Technische Universität München
TUM School of Natural Sciences

NMR-based studies of nucleotide states encountered during the ATP hydrolysis cycle of the p97 enzyme

Mikhail Shein

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

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Martin Zacharias

Prüfende der Dissertation:

1. Prof. Dr. Anne Schütz
2. Prof. Dr. Bernd Reif

Die Dissertation wurde am 23.01.2025 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 24.02.2025 angenommen.

Abstract

During its ATP hydrolysis cycle, the human AAA+ ATPase p97 transitions between four distinct protein and nucleotide states: apo, ATP-, ADP·P_i- and ADP-bound. Among these, the ADP·P_i state has remained elusive and has not been characterised as an authentic post-hydrolysis state of a AAA+ protein. Recently, the ADP·P_i state of p97 has been observed using solution-state NMR experiments in combination with an ATP regeneration system. Here, we attempt to characterise this rate-determining step of ATP hydrolysis using solution-state and solid-state NMR methods in combination with biophysical and –chemical techniques. As tools for structural analysis, we test several nucleotide analogues for their biological relevance and practical usefulness.

We observe that the ADP·P_i state shares several similarities with the pre-hydrolysis ATP-bound state, but the cleaved phosphate ion is found in a heterogeneous environment. This heterogeneity is linked to phosphate binding in multiple locations, including two sub-states of the ADP·P_i state and the subsequent phosphate ion dissociation through a positively charged channel towards a central pore. This phosphate release process is coupled to structural rearrangements in the sensor loop, a secondary structure element connecting two neighbouring binding pockets, and thus plays a role in intersubunit communication. Finally, we investigate the contributions of several conserved residues to the stability of the ADP·P_i state. Using mutational analysis, we find that most mutant constructs concerning these conserved residues are still able to form the ADP·P_i state, but with reduced P_i release rates.

Zusammenfassung

Während ihres ATP-Hydrolysezyklus durchläuft die menschliche AAA+-ATPase p97 vier verschiedene Protein- und Nukleotidzustände: apo, ATP-, ADP·P_i- und ADP-gebunden. Von diesen, entzog sich der ADP·P_i-Zustand der direkten Untersuchung und wurde bislang noch nicht als authentischer Post-Hydrolyse-Zustand eines AAA+-Proteins charakterisiert. Vor kurzem wurde der p97-ADP·P_i-Zustand mit Hilfe von Lösungs-NMR-Experimenten in Gegenwart eines ATP-Regenerationssystems beobachtet. In dieser Arbeit charakterisieren wir diesen geschwindigkeitsbestimmenden Schritt der ATP-Hydrolyse mit Hilfe von Lösungs- und Festkörper-NMR in Kombination mit biophysikalischen und –chemischen Methoden. Als Hilfsmittel der Strukturanalyse testen wir mehrere Nukleotid-Analoga auf ihre biologische Relevanz und praktische Nützlichkeit.

Der ADP·P_i-Zustand weist Ähnlichkeiten zu dem ATP-gebundenen pre-Hydrolyse-Zustand auf, das gespaltene Phosphat-Ion befindet sich jedoch in einer heterogenen Umgebung. Diese Heterogenität kann mit der Phosphatbindung an mehreren Positionen erklärt werden, u.a. in zwei Unterzuständen des ADP·P_i-Zustands und mit der anschließenden Dissoziation des Phosphat-Ions durch einen positiv geladenen Kanal in Richtung der zentralen Pore. Dieser Phosphatfreisetzungsprozess ist an Konformationsänderungen in der Sensorschleife, einem Sekundärstrukturelement, das zwei benachbarte Bindungstaschen verbindet, gekoppelt und spielt somit eine Rolle bei der Kommunikation zwischen den Untereinheiten. Schließlich untersuchen wir Beiträge mehrerer konservierter Aminosäuren zur Stabilität des ADP·P_i-Zustands. Mithilfe von Mutationsanalysen stellen wir fest, dass die meisten mutierten Konstrukte, die diese konservierten Aminosäuren betreffen, immer noch in der Lage sind, den ADP·P_i-Zustand zu bilden, allerdings mit reduzierten P_i-Freisetzungsraten.

List of publications

The main portion of the results presented in this thesis has been previously published. The following lists my co-authored articles, and my scientific contribution to the publications is explained.

a) Rydzek, S., **Shein, M.**, Bielytskyi, P., Schütz, A.K. "Observation of a transient reaction intermediate illuminates the mechanochemical cycle of the AAA-ATPase p97." *J. Am. Chem. Soc.*, 2020, 142, 34, 14472–14480

b) **Shein, M.***, Hitzenberger M.*, Cheng, T. C.*, Rout, S. R., Leitl, K. D., Sato, Y., Zacharias, M., Sakata, E., Schütz, A. K. "Characterizing ATP processing by the AAA+ protein p97 at the atomic level." *Nat. Chem.*, 2024, 16, 363-372.
* denotes equal contributions

c) Szántó, J. K., Dietschreit, J. C. B., **Shein, M.**, Schütz, A. K., Ochsenfeld, C. "Systematic QM/MM Study for Predicting ^{31}P NMR Chemical Shifts of Adenosine Nucleotides in Solution and Stages of ATP Hydrolysis in a Protein Environment." *J. Chem. Theory Comp.* 2024, 20, 2433-2444.

In publication a), the ADP·P_i state of p97 was observed for the first time. S. Rydzek has performed the methyl TROSY NMR experiments; ^{31}P solid-state NMR measurements were contributed by me (together with P. Bielytskyi).

In publication b), three authors share the first co-authorship. My scientific contribution to that article concerns all NMR measurements and mutational analysis by biophysical and –chemical methods.

To publication c), I have contributed ^{31}P solution-state NMR measurements.

Generally, results obtained by my co-authors are not shown in this thesis; instead, when necessary, I refer to the original publication. Unless otherwise stated, all results presented in this thesis were obtained either directly by me alone or through shared effort.

