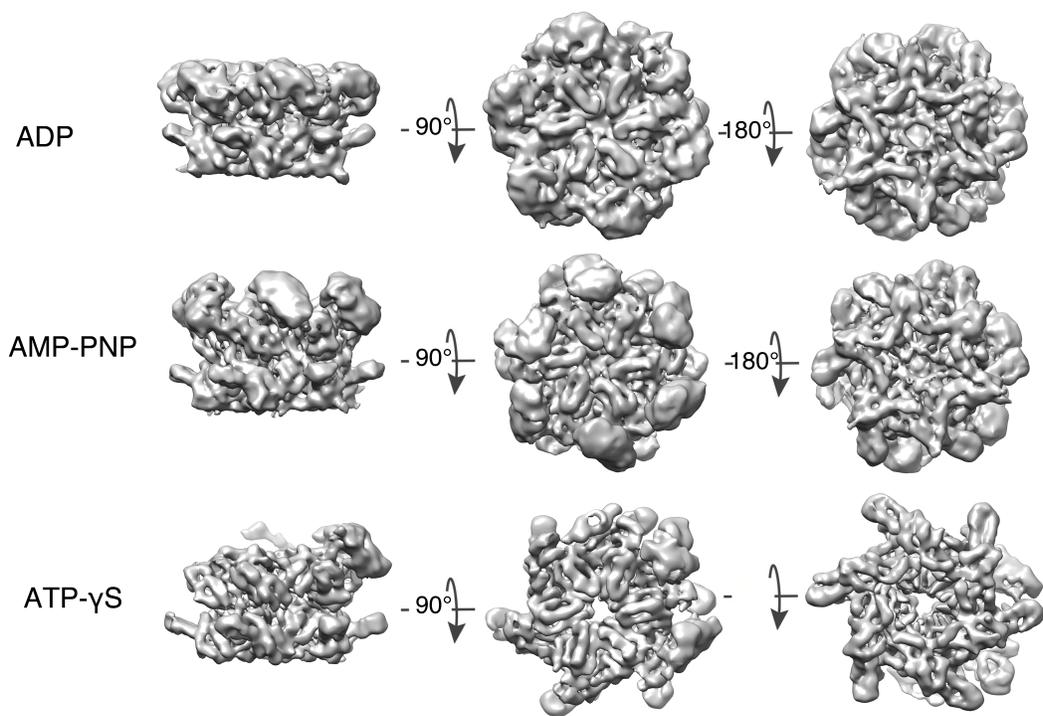
**Figure S3**. **Non-symmetrized ADP, AMP-PNP and ATP-γS densities.** The densities are filtered to their local resolutions.
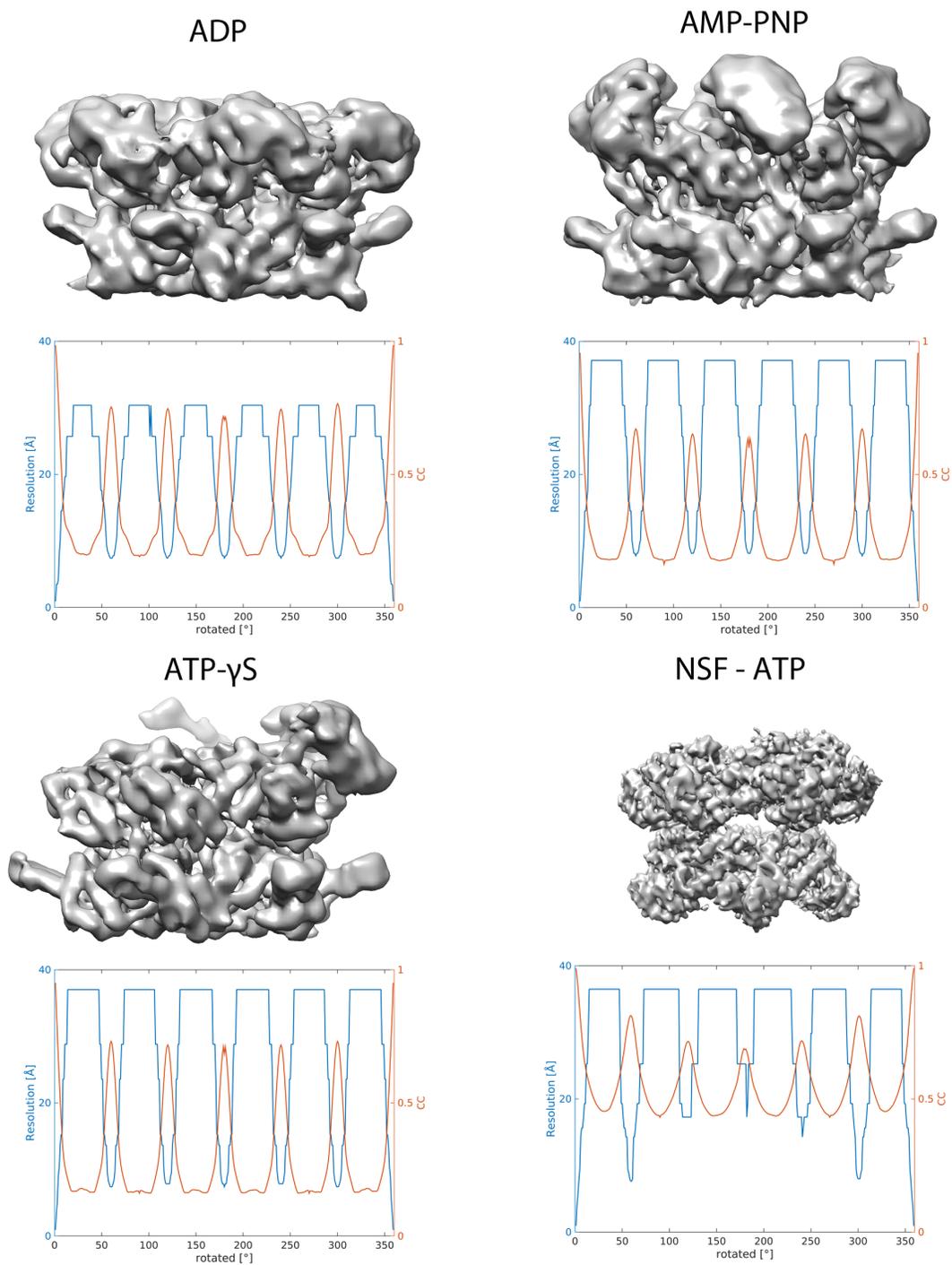
**Figure S4**. **Rotational symmetry analysis of the reconstructions.** For the analysis, the N-domains were masked out. The rotational autocorrelation functions (red) and corresponding resolutions (blue, criterion: FSC=0.143) were computed for the different densities of p97 and NSF (bottom-right). The NSF density of the ATP-state (electron microscopy database code 6204) was assessed using the same protocol.
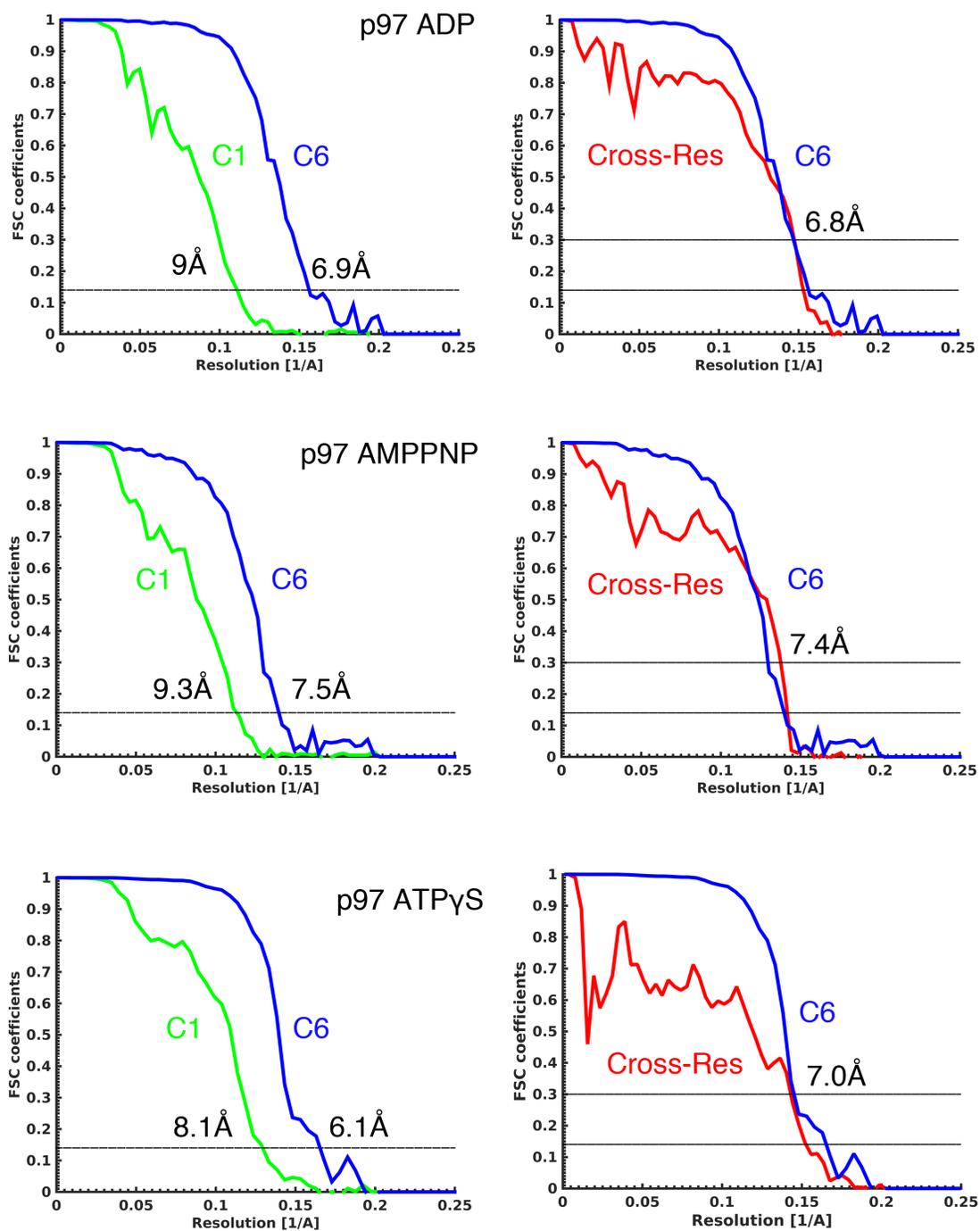
**Figure S5**. **Resolution assessment of reconstructions.** For each of the structures the Fourier Shell Correlations (FSCs) of two densities, each independently refined from 50% of the data ('gold standard FSC'), were determined (left-panel) as well as the Fourier Cross Resolutions (FCRs) of the densities with the corresponding atomic models (right panel).
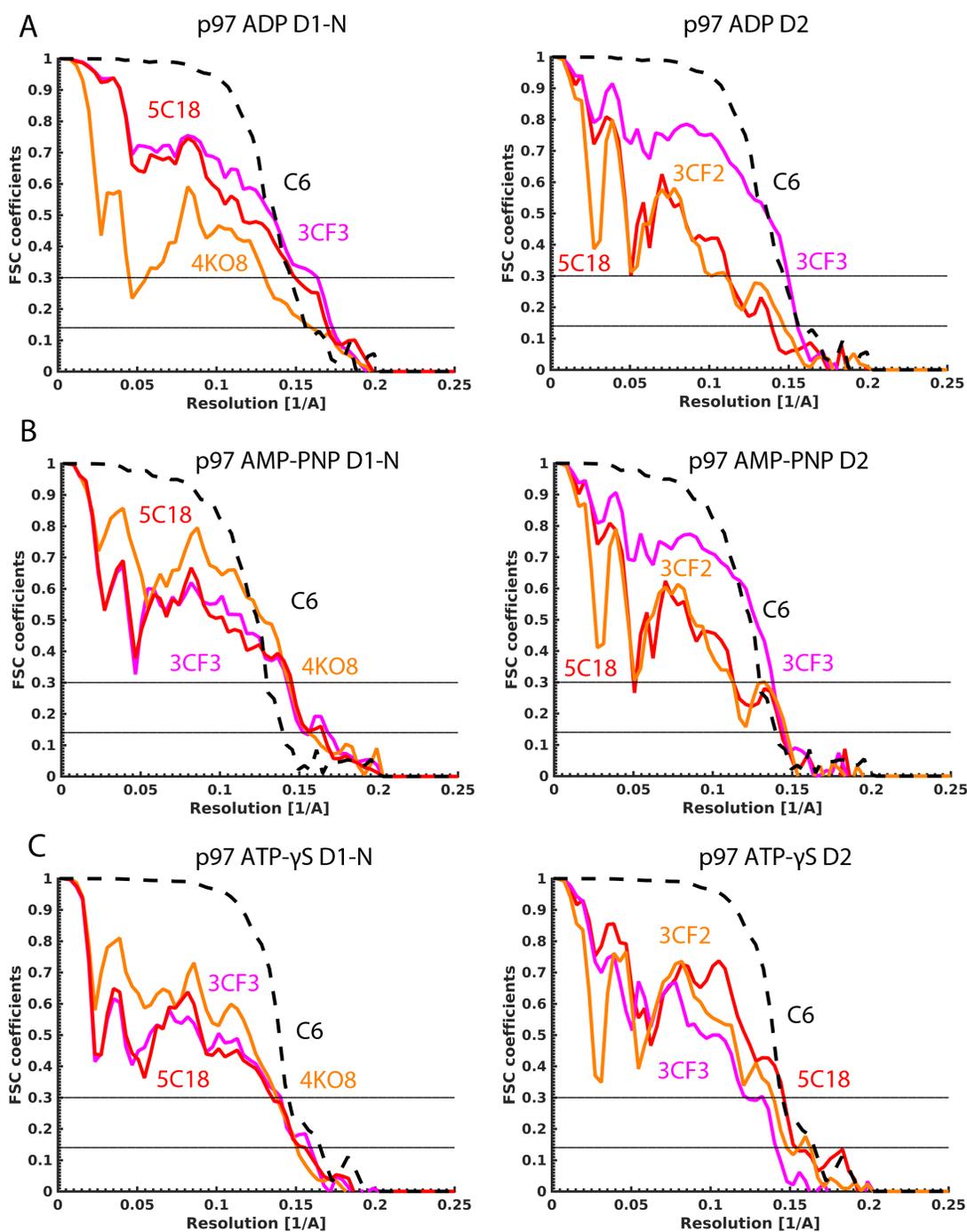
**Figure S6. Domain specific FCR analysis of fitted atomic models, related to Figure 3.** Left: The FCR of the ADP crystal structure (3CF3/purple) and the IBMPFD mutant structure (4KO8/orange) and the ATP- γS bound Δ709-728 p97 (5C18/red) was computed for each D1-N segment of the structures. Right: The FCR of the ADP crystal structure (3CF3/purple) and the AMP-PNP full-length structure (3CF2/orange) and the ATP- γS bound Δ709-728 p97 (5C18/red) was computed for each D2 segment of the structures. (A) ADP structure. (B) AMP-PNP structure. (C) ATP- γS structure.