The following manuscripts have been published over the course of the PhD program but are not covered in this thesis.

a) Wadenpohl, T., **Shein, M.**, Steinberg, J., Lehmann, J. B., Schütz, A. K., Jung, S. Economical large-scale purification of extracellular vesicles from urine. *Sep. Purif. Technol.*, 2024, 335, 126155.

b) Su, J., Harati Taji, Z., Kosinska, A. D., Oz, E. A., Xie, Z., Bielytskyi, P., **Shein, M.**, Hagen, P., Esmaeili, S., Steiger, K., Protzer, U., Schütz, A. K. Introducing adjuvant-loaded particulate hepatitis B core antigen as an alternative therapeutic hepatitis B vaccine component. *JHEP Rep.*, 2024, 6, 100997.

c) Harati Taji, Z., Bielytskyi, P., **Shein, M.**, Sani, M.-A., Seitz, S., Schütz, A.K., Transient RNA Interactions Leave a Covalent Imprint on a Viral Capsid Protein, *J. Am. Chem. Soc.*, 2022, 144, 19, 8536–8550.

d) Albanese, M., Chen, Y.-F. A., Hüls, C., Gärtner, K., Tagawa, T., Mejias-Perez, E., Keppler, O. T., Göbel, C., Zeidler, R., **Shein, M.**, Schütz, A. K., Hammerschmidt, W. MicroRNAs are minor constituents of extracellular vesicles that are rarely delivered to target cells. *PLoS Genet.*, 2021, 17, 12:e1009951.

e) Jung, S., Jacobs, K. F. K., **Shein, M.**, Schütz, A. K., Mohr, F., Stadler, H., Stadler, D., Lucko, A. M., Altstetter, S. M., Wilsch, F., Deng, L., Protzer, U. Efficient and reproducible depletion of hepatitis B virus from plasma derived extracellular vesicles. *J. Extracell. Vesicles.*, 2020**,** 10:e12040.

f) Jung, S., Altstetter, S.M., Wilsch, F., **Shein, M.**, Schütz, A.K., Protzer, U. Extracellular vesicles derived from Hepatitis-D Virus infected cells induce a proinflammatory cytokine response in human peripheral blood mononuclear cells and macrophages. *Matters*, 2020, 6, 2, 1-10.

Table of Contents

List of Figures

List of Tables

Chapter 1: Biomolecular NMR Spectroscopy

1.1 Introduction

Nuclear Magnetic Resonance (NMR) spectroscopy is a method that allows the detection of individual atoms in molecules based on the intrinsic properties of their nucleus. This is achieved by using the spin properties of nuclei, which lead to energy level splitting under static magnetic fields. Oscillating magnetic fields are then used to probe this exact energy splitting, which is highly dependent on the nucleus' precise chemical and magnetic environment. Biomolecular NMR spectroscopy extends the range of samples to biological molecular entities, ranging from peptides and single nucleotides to full-size proteins, nucleic acids and large biomolecular complexes, e.g. virus-like particles, lipids or enzymatic machinery. (e.g. [1,2]).

1.2 Spins

This section has been largely derived from the following textbook:
Levitt, M. H. Spin dynamics: basics of nuclear magnetic resonance. (Wiley, 2013, 2nd ed.).

A *spin* is an intrinsic property of an elementary particle (e.g. a nucleus or an electron) and is akin to the quantum angular momentum. A spin angular momentum S of a particle is quantised into $(2S+1)$ degenerate levels. In the case of a nucleus, the nuclear spin quantum number is denoted as I and can take integer or half-integer values: $I=0, 1/2, 1, 3/2, 2, 5/2\ldots$
The spin quantum number is inherent to a nucleus and depends on the number of protons and neutrons, meaning that different isotopes of the same element can differ in their nuclear spin quantum number. While $I=0$ nuclei are NMR-inactive and unobservable, nuclei with spin quantum numbers $I>½$ possess quadrupolar moments, meaning that their electric charge distribution is non-spherical. Quadrupolar moments facilitate relaxation (see below), leading to broader spectral lines. NMR spectroscopy of such nuclei is technically more difficult. Thus, first and foremost, spin ½-nuclei are of importance and usefulness to NMR spectroscopists.

In biomolecular NMR spectroscopy, spin ½-nuclei such as ^{1}H, ^{13}C and ^{15}N play the most significant roles due to the availability of these elements in biological samples, such as proteins and nucleic acids. The limitation of their application lies in the low natural abundance of these nuclei, with ~1% for ^{13}C and only 0.4% for ^{15}N (Table 1). Additionally, low gyromagnetic ratios of these isotopes diminish experimental sensitivity and limit the usefulness of a direct detection. Instead, other spectroscopic techniques, such as indirect observation or signal enhancement, are employed in order to detect signals originating from those spins.

Table 1. Overview of NMR relevant properties of selected isotopes. The values were derived from ref. [3].

Isotopes	**Natural abundance (%)**	**Gyromagnetic ratio $\frac{\gamma}{2\pi}$ (MHz/T)**	**Spin quantum number I**
^{1}H	99.99	42.58	½
^{2}H	0.01	6.54	1
^{12}C	98.93	-	0
^{13}C	1.07	10.71	½
^{14}N	99.63	3.08	1
^{15}N	0.37	4.31	½
^{16}O	99.76	-	0
^{17}O	0.037	5.77	5/2
^{19}F	100	40.08	½
^{29}Si	4.7	8.47	½
^{31}P	100	17.25	½
^{103}Rh	100	1.35	½
^{195}Pt	33.8	9.29	½

These limitations are absent for a few spin ½-nuclei, most notably ^{1}H, ^{19}F and ^{31}P, which possess a (close to) 100% natural abundance and high gyromagnetic ratios and are often used in biomolecular NMR spectroscopy.

1.3 Zeeman splitting and Larmor frequency

For a spin ½-nucleus, two spin states are possible, denoted as $m = -\frac{1}{2}$ and $m = +\frac{1}{2}$. In the presence of an external magnetic field **B₀**, the degenerate nature of these states is lost, and an energetic splitting, called nuclear Zeeman splitting, is observed (Figure 1).