# Chapter III: Discussion

## 3.1 Common themes in the ASCE ATPase superfamily

The ASCE superfamily of proteins represents the largest and most diverse clade of ring-shaped P-loop NTPases (Lyubimov et al., 2011). The AAA+ proteins and RecA-type translocases have a general, common architecture. They are characterized by an N-terminal domain, one or two copies of the P-loop ATPase, with or without an insertion domain between the central beta-strands, and a small, α-helical subdomain, which folds back on the P-loops ATPase and is important for its chemo-mechanical force-transduction. The N-terminal domains are more diverse.

For p97 we showed that the up-and-down swinging motion of the N-domain, driven by nucleotide hydrolysis in the D1 domain, could provide the necessary force to segregate proteins from their environment. The N-domains act together with substrate recruiting co-factors, for which the N-terminal domain forms the binding platform (Forster et al., 2014). For the HerA-NurA complex the N-terminal HAS-domain of HerA forms the binding platform for NurA. This binding pattern has also been observed for the SNAP protein in case of the NSF AAA+ ATPase (Zhao et al., 2015), the Pex5 protein in case of the Pex1/6 complex (Blok et al., 2015) and degron-containing peptides in case of the ClpXP machinery (Olivares et al., 2016). Thus, the binding of adaptors or substrates to the N-terminal domains is a common theme for all ring forming ATPases as exemplified by the two structures determined in this thesis.

## 3.2 HerA-NurA – a RecA-like DNA motor

The molecular architecture of the HerA-NurA complex solved by cryo-EM single particle analysis allowed us to propose a functional model of how the translocation-coupled DNA-end processing of this integrated helicase-nuclease complex works (Byrne et al., 2014). After the acceptance of the manuscript the crystal-structure of the full-length HerA helicases bound to AMP-PNP was solved (Rzechorzek et al., 2014). This structure allows us to critically assess the quality of our atomic model. Also by using information from both

sources, we can now generate a pseudo-atomic model of the entire HerA-NurA resection complex, permitting a more in depth analysis of the biological mechanisms.