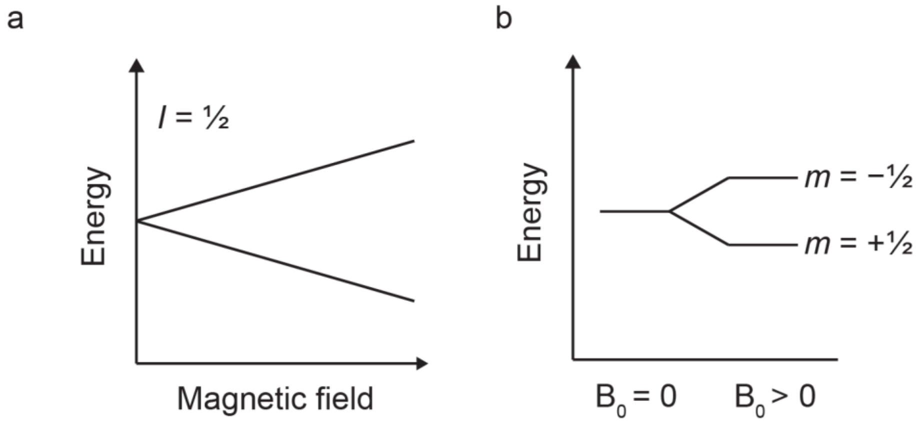


Figure 1. The presence of an external magnetic field induces energetic splitting of spin states. a) Nuclear Zeeman splitting is proportional to the strength of the magnetic field. **b**) In the absence of an external magnetic field, the two spin states of a spin ½ are degenerate (left). In the presence of an external magnetic field, the spin "up" state with $m = +½$ is energetically favoured, while the spin "down" state with $m = -½$ has a higher energy. Both panels exemplify the splitting for a spin-½ nucleus with γ>0.

This strength of the splitting is both proportional to the magnetic field and the gyromagnetic ratio γ, a nucleus-specific proportionality constant describing the ratio between the spin angular momentum and a magnetic moment μ:

$$E = -\vec{\mu} \cdot \hbar \cdot \vec{B}_0 = -\gamma \cdot \vec{S} \cdot \hbar \cdot \vec{B}_0,$$

where ħ is the reduced Planck constant. The negative sign indicates that the energy is lower for a state in which the magnetic moment is aligned parallel to the magnetic field, and the energy is maximised for an antiparallel alignment. For a spin ½ nucleus with γ>0, the energy level splitting is depicted in Figure 1b. These states are conventionally labelled as α ($m = +\frac{1}{2}$) and β ($m = -\frac{1}{2}$).

In thermal equilibrium, the nuclei populate both states almost equally according to the Boltzmann distribution:

$$\frac{N_\alpha}{N_\beta} = e^{-\frac{\Delta E}{k_B T}}$$

where $N_{\alpha,\beta}$ represent the respective populations, ΔE is the energy difference between the two states, k_B is the Boltzmann constant and T the temperature.

The presence of a magnetic field forces the spin to precess around the field axis on a precession cone with a constant angular frequency, called the Larmor frequency ω_0 (Figure 2a):

$$\omega_0 = -\gamma \cdot B_0$$

The selection rule for the transition between different spin states is given by $\Delta m = \pm 1$, implying that only transitions between $m = -\frac{1}{2}$ and $m = +\frac{1}{2}$ states are allowed. Absorption of electromagnetic radiation by a spin can change its spin state if the resonance condition is satisfied:

$$\Delta E = \hbar\omega = \left[-\left(-\frac{1}{2}\right)\gamma \cdot \hbar \cdot \mathrm{B}_0\right] - \left[-\left(+\frac{1}{2}\right)\gamma \cdot \hbar \cdot \mathrm{B}_0\right] = \gamma \cdot \hbar \cdot \mathrm{B}_0 = \hbar\omega_0$$

This shows that an irradiation with a radio frequency corresponding to the Larmor frequency leads to a spin state transition.

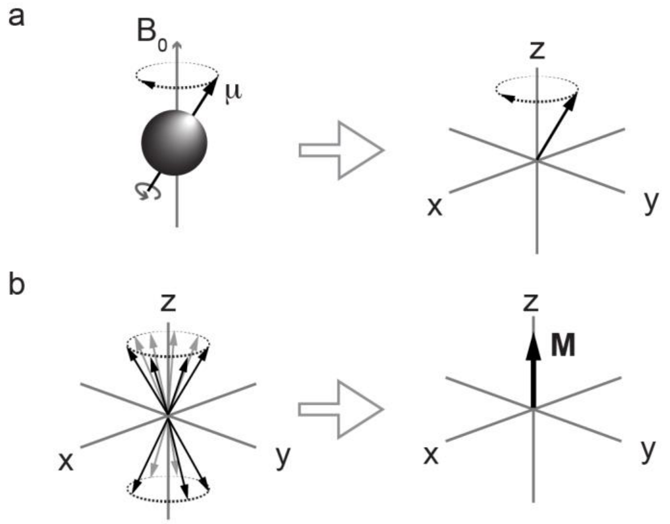


Figure 2. Spin precession around a magnetic field. a) Spins precess around an external magnetic field B_0, aligned along the z-axis. The individual spin can be designated by an arrow in a vector model. **b**) The precession of multiple spins forms a double cone. The resulting macroscopic bulk magnetisation **M** is aligned along $\mathbf{B_0}$, and thus the z-axis.

Even at moderately high fields of $B_0 = 11.74$ T (corresponding to a ^{1}H spin Larmor frequency of 500 MHz), the energy difference between the two states is approximately four orders of magnitude smaller than the available thermal energy at room temperature. This leads to a minimal overpopulation of the lower energy state and explains the relative insensitivity of NMR in comparison to other spectroscopic techniques in which the ground and excited states have higher energy differences.

Nonetheless, macroscopically, a bulk magnetisation **M**, originating from the collection of orientations of all the individual spins, can be observed. Since the spins preferentially align parallel to the magnetic field, in thermal equilibrium the bulk magnetisation vector $\mathbf{M}_0$ points along the axis of the external field, conventionally the z-axis (Figure 2b). Due to the spin precession around the field axis no preferential orientation in x or y directions exist. In the vector model representation, this is shown by a double cone along the z-axis. The spin magnetisation along the z-axis is termed longitudinal magnetisation, while the magnetisation along the x- and y-directions is called transverse magnetisation.

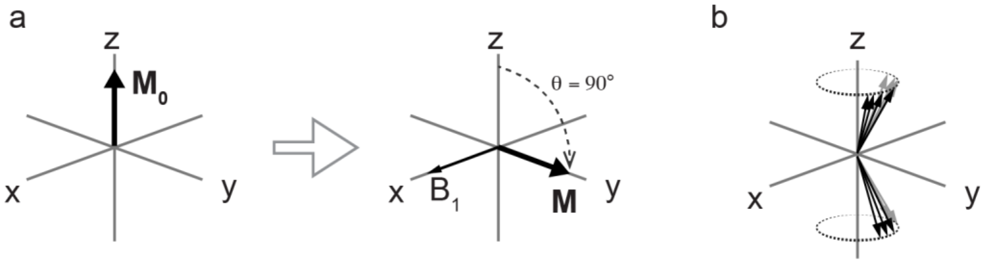


Figure 3. Macroscopic magnetization is manipulated by the application of radiofrequency pulses. This leads to the manipulation of individual spins. **a**) Equilibrium magnetization M_0 is aligned along the z-axis. Upon irradiation with a radiofrequency pulse $\mathbf{B}_1$ along the x-axis, the macroscopic magnetization is flipped by a flip angle θ=90° towards the y-axis. **b**) The macroscopic alignment of the magnetization vector equates to a phase coherence of the individual spin vectors along the y-axis. After the initial flip, the bulk magnetization (and with it the individual spins) continue to precess around the z-axis.