## 3.2.1 The HerA structure – comparison between atomic model based on related structures and medium resolution EM map

The HerA helicase from the thermophilic archaeon *Sulfolobus solfataricus* was crystalized using a total of 10 surface entropy reducing mutants at a resolution of 2.85 Å, with two protomers per asymmetric unit bound to the non-hydrolyzable ATP analogue, AMP-PNP (Rzechorzek et al., 2014). The structure of the hexameric complex has a compact architecture with a central channel of approximately 25 Å in diameter. All three domains, the N-terminal β-barrel domain, the central RecA-like catalytic core and the helical insertion domain are resolved. A C-terminal extension was found to act as a brace between adjacent protomers. The C-terminal residue D471 was also found to directly interact with the nucleotide-binding pocket and the arginine-finger (R142) of the neighboring protomer (Rzechorzek et al., 2014). Mutation of this residue shows a decoupling of the ATPase activity from the nucleotide binding and thus the c-terminal brace seems to be of importance for directional DNA translocation.



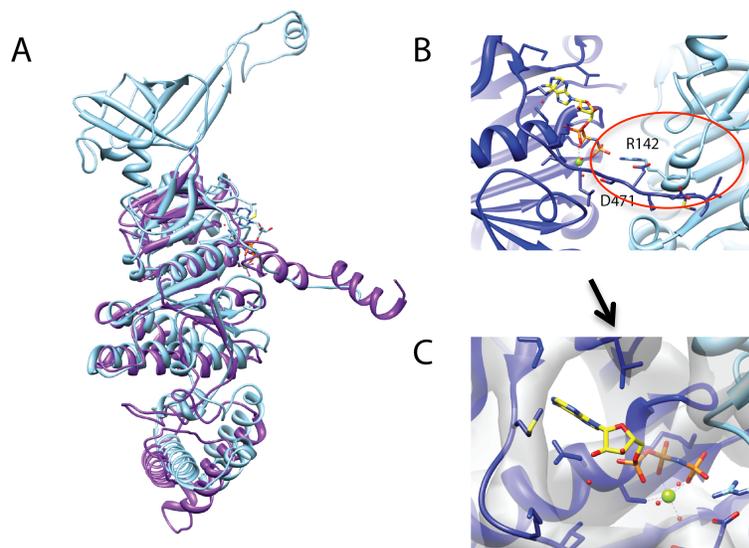Figure 14: Comparison of different structural models. (A) Superposition of the different models. Light blue: HerA-crystal structure (PDB: 4d2i), purple our homology model from Chapter 2.1. (B) close-up on the C-terminal brace, highlighted by a red circle. (C) Overlay of the AMP-PNP X-ray model and the density envelope. The nucleotide is depicted in yellow – indicating no density visible for a bound nucleotide in our map.

In my work I built an atomic model for HerA based on the 7.4-Å cryo-EM map, which does not allow *de novo* chain tracing, and crystal structures of structurally related proteins. The model of the RecA-like and the all-helical insertion domain was obtained by comparative modeling based on the VirB4 Type IV secretion DNA pump ((Wallden et al., 2012), 4AG5) and subsequent refinement in the context of the EM map by MDFF (Trabuco et al., 2008). A comparison of our HerA MDFF model and the crystal structure reveals a high similarity between the two models with an overall root-mean square deviation (rmsd) of 1.1 Å (Figure 14A). The major difference between our model and the crystal structure is that, based on secondary structure prediction, the C-terminal extension was modeled as an α-helix, while it forms a coil in the crystal structure (Figure 14B). The coil is clearly a better fit to the EM map indicating that, albeit the C-terminus was positioned correctly in the EM map, the resolution was not sufficient to unambiguously detect the correct secondary structure.

Our EM map does not show density that could account for bound nucleotides (Figure 14C), whereas the crystal structure had AMPPNP bound. Thus, the negligible differences between the two states are somewhat unexpected. Most likely, similar to the case of p97 (see also section 2.2), the conformational freedom of HerA seemed to be strongly influenced by the crystal packing. Thus, it would be required to generate single particle reconstructions of the HerA-NurA complex in presence of different nucleotides in order to understand the nucleotide dependent conformational changes of HerA (analogous to p97 in the second part of this thesis).

### 3.2.2 The HerA-NurA Interface – a detailed description of the molecular basis for the rigid body rotation

Our model lacked the N-terminal HAS domain of HerA because the sequence identity (10%) to any other known atomic structure was not high enough to build an accurate model. We could however conclude that it forms a rigid structural scaffold, similar to the oligosaccharide binding domain (OB) ring of the Rpt-ATPase of the 26S proteasome (Beck et al., 2012).

Three adjacent HAS domain regions form a binding interface for one subunit of the NurA-dimer. The HerA crystal structure fitted into our EM density envelope now enables a more

detailed description of the HerA-NurA interface (Figure 15). It is mostly hydrophobic in nature, which forms the basis for the rigid body rotation of the NurA dimer by 23°. The helical extension between the fourth and fifth beta-strand of the six-stranded HAS-domain contributes a series of hydrophobic residues that form the majority of the NurA binding site. In NurA the hydrophobic binding patch is centered on residue F300 (Blackwood et al., 2012). However, these residues are only close enough for binding after the rigid body rotation of the NurA dimer (Byrne et al., 2014).
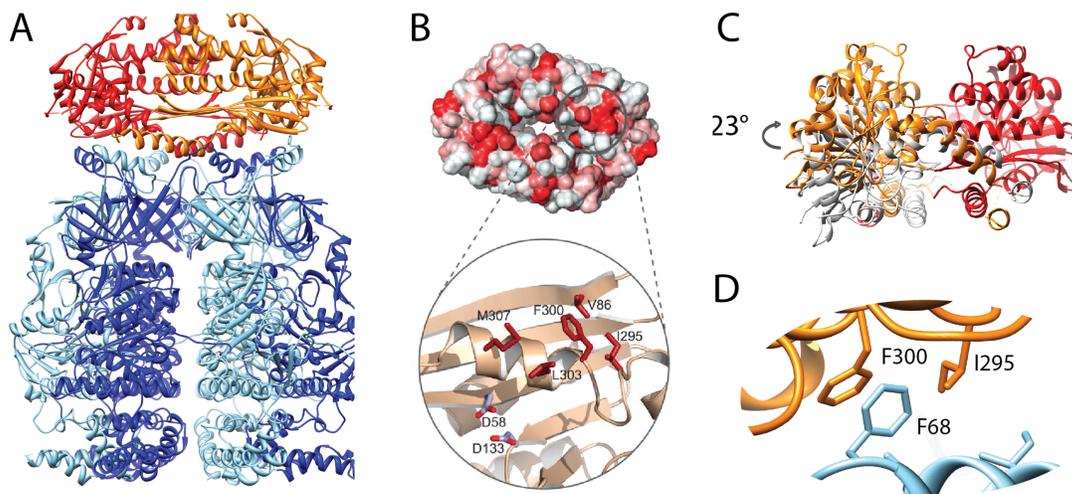


Figure 15: A detailed description of the structural interface of the HerA-NurA complex. (A) Structural model of the entire HerA-NurA complex. (B) Hydrophobic nature of the NurA surface, plotted red (most hydrophobic) to white (least hydrophobic) on the solvent-accessible surface of the protein. A hydrophobic patch is shown in close-up view (C) In the NurA-homodimer the relative orientations of the monomers differ by 23° with respect to the X-ray crystal structure (white). (D) Key residues forming the hydrophobic binding interface of the complex. Panel (B) is adapted from (Blackwood et al., 2012).

### 3.2.3 The HerA-NurA molecular machine – an integrated functional model

The apo structure of the complex allows us to postulate a mechanism for single strand resection by HerA-NurA. The HerA-helicase and the NurA-nuclease form an integrated coaxial structure that threads dsDNA substrates towards the NurA nuclease. The HerA-NurA complex preferably acts on substrates with a 5' single-strand overhang, a structure that is commonly generated by the processive nuclease activity of the Mre11-Rad50 complex after double-strand break recognition (Blackwood et al., 2013). The central cavity inside the HerA–helicase is wide enough to accommodate up to three turns of dsDNA. After

the rigid-body rotation of the NurA-dimer one central channel remains that is only wide enough to accommodate the 5' single-strand overhang, but not the dsDNA. The dsDNA is now translocated towards the NurA dimer, which would act as a molecular 'ploughshare', melting the dsDNA. The 5'-free end passes through the annulus and the unwound 3' end meets the NurA active sites where it is digested. To elucidate the exact mechanism of how the HerA-NurA complex threads DNA towards the NurA channel a cryo-EM structure of the substrate bound complex would be required.

### 3.3 p97–an ATP dependent molecular segregase

We were able to determine the structure of p97 bound to different nucleotides at resolutions between 6.1-7.4 Å, after applying the appropriate $C_6$-symmetry. By comparing our reconstructions with available crystal structures we could build pseudo-atomic models for all nucleotide states. In the following, I will critically re-evaluate the applied-six fold symmetry and compare our structures with a higher resolution cryo-EM study of p97 published after submission of our manuscript (Banerjee et al., 2016) and discuss potential functional models.