In NMR spectroscopy, radiofrequency pulses are described by their corresponding field strengths instead of wavelengths and denoted as $\mathbf{B}_1$. Spin orientations can be manipulated using resonant radiofrequency pulses applied along the x or y-axes. Since all identical spins are resonant with the same frequency and experience the $\mathbf{B}_1$ field identically, the bulk magnetisation **M** can thus be manipulated. A $\mathbf{B}_1$ pulse along the x-axis with an appropriate duration can thus flip the magnetisation vector towards the -y direction by a flip angle θ. The most important pulse, a 90° pulse, flips the magnetisation vector onto the transverse x-y plane, equalising the populations of the α and β states, resulting in zero z-magnetisation (Figure 3a). Following such a pulse, the individual spin vectors become aligned (or phase coherent) but continue to precess along the z-axis (Figure 3b).

1.4 Relaxation

Any perturbation from the equilibrium state of the bulk magnetisation is met with processes trying to re-establish the original state. Two mechanisms are of importance: the longitudinal or T_1-relaxation and the transverse or T_2-relaxation.

Longitudinal relaxation describes a process to recover the magnetisation along the z-axis $\mathbf{M}_z$ from the starting magnetisation $\mathbf{M}_0$, diminishing the transverse magnetisation at the same time. An important source of longitudinal relaxation lies with fluctuating magnetic fields at Larmor frequency caused by thermal molecular motions leading to stimulated emissions. The main contributor to these fluctuating magnetic fields are nuclear spin-spin couplings, which scale with the gyromagnetic ratios and are thus particularly strong for ^1H spins (see below). The magnetisation recovery follows an exponential behaviour:

$$M_z = M_0 \cdot \left(1 - e^{-\frac{t}{T_1}}\right)$$

The time constant T_1 falls typically in the range of 0.5-10 s for protons but can reach many tens of seconds for other nuclei such as ^{13}C or ^{31}P, especially, if no ^1H spins are located in close proximity [4,5]. Normally, a full relaxation of spins after a 90° pulse is achieved after a time corresponding to $5 \cdot T_1$, which can significantly increase the time between repeated measurements, and this so-called relaxation delay time is often the limiting factor for better signal-to-noise in NMR spectra. In real-life measurements, one must consider the variability of T_1 times among same spins located in different chemical environments, especially, if spins need to fully relax between a subsequent measurement (e.g. for quantitative analysis).

Transverse relaxation describes the loss of the magnetisation in the x-y plane by dephasing of the individual spins out of coherence. Its mains source are local magnetic fields originating from intra- and intermolecular interactions, most notably from dipole-dipole interactions. An additional source of coherence loss comes from the static magnetic field inhomogeneity, causing inhomogeneous broadening. This refers to a broadening of line widths of identical spins in identical environments experiencing slightly different magnetic fields due to technical limitations.

As with the longitudinal relaxation, the magnetisation recovery follows an exponential behaviour:

$$M_{x,y} = M_0 \cdot \left(1 - e^{-\frac{t}{T_2}}\right)$$

T_2 times fall in the range of milliseconds to seconds. The most noticeable effect of the transverse relaxation lies in its connection to half-height linewidths of Lorentzian-shaped peaks $\Delta\nu_{1/2}$:

$$\Delta\upsilon_{1/2} = \frac{1}{\pi \cdot T_2^*}$$

Here, T_2^* denotes the effective time constant, combining contributions from the genuine transverse relaxation and the one driven by field inhomogeneity. Faster T_2 relaxation increases line widths and decreases peak resolution while also lowering the signal-to-noise due to smaller but broader signals.

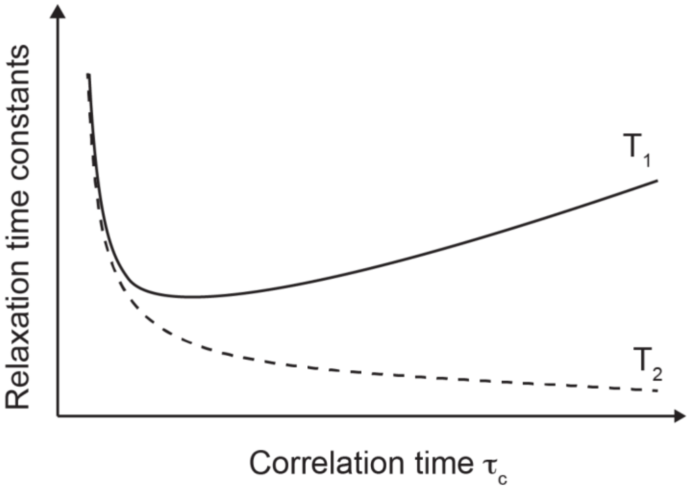


Figure 4. Dependence of the relaxation time constants on the correlation time τ_c. While both T_1 and T_2 equalise for small correlation times which are found for small, fast tumbling molecules, T_1 passes through a minimum and increases again with larger correlation time. On the other hand, T_2 decreases with correlation time.

Generally, since the longitudinal magnetisation along the z-axis cannot be increased without decreasing the transverse magnetisation, $2\,T_1 \geq T_2$ holds true. In practice, the relationship between T_1 and T_2 depends on prevalent molecular motions present in the sample. Fluctuating magnetic fields at resonant frequencies originate from random molecular tumbling, which occurs at a range of different frequencies. The rate of molecular tumbling is characterised by a correlation time $\tau_{c,}$ which describes the average time for the molecule to rotate through one radian. Generally, small molecules in non-viscous solutions tumble fast and have short correlation times (~10 ps), while large molecules, including biomolecular entities, display slow motions and long correlation times (~10 ns; Figure 4). Empirically, rotational correlation time is approximately proportional to the molecular weight [6]. Thus, for small molecules, T_1 and T_2 can be equal, but T_2 becomes significantly shorter than T_1 for proteins, leading to line broadening. In solids, static dipolar interactions are very large, so T_2 relaxation becomes especially efficient, leading to $T_2 << T_1$.

1.5 NMR experiment

In simple terms, the NMR spectrometer consists of a superconducting magnet (generating a large static magnetic field), a probe (holding the sample and housing the radiofrequency coil), a radiofrequency synthesiser and a computer for processing input and output. Since magnetisation can be only recorded in the x-y plane, in the simplest NMR experiment, the bulk magnetisation is first flipped by a 90° pulse into the transverse plane by a radiofrequency pulse applied along the x-axis. The pulse is generated by the synthesiser, which sends it to the coil located in the probe and is as close to the sample as possible. The coil geometry allows the pulse to be applied perpendicularly to the static magnetic field.

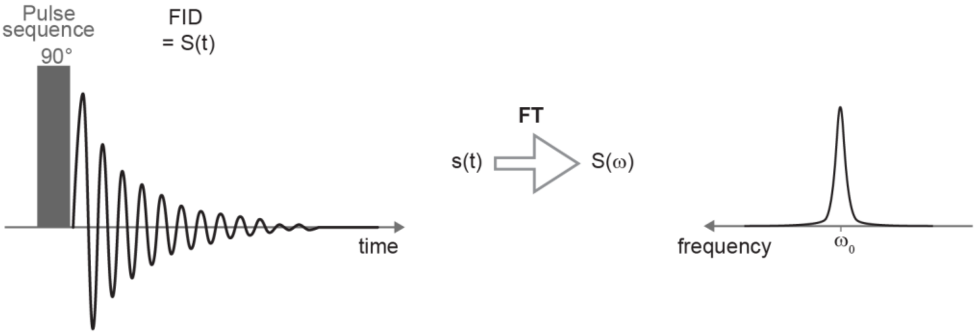


Figure 5. Schematic representation of the measurement process. A pulse sequence (here: a single 90° pulse) is applied, after which the free induction decay (FID) is recorded. The resulting oscillation corresponds to a signal in the time domain. In order to translate the signal to the frequency domain, Fourier Transformation (FT) is performed, giving rise to a signal at frequency ω_0.

After the pulse, the bulk magnetisation starts to precess around the z-axis. This rotating magnetic field induces weak electric currents in the coil, which are then recorded as an oscillatory NMR signal. Due to relaxation effects, the signal intensity dampens exponentially and quickly fades. This NMR signal is thus referred to as the free induction decay, FID. The FID, a signal recorded in the time domain, is then transformed into the frequency domain using Fourier Transformation (Figure 5).

1.6 Chemical shift

While all spins experience the same external static magnetic field, the individual spins experience additional effects. The total magnetic field felt by a spin is altered by local environmental effects, which can add to or subtract from the external magnetic field, effectively shielding or deshielding the nucleus. The main source of such an effect is the immediate electronic structure of the spin, namely electrons in the outer atomic shells and those that are localised in chemical bonds. These electrons generate an additional magnetic field and thus change the Larmor frequency of the spin, which is referred to as the chemical shift δ of the spin. The overall change in Larmor frequency due to the chemical shift is generally small in comparison to the size of the external field. Since both the Larmor frequency and the chemical shift are proportional to the external field and the differences between the individual shifts are very small, chemical shifts are generally reported using a relative deshielding scale normalised to a reference compound:

$$\delta = \frac{\omega_0 - \omega_{\text{ref}}}{\omega_{\text{ref}}}$$

where ω_{ref} is the Larmor frequency of the reference compound. In this scale, chemical shifts are reported in units of parts per million (ppm) from the reference shift. For aqueous samples, typically, sodium trimethylsilylpropanesulphonate (DSS) is used as an internal reference compound of ^{1}H spectra with a sharp peak calibrated to 0 ppm.

1.6.1 Chemical shift anisotropy

In reality, chemical shifts need to be treated as tensors instead of as scalars, implying a chemical shift anisotropy (CSA), i.e. an orientation dependence on the induced fields $\vec{B}^{\text{induced}}$:

$$\vec{B}^{\text{induced}} = \vec{\delta} \cdot \vec{B}_0$$

leading to

$$\vec{B}^{\text{induced}} = \begin{pmatrix} \delta_{xx} & \delta_{xy} & \delta_{xz} \\ \delta_{yx} & \delta_{yy} & \delta_{yz} \\ \delta_{zx} & \delta_{zy} & \delta_{zz} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix}$$

Using a spin centred, molecule specific (so-called “principal”) axis frame of reference aligning with the local molecular symmetries in order to transform the respective chemical shift tensor, one can get the isotropic chemical shift

$$\delta^{iso} = \frac{1}{3} \cdot (\delta_{XX} + \delta_{YY} + \delta_{ZZ})$$

where δ_{JJ} are the principal values of the chemical shift tensor. In usual cases, the external magnetic field is not applied along the principal axes and molecular orientation needs to be accounted for as well. In a tensor notation, Θ describes the relative orientation of the principal axis and the external field. Thus, motional averaging adds an orientational dependency to the chemical shift in the form of

$$\omega_0 = -\gamma \cdot B_0 \cdot (1 + \overline{\delta_{ZZ}(\Theta)})$$

where $\overline{\delta_{ZZ}(\Theta)}$ is the motionally averaged term δ_{ZZ}. In isotropic liquids, however, due to rapid tumbling, this averages to

$$\omega_0 = -\gamma \cdot B_0 \cdot (1 + \delta^{iso})$$

1.7 Nuclear spin-spin coupling

Additionally, spins feel magnetic fields created by neighbouring nuclei, leading to a so-called nuclear spin-spin coupling. Direct spin-spin coupling, also called dipolar coupling, emanates from nuclear spins that generate directional magnetic fields that possess dipolar qualities. These fields act through space and are felt by neighbouring spins. The magnitude of the dipole-dipole coupling constant b between spin j and k is given by

$$b_{jk} = -\frac{\mu_0}{4\pi} \cdot \frac{\gamma_j \gamma_k \hbar}{r_{jk}^3}$$

with μ_0 being the magnetic constant and *r* the distance between the two spins. By accounting for the orientation between the two spins and the (very strong) external magnetic field, the dipole-dipole coupling is found:

$$d_{jk} = b_{jk} \frac{1}{2} \cdot (3\cos^2 \Theta_{jk} - 1)$$

Here, Θ_{jk} denotes the angle between the unit vector joining the two spins *j* and *k* and the vector of the external field (Figure 6). While this interaction is strong in nature, in isotropic liquids, the dipolar coupling is averaged out due to rapid molecular tumbling and internal motion and its effect on the NMR spectrum essentially disappears.