### 3.3.1 p97 six-fold symmetric structure, but asymmetric functionality

In our study we probed the nucleotide binding of p97 after incubation with different nucleotides by native mass-spectroscopy. We showed that without incubation with additional nucleotide p97 carried pre-bound ADP and that after incubation with ATP-γS, all pre-bound ADP was exchanged with ATP-γS. In all cases, the major mode of p97 was loaded with approximately 10±1 nucleotides, indicating a mechanism of cooperative nucleotide binding and exchange, in which 1-2 nucleotide binding sites are not occupied per hexamer. These numbers would exclude a strict six-fold symmetry in the presence of ATP or an ATP analogue.

It was thus unexpected to find obvious six-fold symmetry in our p97 reconstructions. However, the nucleotide binding studies and our symmetric sub-nanometer reconstructions are only at the first sight contradictory. First, the native-mass spectrometry data has a non-negligible error margin. Solely based on this method it would therefore not be possible to exculde a fully loaded hexamer with real six-fold symmetry. However, not fully saturated binding was also observed in a variety of independent studies, using

different techniques, such as isothermal titration calorimetry (ITC) measurements (Briggs et al., 2008), heat-denaturation experiments (Davies et al., 2005) and kinetic measurements, probing the cooperativity of the two ATPase rings (Chou et al., 2014). Secondly, we used a hierarchical clustering procedure to classify p97 particles based on the conformation of their N-domain in order to get a conformational homogenous dataset. Only in the AMP-PNP dataset, we were able to obtain a distribution of different conformation of N-domains, providing evidence for conformational heterogeneity of N-domains in p97 (Figure 16A). This can most likely be explained by the low affinity of AMP-PNP for p97, compared with ATP-γS (Briggs et al., 2008). In case of the ADP dataset, no up-conformation was detected and in the ATP-γS dataset no down-conformer could be visualized. Thus, these data sets seem to be structurally more homogenous. Third, we are limited by the resolution of our reconstructions. All our reconstructions reached sub-nanometer resolution without applying any symmetry, thus we were able to resolve secondary structures at the AAA core of the molecule, but we cannot distinguish the bound nucleotide. In consequence we cannot determine the nucleotide occupancy in the non-symmetrized maps. On the other hand, a local resolution map of the ADP and AMP-PNP reconstructions indicate that one of the six N-domains is lower resolved than the other five, thus hinting at structural asymmetry (Figure 16B).
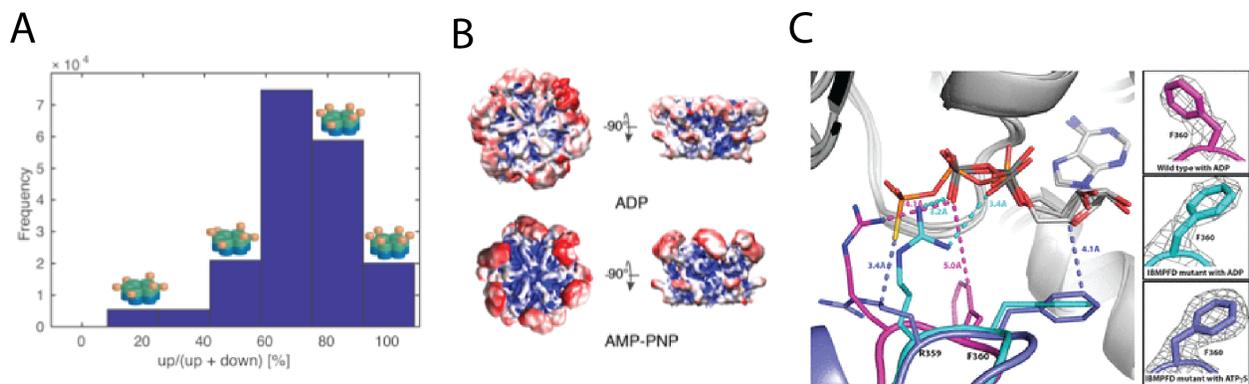


Figure 16: Structural asymmetry in the N-D1 Interface. (A) Frequencies of up- to down- ratios in the hexamers in case of the AMP-PNP dataset. A cartoon representation illustrates the possible corresponding asymmetric N-domain arrangements. (B) p97 nucleotide states colored according to their local resolution. (C) alignment of three p97 N-D1 structures (wild type with ADP bound (magenta, PDB code 1E32), R155H mutant with ADP bound (cyan), and R155H mutant with ATPγS bound (blue)), based on the superposition of bound nucleotides. The position of the residue F360 is responsible for the functional asymmetry of p97 IBMPFD mutants. Panel (C) is adapted from (Tang and Xia, 2013).

An asymmetric feature that may lead to a more loosely bound N-domain is the nucleotide binding and in direct consequence the side chain positioning, which can only be observed at higher resolution. Especially, in the D1 domain of p97 the positioning of the side-chains seems to be of greatest importance for its function. This is exemplified by high-resolution crystal structures of the IBMPFD mutants, which show that an altered conformation of residue F360 probes the adjacent nucleotide-binding pocket and thus leads to a lower affinity for ADP compared to wild-type proteins (Figure 16C) (Tang et al., 2010; Tang and Xia, 2013). Another study published after the acceptance of our manuscript showed that the N-terminal half of the D1-D2 linker is evolutionary conserved and capable of introducing asymmetry in the association of subunits (Tang and Xia, 2016). The activation of the D1 ATPase by this linker is mainly based on the side-chain conformation of the trans-acting arginine finger R359. Thus, implying that breaking the perfect six-fold symmetry by side-chain conformations is at least required for D1 AAA activity.

### 3.3.2 Comparison of our p97 structure with an external cryo-EM structure

The Subramanian group solved the structures of p97 bound to ADP, ATP-γS and of ADP bound to an allosteric inhibitor to near atomic resolutions of 2.3-3.4 Å (Banerjee et al., 2016). All these reconstructions imposed $C_6$-symmetry; asymmetric reconstructions were not provided.

The Banerjee ADP-reconstruction is essentially identical to published crystal structures (Davies et al., 2008) and to our map, albeit at higher resolution. They classified the ATP-γS dataset into three classes, from which two had the N-domains in the down conformation. The first of the Banerjee ATP-γS classes had ADP bound to both rings. This class was essentially identical to the ADP structure. The second class had ATP-γS bound to the D2 domain and ADP bound to the D1 domain. The N-domains where thus in the down-conformer and the D2 ring showed the same nucleotide-dependent conformational changes as observed in the D2 domain of the p97-crystal structure in presence of ATP-γS (Hanzelmann and Schindelin, 2016b). The third class had ATP-γS bound to both ATPase rings. This structure showed the same nucleotide dependent conformational changes as our structure. From these final refined structures, a step-wise evolution of nucleotide-driven structural changes in p97 is proposed. They conclude that the effect of ATP-γS binding is a

two-step, sequential and cooperative conformational change in the D2 and D1 ATPase rings (Banerjee et al., 2016). However, our ATP-$\gamma$S data indicates no down-conformer after incubation with ATP-$\gamma$S. Our results do therefore not fully corroborate their model. In addition, our AMP-PNP structure showed the ATP dependent up-conformer of the N-domain, while the D2 ring showed the same conformation as observed in the ADP structure. Thus, suggesting a model, in which ATP is first exchanged in the D1 domain, while the D2 domain is still occupied by ADP.

### 3.3.3 Structural studies of p97 and the NU-complex

For a final model on p97 function it is important to note that p97 in the cell always works in concert with its co-factors (Forster et al., 2014). We suggest that the repositioning of the N-domain as a result of nucleotide hydrolysis may provide the necessary force to disassemble targeted complexes by their interaction with the cofactors. The substrate recruiting cofactors all bind to two neighboring N-domains in p97 using bipartite interaction motifs (Figure 17A, B) (Hanzelmann and Schindelin, 2016a). The binding of the heterodimer NU to p97 shows a fixed 1 to 6 ratio. Why the NU:p97 stoichiometry is 1:6 and how co-factors influence the ATPase activity of p97 cannot be explained without further structural data.