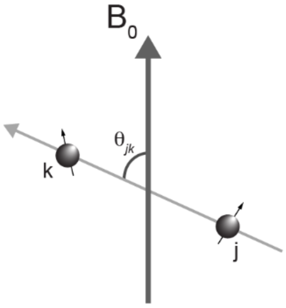


Figure 6. Orientation dependence of dipole-dipole coupling. Dipolar spin-spin coupling between spins k and j is defined by the angle Θ_{jk} between the unit vector joining the two spins and the vector of the external magnetic field.

In contrast, the indirect spin-spin coupling has a major effect on spectral forms and chemical shifts in liquids as well. These coupling interactions, called *J*-coupling, act through chemical bonds and survive the averaging of molecular motions. Effectively, mediated through bonding electrons, neighbouring nuclear spins align their orientation either parallel or antiparallel to each other, leading to a further energy level splitting. Phenomenologically, *J*-coupling to neighbouring nuclei creates a splitting of individual peaks into a multiplet structure. The size of the coupling is independent of the external magnetic field but depends on the number of bonds between the coupled spins, the type of involved spins, bond and torsion angles and other factors. Directly bonded (^1J) spin coupling constants between ^{1}H and ^{13}C spins can reach 80-150 Hz in size, while three-bond *J*-coupling (^3J) constants are only 1-10 Hz.

1.8 Sensitivity of NMR

NMR spectroscopy is a rather insensitive technique that requires a long measurement time due to experimental repetition (either to improve signal-to-noise or to attain multi-dimensional spectra) and inherently long spectral acquisition times of several hundred milliseconds. Signal-to-noise ratio (SNR) of NMR spectra is generally proportional to

$$\mathrm{SNR} \sim N \cdot B_0^{3/2} \cdot \gamma^{5/2} \cdot T_2 \cdot \sqrt{\mathrm{NS}}$$

with *N* being the number of identical spins in the sample and NS the number of scans. This implies that SNR can be improved by increasing the sample concentration, recording the spectra at higher field magnets and exciting/detecting nuclei with high gyromagnetic ratios (e.g. ^{1}H). Additionally, longer T_2 values are favourable, but this is often dictated by sample properties. Finally, SNR can be improved by simply repeating the experiment and collecting more scans; however, since the SNR only increases with a square root of the number of scans, this approach has its limits: if a spectrum takes 10 minutes to record, improving its SNR by a factor of two would require a forty-minute measurement, but to improve its SNR by a factor of 10 would already take 16.7 hours. Multidimensional spectra inherently require more measurement time since each experiment needs to be repeated many times to increment the evolution period(s) (see below). Direct observation of nuclei other than ^{1}H additionally suffers from lower gyromagnetic ratios, further decreasing SNR. Additionally, heteronuclei often display longer T_1 relaxation times, extending the relaxation delay between repeated measurements even further. In practice, NMR measurements often take hours to days to be recorded.

1.9 Biomolecular solution-state and solid-state NMR

The low natural abundance of ^{13}C and ^{15}N nuclei prevented NMR spectroscopists from studying proteins for a long time. The development of recombinant protein expression opened doors for the production of labelled macromolecular samples [7]. Specifically, the production of recombinant proteins in minimal media supplied with isotope labelled nutrients such as ^{13}C-glucose or $^{15}NH_4Cl$ by host organisms such as Escherichia coli (*E.coli*) allows labelled sample production in large enough quantities that are needed for NMR analysis.

Biological samples like proteins and nucleic acids usually require solvating water to support their tertiary and quaternary structures. Additionally, pH value, ionic strength and composition, etc., need to be precisely controlled for the samples to be stable over a long period of time – a prerequisite to being able to record the desired NMR spectra. Many biological species are highly soluble in water, leading to solution-state NMR being an obvious method for sample characterisation. Often, two factors determine whether this approach is viable: 1) sample concentration and 2) molecular weight. Protein stability and material availability can limit achievable sample concentrations, leading to a low number of spins in the sample tube. For biomolecular NMR, sample concentrations typically range from 10 to 1000 µM. At the same time, molecular weight has a major impact on the SNR as well. As discussed previously, a higher molecular weight increases the correlation time and thus leads to a faster T_2 relaxation. The fast coherence dephasing of spins broadens the NMR signals, called peaks, decreasing SNR or even broadening beyond recognition. Typically, solution-state NMR spectroscopy is readily applicable up to a molecular weight of 30-50 kDa, although certain labelling schemes and pulse sequences (see Section 1.12.4) can increase acceptable molecular weight ranges to even 100-500 kDa, although not without limiting the obtainable information. Finally, temperature plays a major role in improving signal-to-noise ratios. However, its effect on the NMR spectrum is not trivial to predict: increasing the temperature increases molecular motion and reduces the correlation time, thus increasing T_2 and decreasing line broadening. At the same time, chemical or conformational exchange processes, the rates of which are temperature dependent, also influence the spectral quality. For instance, decreasing the temperature slows down conformational exchange rates and thus potentially stabilises a certain structure. Moreover, in the case of a kinetic measurement, the temperature can influence the reaction rate and thus affect the maximal achievable experiment time. Finally, temperature affects the macroscopic magnetisation, since the population of spin states is defined by the (temperature-dependent) Boltzmann distribution.

Molecular entities exceeding the weight limit of 50 kDa are poorly suitable for solution-state NMR. Instead, solid-state NMR techniques might be applicable, which can be applied not only to powders and crystals but to immobilised samples in general. For instance, particularly large complexes like virus capsids and fibrils have been studied with solid-state NMR [1]. In viscous solutions, tumbling rates of very large biological entities are reduced so much that dipolar interactions become prevalent, a characteristic otherwise found in solid materials. The basis for this is the formation of a gel-like structure inside the sample container and the subsequent measurement of this sediment, sometimes referred to as sedNMR [8]. Thus, the limit of solid-state NMR applications is that molecules have a sufficient molecular weight in order to be sedimented into the sample container, a solid-state rotor. Using commercial ultracentrifuges reaching >100,000 g, sample sedimentation can be achieved for molecular entities of 100 kDa [8] and above in a reasonable time. However, it should still be noted that the viscosity of the gel has a great effect on the mobility of the sample, so dipolar-coupling based pulse sequences give better signals with more viscous gels.

1.10 Magic angle spinning

Dipolar couplings and chemical shift anisotropy are dominant sources of line broadening, homogeneous or inhomogeneous, respectively, and relaxation. These effects significantly reduce the spectral resolution and increase the relaxation rates. As previously described, these contributions disappear in isotropic liquids due to fast molecular motions. Molecular rotations and translations are absent (or significantly slowed down) in solids, so the orientational averaging cannot take place for either anisotropic interaction. In order to overcome this limitation, in solid-state NMR, Magic angle spinning (MAS) is performed: the sample is rotated at a high spinning speed in a rotor. Hereby, the angle of rotation is chosen such that the orientational part of the dipolar coupling (Section 1.7) becomes zero:

$$3\cos^2\Theta_{jk} - 1 = 0$$

An identical orientational dependence is found for the chemical shift anisotropy. This equation is fulfilled for the so-called magic angle:

$$\Theta_{magic} = \tan^{-1}\sqrt{2} = 54.74°$$

Rapid rotation of the sample about this angle with respect to the static field (Figure 7) can effectively average out CSA and dipolar interactions and ideally lead to a solution-state-like spectrum.

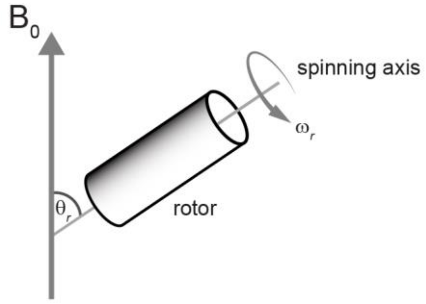


Figure 7. Magic angle spinning. In solid-state NMR, the sample containing rotor is spun rapidly about the spinning axis. If the angle of the spinning axis with respect to the external magnetic field θ_r equals 54.74° (the magic angle), then the CSA and dipolar interactions are effectively averaged out over one rotor period.

In practice, MAS zirconium oxide rotors with outer diameters of 0.7 mm to 4 mm and active volumes of ~0.5 to 80 µL with top rotation rates of 111 to 15 kHz, respectively, are available. While faster spinning speeds improve the desired averaging of anisotropic interactions (especially when it comes to homonuclear dipolar coupling), a further advantage can be found for sedimented samples: MAS applies a centrifugal force of >1,000,000 g, redistributing the sample towards the rotor walls [8]. Gel-like samples are thus further sedimented, improving the solid-like behaviour.

MAS rotation generally introduces an additional time dependence since the sample experiences orientational modulation with the rotational frequency ω_r. As a general rule, efficient averaging of anisotropic effects is achieved when ω_r is approx. 3-4 times larger than the magnitude of the interaction (in frequency units). For very fast spinning, the anisotropic effects become negligible. For lower spinning speeds, however, MAS can create additional spectral lines: the chemical shift tensor in the frame of the rotor is modulated by further terms that are dependent on the rotational frequency of the rotor. This creates new observable signals with frequencies $\omega = \omega_0 \pm n \cdot \omega_r$, with n being an integer. The position of these so-called spinning sidebands is dependent on the MAS speed as the sidebands “move” away from the main peak and, for sufficiently fast spinning, disappear completely.

1.11 Two-dimensional NMR

One-dimensional NMR spectra often lack resolution and only contain information about chemical shifts of one type of nuclei. Two-dimensional (or multidimensional in general) NMR spectroscopy extends the information content by adding another dimension to the spectra. That way, signals are resolved in a plane instead of a single axis, reducing peak overlap. Additionally, several spins can be coupled via a correlation of interest by observing cross peaks connecting resonances corresponding to the said spins (Figure 8).

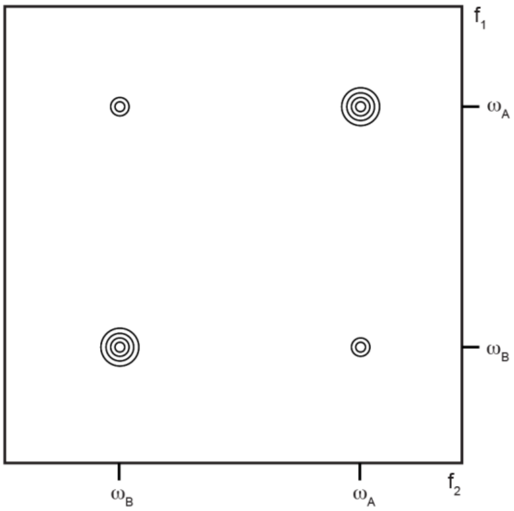


Figure 8. Simplified example of a two-dimensional NMR correlation map. The correlation map is shown as a contour plot, in which points of equal intensity (located in the 3rd dimension) are joined by contour lines. The spectrum correlates two resonances (with Larmor frequencies ω_A and ω_B) to each other. This correlation is indicated by the presence of cross peaks. The self correlation leads to diagonal peaks instead.

The principle of two-dimensional NMR is shown in Figure 9. Similar to a one-dimensional experiment, the pulse sequence for a 2D NMR experiment consists of a series of pulses and time periods, during which no pulses are applied, and a final recording of an FID at the end. The types and number of pulses can vary between experiments and depend on the correlation of interest. By definition, in a two-dimensional spectrum data points must be sampled in two separate dimensions. However, only one dimension can be sampled directly in form of an FID, leading to the observation of the direct dimension, while the second – indirect – dimension must be sampled by other means.

The first pulses are applied during the preparation period, generating some coherence which evolves during the evolution period of length t_1. However, the length of t_1 is not fixed, but incremented over a series of otherwise identical 1D experiments, creating a time dependence (t_1 domain). In the final two-dimensional spectrum, the t_1 domain is represented by the indirect dimension.

In a typical 2D experiment, during the preparation period magnetisation is excited for one type of nuclei, so that during this evolution period the magnetisation acquires a “frequency label”, meaning it is amplitude and/or phase modulated by the resonance frequency of these spins. The resonance frequency of these nuclei is then found in the indirect dimension of the final spectrum. Then, the magnetisation is transferred to the second type of nuclei in the mixing step and into an observable coherence. During the last step, the detection, a regular FID is acquired, collecting a signal over a time t_2. The magnetisation is again frequency modulated by the resonance frequency of these nuclei, and this modulation is directly observed through sampling in this t_2 domain. Afterwards, the 2D time-domain dataset is Fourier transformed in both dimensions. The resulting frequency-domain spectrum produces a correlation map, connecting the characteristic evolution frequency acquired in t_1 to the directly observed frequency of t_2.