In addition to the above outlined published results I also attempted to elucidate the molecular architecture of the p97-NU complex as part of this thesis. To this end we used an E305Q walker B mutation of the p97 D1 domain (p97-WB) to abolish the hydrolysis of ATP in the D1 domain. In analogy to our structures of p97 in presence of ATP-$\gamma$S, p97-WB in presence of ATP should adopt the "up" conformation. Binding of ATP in the D1 domain has been shown to increase the affinity of p97 to NU (Chia et al., 2012). The NU complex could be successfully purified after co-expression in *Escherichia Coli* (Figure 17C). After addition of purified p97, the octameric complex was further separated from the excess of NU by gel filtration yielding a complex with the expected 6:1 stoichiometry. However no additional density corresponding to NU could be observed in cryo-EM analysis and 3D reconstruction of this p97-cofactor complex. To stabilize the complex the sample was subjected to chemical cross-linking prior to cryo-EM using the soluble BS3 lysine-specific cross-linking reagent. Successful cross-linking was visualized using SDS-PAGE (Figure 17C). However, the resulting reference-free 2D classes did not indicate the binding of extra-
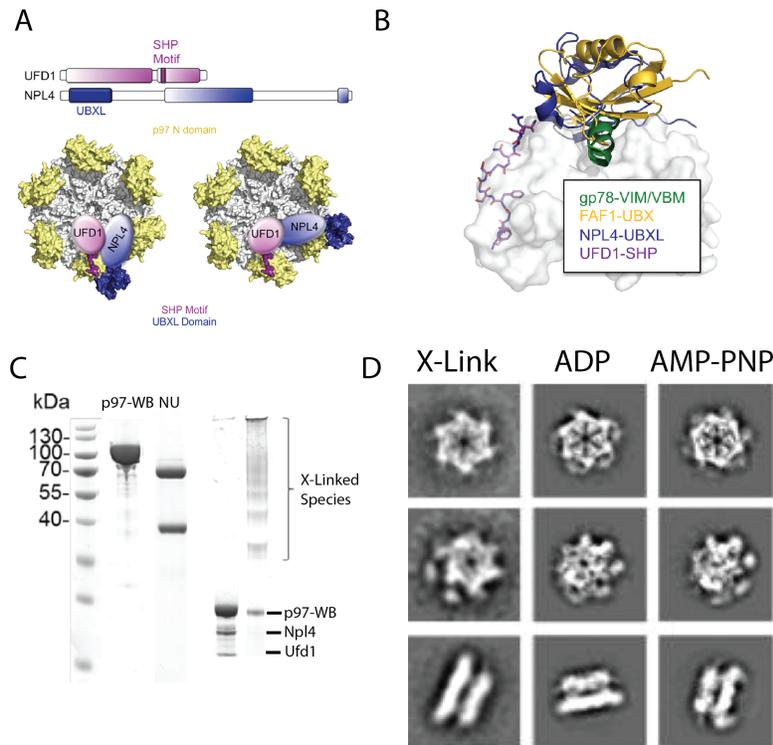
Figure 17: NU binding to the p97 N-domain. (A) Domain architecture of NU and representation of the possible bipartite interactions binding modes (B) p97N in complex with the VIM of gp78 (PDB: 3TIW, colored in green), FAF1-UBX (PDB: 3QQ8, colored in gold) and NPL4-UBXL (PDB: 2PJH, colored in blue). The UFD1-SHP1 peptide (colored according to atom type with C-atoms in purple). (C) Purification of the p97-NU complex and stabilization of the complex by cross-linking. (D) Reference-free 2D class-averages indicate no extra-density bound to the cross-linked complex. The N-domains are locked in the down-conformation as seen for the ADP bound state. Panel (A) and (B) are adapted from (Hanzelmann and Schindelin, 2016a).

density to the p97 N-domains. Moreover, the classes showed the N-domains exclusively in the down-conformation, which is likely, a consequence of the cross-linking (Figure 17D).

In summary, the NU-p97 structure could not be determined by cryo-EM. One possible reason is that cross-linking forced p97-WB into an ADP-like state, leading to complex disassembly. Another possibility is that NU was actually bound and cross-linked to p97-WB, but could not be visualized due to conformational flexibility. Future attempts could include reconstitution of the p97-NU complex in presence of a poly-ubiquitylated model substrate. Presence of such a substrate could increase the affinity of NU for p97 and limit the flexibility of the NU heterodimer, which would both make structure determination by cryo-EM easier.

# Chapter IV: Conclusion and Future Challenges

This work makes use of recent developments in cryo-EM single particle analysis to determine the high-resolution structures of two ring forming ASCE ATPases. Until direct-detector technologies and improvements in computational analysis became recently available even quaternary structure determinations for such molecules were ambiguous. One of the most prominent examples of an erroneous single particle reconstruction was the study of the Hsp104 AAA+ protein disaggregase (Wendler et al., 2007).

In this thesis I solved the structure of the DNA-end resection complex HerA-NurA and thereby demonstrate that single particle analysis can be used to place individual crystal structures in a molecular context. Furthermore, by studying the conformational cycle of p97 I showed that cryo-EM methods could provide useful insights into the functional mechanisms for dynamic protein complexes in solution, avoiding artifacts from crystal-packing forces. (chapter 2.2, (Schuller et al., 2016)). Other recent examples include, the studies on NSF (Zhao et al., 2015) and the 26S proteasome (Unverdorben et al., 2014). Taken together this demonstrates that cryo-EM single particle analysis is the method of choice to study nucleotide-dependent conformational changes of ring-forming ASCE ATPases.

The technology of single-particle cryo-EM is, however, still far from perfection and meets several challenges. Most of them I encountered during the p97 project (see section 2.2). (i) A major issue was the sample preparation, especially the reproducibility of continuous, thin ice areas on a specimen grid. Furthermore, the water-air interface, which is extremely large for thin ice, seems to cause the disassembly of protein complexes in a poorly understood manner (Glaeser et al., 2016). This might also have affected the stability of the HerA-NurA complex in thin ice areas (Chapter 2.1). As exemplified in case of the NU-p97 complex, the successful vitrification of large, transient and potentially flexible protein assemblies is still challenging. They often disassemble during the vitrification process and need to be stabilized by cross-linking (Chang et al., 2015; Plaschka et al., 2015). As cross-linkers, either

glutaraldehyde or bi-functional lysine cross-linkers are employed (Kastner et al., 2008; Plaschka et al., 2015). It is often beneficial to further separate cross-linked transient complexes by size separation techniques (Chang et al., 2015). For this type of samples the biochemical sample preparation and the computational analysis will become the main bottleneck and mandates further developments in the future. (ii) Furthermore, particles rarely show a uniform orientation distribution in the vitreous ice, again due to the air-water interface. Instead, particular views, which are required for 3D reconstruction, are often largely absent. This preferred orientation bias leads to distortions in the 3D reconstructions, similar to the missing-wedge effect in tomography (Lucic et al., 2005). In severe cases, the computational algorithms for single-particle analysis fail entirely because they inherently assume that all required views are sampled (Wendler et al., 2007). This turned out to be largest problem in the case of p97. Thus, the addition of surfactants was tested, as in some cases the addition of detergents has triggered a more uniform orientation distribution (Chowdhury et al., 2015; Lander et al., 2013; Zhao et al., 2015). However, the use of detergents requires large concentration of the protein sample, making it not suitable for many studies. In other studies, the data acquisition at areas of thick ice was reported to make the orientation distribution of the particles more uniform, but in turn limits the achievable resolution (Blok et al., 2015).

The rapid advances in cryo-EM promise an exciting future for this technique. It is not only expected that high-resolution structures will propel our understanding of macro-molecular machines, but also a resolution range that rivals X-ray crystallography could be used to study the binding of small-molecule compounds, such as co-factors or inhibitors. The wide range of potentially applicable targets is exemplified by the studies of inhibitors bound to ribosomes (Wilson et al., 2015; Wong et al., 2014), to the 20S proteasome of *Plasmodium falciparum* (Li et al., 2016), to p97 (Banerjee et al., 2016) as well as to ion channels (Yang et al., 2015; Zhu et al., 2016). This kind of structures, initially used to lock a macromolecule in a specific functional state, shows also big potential for pharmacological studies of drug-target interactions. Thus, cryo-EM studies are likely to contribute in the future to the drug-design field when the targets are not amendable to X-ray crystallography.

# Bibliography

Baek, G.H., Cheng, H., Choe, V., Bao, X., Shao, J., Luo, S., and Rao, H. (2013). Cdc48: a swiss army knife of cell biology. J Amino Acids *2013*, 183421.

Bai, X.C., Fernandez, I.S., McMullan, G., and Scheres, S.H. (2013). Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. Elife *2*, e00461.

Baker, L.A., and Rubinstein, J.L. (2010). Radiation damage in electron cryomicroscopy. Methods Enzymol *481*, 371-388.

Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S.L., Yan, Y., Green, N., Mroczkowski, B., Neitz, R.J., Wipf, P.*, et al.* (2016). 2.3 A resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. Science *351*, 871-875.

Bartesaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J.L., and Subramaniam, S. (2015). 2.2 A resolution cryo-EM structure of beta-galactosidase in complex with a cell-permeant inhibitor. Science *348*, 1147-1151.

Beck, F., Unverdorben, P., Bohn, S., Schweitzer, A., Pfeifer, G., Sakata, E., Nickell, S., Plitzko, J.M., Villa, E., Baumeister, W.*, et al.* (2012). Near-atomic resolution structural model of the yeast 26S proteasome. Proc Natl Acad Sci U S A *109*, 14870-14875.

Blackwood, J.K., Rzechorzek, N.J., Abrams, A.S., Maman, J.D., Pellegrini, L., and Robinson, N.P. (2012). Structural and functional insights into DNA-end processing by the archaeal HerA helicase-NurA nuclease complex. Nucleic Acids Res *40*, 3183-3196.

Blackwood, J.K., Rzechorzek, N.J., Bray, S.M., Maman, J.D., Pellegrini, L., and Robinson, N.P. (2013). End-resection at DNA double-strand breaks in the three domains of life. Biochem Soc Trans *41*, 314-320.

Blok, N.B., Tan, D., Wang, R.Y., Penczek, P.A., Baker, D., DiMaio, F., Rapoport, T.A., and Walz, T. (2015). Unique double-ring structure of the peroxisomal Pex1/Pex6 ATPase complex revealed by cryo-electron microscopy. Proc Natl Acad Sci U S A *112*, E4017-4025.

Bogden, C.E., Fass, D., Bergman, N., Nichols, M.D., and Berger, J.M. (1999). The structural basis for terminator recognition by the Rho transcription termination factor. Mol Cell *3*, 487-493.

Botos, I., Melnikov, E.E., Cherry, S., Khalatova, A.G., Rasulova, F.S., Tropea, J.E., Maurizi, M.R., Rotanova, T.V., Gustchina, A., and Wlodawer, A. (2004). Crystal structure of the AAA+ alpha domain of E. coli Lon protease at 1.9A resolution. J Struct Biol *146*, 113-122.

Braun, S., Matuschewski, K., Rape, M., Thoms, S., and Jentsch, S. (2002). Role of the ubiquitin-selective CDC48(UFD1/NPL4 )chaperone (segregase) in ERAD of OLE1 and other substrates. Embo J *21*, 615-621.

Briggs, L.C., Baldwin, G.S., Miyata, N., Kondo, H., Zhang, X., and Freemont, P.S. (2008). Analysis of nucleotide binding to P97 reveals the properties of a tandem AAA hexameric ATPase. J Biol Chem *283*, 13745-13752.

Brilot, A.F., Chen, J.Z., Cheng, A., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B., Henderson, R., and Grigorieff, N. (2012). Beam-induced motion of vitrified specimen on holey carbon film. J Struct Biol *177*, 630-637.

Buchberger, A., Schindelin, H., and Hanzelmann, P. (2015). Control of p97 function by cofactor binding. FEBS Lett *589*, 2578-2589.

Byrne, R.T., Schuller, J.M., Unverdorben, P., Forster, F., and Hopfner, K.P. (2014). Molecular architecture of the HerA-NurA DNA double-strand break resection complex. FEBS Lett *588*, 4637-4644.

Chae, J., Kim, Y.C., and Cho, Y. (2012). Crystal structure of the NurA-dAMP-Mn2+ complex. Nucleic Acids Res *40*, 2258-2270.

Chang, L., Zhang, Z., Yang, J., McLaughlin, S.H., and Barford, D. (2015). Atomic structure of the APC/C and its mechanism of protein ubiquitination. Nature *522*, 450-454.

Chedin, F., Handa, N., Dillingham, M.S., and Kowalczykowski, S.C. (2006). The AddAB helicase/nuclease forms a stable complex with its cognate chi sequence during translocation. J Biol Chem *281*, 18610-18617.

Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H., and Henderson, R. (2013). High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy *135*, 24-35.

Chen, X., Cui, D., Papusha, A., Zhang, X., Chu, C.D., Tang, J., Chen, K., Pan, X., and Ira, G. (2012). The Fun30 nucleosome remodeller promotes resection of DNA double-strand break ends. Nature *489*, 576-580.

Chia, W.S., Chia, D.X., Rao, F., Bar Nun, S., and Geifman Shochat, S. (2012). ATP binding to p97/VCP D1 domain regulates selective recruitment of adaptors to its proximal N-domain. PLoS One *7*, e50490.

Chou, T.F., Bulfer, S.L., Weihl, C.C., Li, K., Lis, L.G., Walters, M.A., Schoenen, F.J., Lin, H.J., Deshaies, R.J., and Arkin, M.R. (2014). Specific inhibition of p97/VCP ATPase and kinetic analysis demonstrate interaction between D1 and D2 ATPase domains. J Mol Biol *426*, 2886-2899.

Chowdhury, S., Ketcham, S.A., Schroer, T.A., and Lander, G.C. (2015). Structural organization of the dynein-dynactin complex bound to microtubules. Nat Struct Mol Biol *22*, 345-347.

Churchill, J.J., and Kowalczykowski, S.C. (2000). Identification of the RecA protein-loading domain of RecBCD enzyme. J Mol Biol *297*, 537-542.

Constantinesco, F., Forterre, P., and Elie, C. (2002). NurA, a novel 5'-3' nuclease gene linked to rad50 and mre11 homologs of thermophilic Archaea. EMBO Rep *3*, 537-542.

Constantinesco, F., Forterre, P., Koonin, E.V., Aravind, L., and Elie, C. (2004). A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea. Nucleic Acids Res *32*, 1439-1447.

David B. Williams, C.B.C. (2009). Transmission Electron Microscopy, Vol 978-0-387-76500-6

Davies, J.M., Brunger, A.T., and Weis, W.I. (2008). Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. Structure *16*, 715-726.

Davies, J.M., Tsuruta, H., May, A.P., and Weis, W.I. (2005). Conformational changes of p97 during nucleotide hydrolysis determined by small-angle X-Ray scattering. Structure *13*, 183-195.

De Rosier, D.J., and Klug, A. (1968). Reconstruction of three dimensional structures from electron micrographs. Nature *217*, 130-134.

Defenouillere, Q., Yao, Y., Mouaikel, J., Namane, A., Galopier, A., Decourty, L., Doyen, A., Malabat, C., Saveanu, C., Jacquier, A*., et al.* (2013). Cdc48-associated complex bound to 60S

particles is required for the clearance of aberrant translation products. Proc Natl Acad Sci U S A *110*, 5046-5051.

Dillingham, M.S., and Kowalczykowski, S.C. (2008). RecBCD enzyme and the repair of double-stranded DNA breaks. Microbiol Mol Biol Rev *72*, 642-671, Table of Contents.

Dubochet, J. (2016). A Reminiscence about Early Times of Vitreous Water in Electron Cryomicroscopy. Biophys J *110*, 756-757.

Dubochet, J., Lepault, J., Freeman, R., Berriman, J.A., and Homo, J.C. (1982). Electron-Microscopy of Frozen Water and Aqueous-Solutions. J Microsc-Oxford *128*, 219-237.

Enemark, E.J., and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. Nature *442*, 270-275.

Enenkel, C., Lehmann, A., and Kloetzel, P.M. (1998). Subcellular distribution of proteasomes implicates a major location of protein degradation in the nuclear envelope-ER network in yeast. Embo J *17*, 6144-6154.

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu Rev Biophys Biomol Struct *35*, 93-114.

Forster, F., Schuller, J.M., Unverdorben, P., and Aufderheide, A. (2014). Emerging mechanistic insights into AAA complexes regulating proteasomal degradation. Biomolecules *4*, 774-794.

Frank, J. (2002). Single-particle imaging of macromolecules by cryo-electron microscopy. Annu Rev Biophys Biomol Struct *31*, 303-319.

Gai, D., Zhao, R., Li, D., Finkielstein, C.V., and Chen, X.S. (2004). Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. Cell *119*, 47-60.

Gerhold, C.B., Hauer, M.H., and Gasser, S.M. (2015). INO80-C and SWR-C: guardians of the genome. J Mol Biol *427*, 637-651.

Glaeser, R.M. (1971). Limitations to significant information in biological electron microscopy as a result of radiation damage. J Ultrastruct Res *36*, 466-482.

Glaeser, R.M., Han, B.G., Csencsits, R., Killilea, A., Pulk, A., and Cate, J.H. (2016). Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens. Biophys J *110*, 749-755.

Guenther, B., Onrust, R., Sali, A., O'Donnell, M., and Kuriyan, J. (1997). Crystal structure of the delta' subunit of the clamp-loader complex of E. coli DNA polymerase III. Cell *91*, 335-345.

Hanzelmann, P., Buchberger, A., and Schindelin, H. (2011). Hierarchical binding of cofactors to the AAA ATPase p97. Structure *19*, 833-843.

Hanzelmann, P., and Schindelin, H. (2016a). Characterization of an Additional Binding Surface on the p97 N-Terminal Domain Involved in Bipartite Cofactor Interactions. Structure *24*, 140-147.

Hanzelmann, P., and Schindelin, H. (2016b). Structural Basis of ATP Hydrolysis and Intersubunit Signaling in the AAA+ ATPase p97. Structure *24*, 127-139.

Harper, J.W., and Elledge, S.J. (2007). The DNA damage response: ten years after. Mol Cell *28*, 739-745.

Heyer, W.D. (2004). Recombination: Holliday junction resolution and crossover formation. Curr Biol *14*, R56-58.

Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. Nature *411*, 366-374.

Hopfner, K.P., and Tainer, J.A. (2003). Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. Curr Opin Struct Biol *13*, 249-255.

Hopkins, B.B., and Paull, T.T. (2008). The P. furiosus mre11/rad50 complex promotes 5' strand resection at a DNA double-strand break. Cell *135*, 250-260.

Iyer, L.M., Leipe, D.D., Koonin, E.V., and Aravind, L. (2004a). Evolutionary history and higher order classification of AAA+ ATPases. J Struct Biol *146*, 11-31.

Iyer, L.M., Makarova, K.S., Koonin, E.V., and Aravind, L. (2004b). Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. Nucleic Acids Res *32*, 5260-5279.

Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E.*, et al.* (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. Nat Methods *5*, 53-55.

Kelman, Z., and White, M.F. (2005). Archaeal DNA replication and repair. Curr Opin Microbiol *8*, 669-676.

Kimonis, V.E., Mehta, S.G., Fulchiero, E.C., Thomasova, D., Pasquali, M., Boycott, K., Neilan, E.G., Kartashov, A., Forman, M.S., Tucker, S.*, et al.* (2008). Clinical studies in familial VCP myopathy associated with Paget disease of bone and frontotemporal dementia. Am J Med Genet A *146A*, 745-757.

Koch, C.A., Agyei, R., Galicia, S., Metalnikov, P., O'Donnell, P., Starostine, A., Weinfeld, M., and Durocher, D. (2004). Xrcc4 physically links DNA end processing by polynucleotide kinase to DNA ligation by DNA ligase IV. Embo J *23*, 3874-3885.

Korinek, A., Beck, F., Baumeister, W., Nickell, S., and Plitzko, J.M. (2011). Computer controlled cryo-electron microscopy--TOM(2) a software package for high-throughput applications. J Struct Biol *175*, 394-405.

Koster, A.J., Chen, H., Sedat, J.W., and Agard, D.A. (1992). Automated microscopy for electron tomography. Ultramicroscopy *46*, 207-227.

Lander, G.C., Martin, A., and Nogales, E. (2013). The proteasome under the microscope: the regulatory particle in focus. Curr Opin Struct Biol *23*, 243-251.

Lee, J.Y., and Yang, W. (2006). UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. Cell *127*, 1349-1360.

Leipe, D.D., Koonin, E.V., and Aravind, L. (2003). Evolution and classification of P-loop kinases and related proteins. J Mol Biol *333*, 781-815.

Lenzen, C.U., Steinmann, D., Whiteheart, S.W., and Weis, W.I. (1998). Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. Cell *94*, 525-536.

Li, H., O'Donoghue, A.J., van der Linden, W.A., Xie, S.C., Yoo, E., Foe, I.T., Tilley, L., Craik, C.S., da Fonseca, P.C., and Bogyo, M. (2016). Structure- and function-based design of Plasmodium-selective proteasome inhibitors. Nature *530*, 233-236.

Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat Methods *10*, 584-590.

Lucic, V., Forster, F., and Baumeister, W. (2005). Structural studies by electron tomography: from cells to molecules. Annu Rev Biochem *74*, 833-865.

Lupas, A.N., and Martin, J. (2002). AAA proteins. Curr Opin Struct Biol *12*, 746-753.

Lyubimov, A.Y., Strycharska, M., and Berger, J.M. (2011). The nuts and bolts of ring-translocase structure and mechanism. Curr Opin Struct Biol *21*, 240-248.

Maric, M., Maculins, T., De Piccoli, G., and Labib, K. (2014). Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. Science *346*, 1253596.

Massey, T.H., Mercogliano, C.P., Yates, J., Sherratt, D.J., and Lowe, J. (2006). Double-stranded DNA translocation: structure and mechanism of hexameric FtsK. Mol Cell *23*, 457-469.

McMullan, G., Faruqi, A.R., Clare, D., and Henderson, R. (2014). Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. Ultramicroscopy *147*, 156-163.

Medicherla, B., Kostova, Z., Schaefer, A., and Wolf, D.H. (2004). A genomic screen identifies Dsk2p and Rad23p as essential components of ER-associated degradation. EMBO Rep *5*, 692-697.

Merai, Z., Chumak, N., Garcia-Aguilar, M., Hsieh, T.F., Nishimura, T., Schoft, V.K., Bindics, J., Slusarz, L., Arnoux, S., Opravil, S.*, et al.* (2014). The AAA-ATPase molecular chaperone Cdc48/p97 disassembles sumoylated centromeres, decondenses heterochromatin, and activates ribosomal RNA genes. Proc Natl Acad Sci U S A *111*, 16166-16171.

Meyer, H.H., Kondo, H., and Warren, G. (1998). The p47 co-factor regulates the ATPase activity of the membrane fusion protein, p97. FEBS Lett *437*, 255-257.

Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature *455*, 770-774.

Moreau, M.J., McGeoch, A.T., Lowe, A.R., Itzhaki, L.S., and Bell, S.D. (2007). ATPase site architecture and helicase mechanism of an archaeal MCM. Mol Cell *28*, 304-314.

Myers, R.S., and Stahl, F.W. (1994). Chi and the RecBC D enzyme of Escherichia coli. Annu Rev Genet *28*, 49-70.

Neuwald, A.F., Aravind, L., Spouge, J.L., and Koonin, E.V. (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res *9*, 27-43.

Nimonkar, A.V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J.L., Wyman, C., Modrich, P., and Kowalczykowski, S.C. (2011). BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. Genes Dev *25*, 350-362.

Nyquist, K., and Martin, A. (2014). Marching to the beat of the ring: polypeptide translocation by AAA+ proteases. Trends Biochem Sci *39*, 53-60.

Olivares, A.O., Baker, T.A., and Sauer, R.T. (2016). Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nat Rev Microbiol *14*, 33-44.

Penczek, P.A. (2010). Image restoration in cryo-electron microscopy. Methods Enzymol *482*, 35-72.

Penczek, P.A., Grassucci, R.A., and Frank, J. (1994). The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. Ultramicroscopy *53*, 251-270.

Plaschka, C., Lariviere, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., Petrotchenko, E.V., Borchers, C.H., Baumeister, W., Herzog, F.*, et al.* (2015). Architecture of the RNA polymerase II-Mediator core initiation complex. Nature *518*, 376-380.

Quaiser, A., Constantinesco, F., White, M.F., Forterre, P., and Elie, C. (2008). The Mre11 protein interacts with both Rad50 and the HerA bipolar helicase and is recruited to DNA following gamma irradiation in the archaeon Sulfolobus acidocaldarius. BMC Mol Biol *9*, 25.

Richly, H., Rape, M., Braun, S., Rumpf, S., Hoege, C., and Jentsch, S. (2005). A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting. Cell *120*, 73-84.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol *192*, 216-221.

Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J Mol Biol *333*, 721-745.

Rothballer, A., Tzvetkov, N., and Zwickl, P. (2007). Mutations in p97/VCP induce unfolding activity. FEBS Lett *581*, 1197-1201.

Rouiller, I., DeLaBarre, B., May, A.P., Weis, W.I., Brunger, A.T., Milligan, R.A., and Wilson-Kubalek, E.M. (2002). Conformational changes of the multifunction p97 AAA ATPase during its ATPase cycle. Nat Struct Biol *9*, 950-957.

Ruskin, R.S., Yu, Z., and Grigorieff, N. (2013). Quantitative characterization of electron detectors for transmission electron microscopy. J Struct Biol *184*, 385-393.

Rzechorzek, N.J., Blackwood, J.K., Bray, S.M., Maman, J.D., Pellegrini, L., and Robinson, N.P. (2014). Structure of the hexameric HerA ATPase reveals a mechanism of translocation-coupled DNA-end processing in archaea. Nat Commun *5*, 5506.

Scheres, S.H. (2010). Classification of structural heterogeneity by maximum-likelihood methods. Methods Enzymol *482*, 295-320.

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol *180*, 519-530.

Scheres, S.H. (2014). Beam-induced motion correction for sub-megadalton cryo-EM particles. Elife *3*, e03665.

Scheres, S.H., and Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. Nat Methods *9*, 853-854.

Scheres, S.H., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P., Frank, J., and Carazo, J.M. (2007). Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. Nat Methods *4*, 27-29.

Schuller, J.M., Beck, F., Lossl, P., Heck, A.J., and Forster, F. (2016). Nucleotide-dependent conformational changes of the AAA+ ATPase p97 revisited. FEBS Lett *590*, 595-604.

Singleton, M.R., Dillingham, M.S., Gaudier, M., Kowalczykowski, S.C., and Wigley, D.B. (2004). Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. Nature *432*, 187-193.

Spies, M., Bianco, P.R., Dillingham, M.S., Handa, N., Baskin, R.J., and Kowalczykowski, S.C. (2003). A molecular throttle: the recombination hotspot chi controls DNA translocation by the RecBCD helicase. Cell *114*, 647-654.

Stein, A., Ruggiano, A., Carvalho, P., and Rapoport, T.A. (2014). Key steps in ERAD of luminal ER proteins reconstituted with purified components. Cell *158*, 1375-1388.

Sung, P., and Klein, H. (2006). Mechanism of homologous recombination: mediators and helicases take on regulatory functions. Nat Rev Mol Cell Biol *7*, 739-750.

Tang, W.K., Li, D., Li, C.C., Esser, L., Dai, R., Guo, L., and Xia, D. (2010). A novel ATP-dependent conformation in p97 N-D1 fragment revealed by crystal structures of disease-related mutants. Embo J *29*, 2217-2229.

Tang, W.K., and Xia, D. (2013). Altered intersubunit communication is the molecular basis for functional defects of pathogenic p97 mutants. J Biol Chem *288*, 36624-36635.

Tang, W.K., and Xia, D. (2016). Role of the D1-D2 Linker of Human VCP/p97 in the Asymmetry and ATPase Activity of the D1-domain. Sci Rep *6*, 20037.

Thomsen, N.D., and Berger, J.M. (2008). Structural frameworks for considering microbial protein- and nucleic acid-dependent motor ATPases. Mol Microbiol *69*, 1071-1090.

Trabuco, L.G., Villa, E., Mitra, K., Frank, J., and Schulten, K. (2008). Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. Structure *16*, 673-683.

Unverdorben, P., Beck, F., Sledz, P., Schweitzer, A., Pfeifer, G., Plitzko, J.M., Baumeister, W., and Forster, F. (2014). Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. Proc Natl Acad Sci U S A *111*, 5544-5549.

Vale, R.D. (2000). AAA proteins. Lords of the ring. J Cell Biol *150*, F13-19.

van Heel, M., and Schatz, M. (2005). Fourier shell correlation threshold criteria. J Struct Biol *151*, 250-262.

van Heel, M., and Stoffler-Meilicke, M. (1985). Characteristic views of E. coli and B. stearothermophilus 30S ribosomal subunits in the electron microscope. Embo J *4*, 2389-2395.

Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S., and Wigley, D.B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. Cell *97*, 75-84.

Verma, R., Oania, R., Fang, R., Smith, G.T., and Deshaies, R.J. (2011). Cdc48/p97 mediates UV-dependent turnover of RNA Pol II. Mol Cell *41*, 82-92.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. Embo J *1*, 945-951.

Wallden, K., Williams, R., Yan, J., Lian, P.W., Wang, L., Thalassinos, K., Orlova, E.V., and Waksman, G. (2012). Structure of the VirB4 ATPase, alone and bound to the core complex of a type IV secretion system. Proc Natl Acad Sci U S A *109*, 11348-11353.

Wang, J. (2004). Nucleotide-dependent domain motions within rings of the RecA/AAA(+) superfamily. J Struct Biol *148*, 259-267.

Wang, J., Song, J.J., Seong, I.S., Franklin, M.C., Kamtekar, S., Eom, S.H., and Chung, C.H. (2001). Nucleotide-dependent conformational changes in a protease-associated ATPase HsIU. Structure *9*, 1107-1116.

Wendler, P., Ciniawsky, S., Kock, M., and Kube, S. (2012). Structure and function of the AAA+ nucleotide binding pocket. Biochim Biophys Acta *1823*, 2-14.

Wendler, P., Shorter, J., Plisson, C., Cashikar, A.G., Lindquist, S., and Saibil, H.R. (2007). Atypical AAA+ subunit packing creates an expanded cavity for disaggregation by the protein-remodeling factor Hsp104. Cell *131*, 1366-1377.

Wigley, D.B. (2013). Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. Nat Rev Microbiol *11*, 9-13.

Williams, G.J., Lees-Miller, S.P., and Tainer, J.A. (2010). Mre11-Rad50-Nbs1 conformations and the control of sensing, signaling, and effector responses at DNA double-strand breaks. DNA Repair (Amst) *9*, 1299-1306.

Wilson, D.N., Guichard, G., and Innis, C.A. (2015). Antimicrobial peptides target ribosomes. Oncotarget *6*, 16826-16827.

Wolf, D.H., and Stolz, A. (2012). The Cdc48 machine in endoplasmic reticulum associated protein degradation. Biochim Biophys Acta *1823*, 117-124.

Wong, W., Bai, X.C., Brown, A., Fernandez, I.S., Hanssen, E., Condron, M., Tan, Y.H., Baum, J., and Scheres, S.H. (2014). Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. Elife *3*.

Wu, C.G., Bradford, C., and Lohman, T.M. (2010). Escherichia coli RecBC helicase has two translocase activities controlled by a single ATPase motor. Nat Struct Mol Biol *17*, 1210-1217.

Yang, F., Xiao, X., Cheng, W., Yang, W., Yu, P., Song, Z., Yarov-Yarovoy, V., and Zheng, J. (2015). Structural mechanism underlying capsaicin binding and activation of the TRPV1 ion channel. Nat Chem Biol *11*, 518-524.

Yang, Z., Fang, J., Chittuluru, J., Asturias, F.J., and Penczek, P.A. (2012). Iterative stable alignment and clustering of 2D transmission electron microscope images. Structure *20*, 237-247.

Zhang, X., Gui, L., Bulfer, S.L., Sanghez, V., Wong, D.E., Lee, Y., Lehmann, L., Lee, J.S., Shih, P.Y., Lin, H.J.*, et al.* (2015). Altered cofactor regulation with disease-associated p97/VCP mutations. Proc Natl Acad Sci U S A *112*, E1705-1714.

Zhao, M., Wu, S., Zhou, Q., Vivona, S., Cipriano, D.J., Cheng, Y., and Brunger, A.T. (2015). Mechanistic insights into the recycling machine of the SNARE complex. Nature *518*, 61-67.

Zhu, J., Penczek, P.A., Schroder, R., and Frank, J. (1997). Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70S Escherichia coli ribosome. J Struct Biol *118*, 197-219.

Zhu, S., Stein, R.A., Yoshioka, C., Lee, C.H., Goehring, A., McHaourab, H.S., and Gouaux, E. (2016). Mechanism of NMDA Receptor Inhibition and Activation. Cell.

Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E., and Ira, G. (2008). Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell *134*, 981-994.

# Abbreviations

| | |
|---|---|
| 3D | three-dimensional |
| AAA | ATPases Associated with diverse cellular Activities |
| ADP | Adenosine diphosphate |
| AMP-PNP | Adenylyl-imidodiphosphate |
| ASCE | additional strand catalytic glutamate |
| ATP | adenosine triphosphate |
| ATP-$\gamma$S | Adenosine 5′-[γ-thio]triphosphate |
| CCD | charge coupled device |
| CCF | cross correlation function |
| CET | cryo electron tomography |
| CMOS | Complementary metal-oxide-semiconductor |
| CTF | contrast transfer function |
| D-loop | displacement loop |
| DQE | detective quantum efficiency |
| DSB | doubel-strand break |
| dsDNA | double stranded DNA |
| EM | electron microscopy |
| EMDB | electron microscopy data bank |
| ERAD | endoplasmic reticulum associated degadation |
| FCR | Fourier cross resolution |
| FSC | Fourier shell correlation |
| GTP | guanosine triphosphate |
| HAS | HerA-ATP Synthase barrel |

| | |
|---|---|
| HR | homogous recombination |
| MS | mass spectrometry |
| NHEJ | non-homolgous end-joining |
| NU | Npl4-Ufd1 |
| PDB | protein data bank |
| P-loop | phophate-binding-loop |
| RNA | ribonucleic acid |
| ssDNA | single-stranded DNA |
| SAXS | Small- angle X-ray scattering |
| SNR | signal-to-noise ratio |
| SPA | single particle analysis |
| SRH | second region of homology |
| TEM | transmission electron microscopy |
| WA | Walker A motif |
| WB | Walker B motif |
| WT | wild type |
| XL | cross-linking |

# Acknowledgments

First of all I would like to thank my supervisor Dr. Friedrich Förster for giving me the opportunity to work in his group. I am grateful for his constant advice and leaving me the freedom to design my own experiments and projects. In particular, I appreciate very much all the scientific discussions and all the writing advice he gave me during the last years.

I am very grateful to Prof. Dr. Wolfgang Baumeister for giving me the opportunity to carry out my Ph.D. work in his lab and being able to use the state-of-the-art electron microscope equipment.

I would also like to thank Dr. Jürgen Plitzko for being such a great teacher and personal mentor. Thank you for always putting everything in order and for your constant support of my professional development.

Thanks to Oana Mihalache for joining me in the endless battle against the "preferred orientation bias" of various AAA+ complexes. In particular, I am grateful for several dozen manually plunged grids.

I am also grateful to Florian Beck, Dr. Radostin Danev, Antje Aufderheide, Dr. Stefan Pfeffer and Dr. Pia Unverdorben for your constant assistance in data processing and modelling. Thank you for your patience.

I also want to express my thanks all members of the Förster group and the Baumeister Department. It was a great time and I enjoyed working together with you.

In addition, I want to express my gratitude to Prof. Dr. Karl-Peter Hopfner, Dr. Robert Byrne and Dr. Agata Butryn from the LMU Gene Center and to Prof. Dr. Albert J.R. Heck and Philip Lössl from Utrecht University for great collaborations and inspiring scientific discussions.

Finally, I want to thank my parents and my girlfriend Sandra, with your constant support and encouragement you certainly also helped to finish this work.