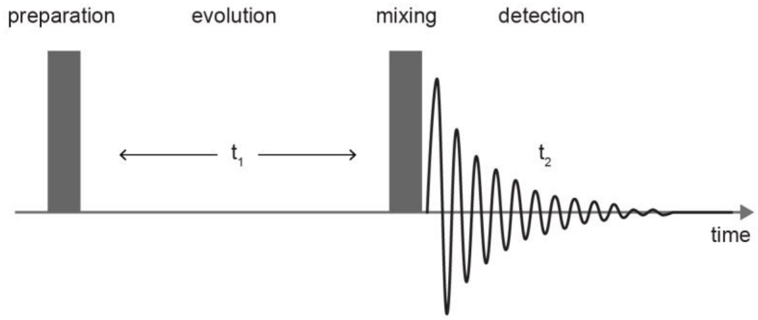


Figure 9. Principle of two-dimensional NMR spectroscopy. Four steps are indicated: preparation and mixing, during which radiofrequency pulses are applied; evolution, during which the magnetisation is modulated by the interaction of interest and a frequency label is acquired and which is dependent on the evolution time t_1; and the detection of length t_2, during which the FID is recorded.

1.12 NMR experiments employed in this thesis

In the following, selected NMR experiments used in this work in solution and solid-state NMR are presented and explained.

1.12.1 Direct excitation

The simplest type of experiment in solid-state NMR spectroscopy is direct excitation followed by acquisition of the FID. In such a 1D experiment, all nuclei of the same type are excited, and an FID is recorded immediately.

For example, ^{31}P nuclei are directly excited by a 90° pulse, and subsequently, the FID is recorded on the ^{31}P channel (Figure 10a). An additional element is responsible for proton decoupling during acquisition to remove broadening (due to ^{31}P-^{1}H dipolar coupling). The resulting spectrum provides information on all ^{31}P nuclei present in the sample, both in flexible and rigid moieties. However, due to long T_1 relaxation times of phosphorus, longer relaxation delays (~5-10 s) between subsequent scans are required.

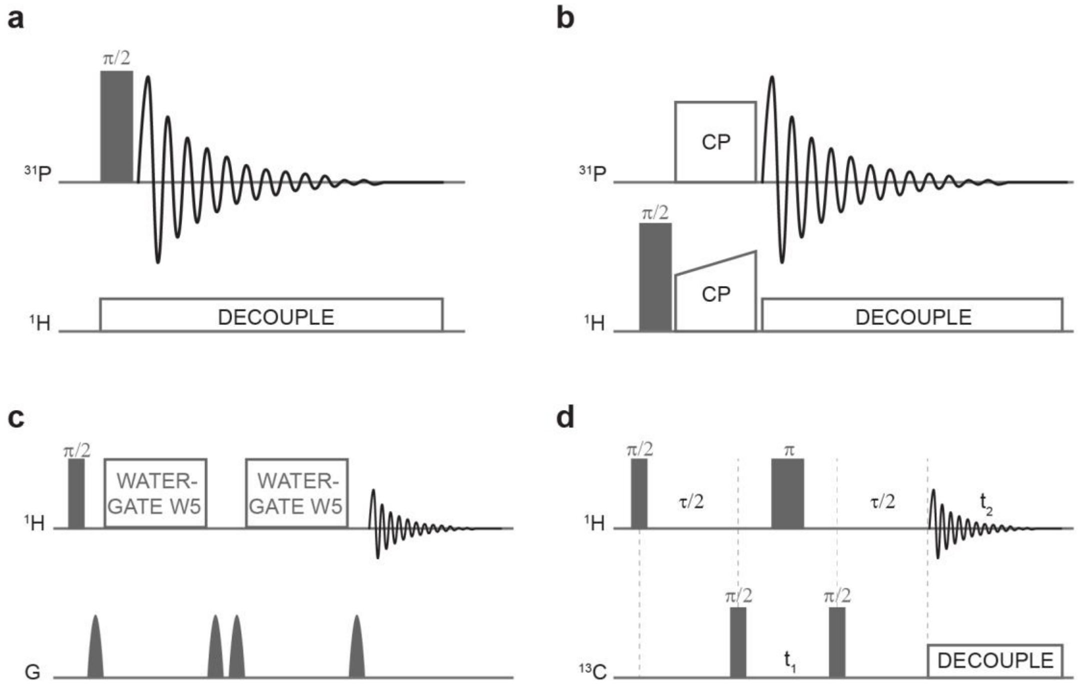


Figure 10. Depiction of simplified pulse sequences used in this work. The following pulse sequences are presented: **a**) 1D ^{31}P direct excitation, **b**) 1D ^{31}P cross-polarisation, **c**) 1D ^{1}H direct excitation with a WATERGATE W5 water suppression scheme, **d**) 2D ^{13}C-^{1}H HMQC used for methyl TROSY.

1.12.2 Cross-polarization

Cross-polarization (CP) is an important building block of countless pulse sequences in solid-state NMR spectroscopy. The technique allows the transfer of magnetisation between two different nuclei through space by virtue of their dipolar coupling. The CP transfer is very efficient for solids and crystals and, in contrast, less efficient in gels and inefficient for freely moving molecules.

Overall, cross-polarization sequences provide three advantages: 1) improving SNR in comparison to direct excitation due to magnetisation transfer from a sensitive (e.g. ^{1}H) to an insensitive (e.g. ^{31}P) spin by a factor of $\gamma_H/\gamma_P \sim 2.5$ [3], 2) acting as a filter to select for immobilised or rigid components and 3) reducing relaxation delay time, because the relaxation time only depends on relaxation rates of the sensitive spins, which is usually fast.

The theory is outlined as follows: both nuclei (^{1}H and ^{31}P) are radiofrequency irradiated with **B_1** field amplitudes set so that the so-called Hartmann-Hahn condition [9] is fulfilled:

$$\gamma_H B_1(^1\mathrm{H}) = \gamma_P B_1(^{31}\mathrm{P})$$

Hereby, the energy gaps between the respective spin states are set to be equal and to "spin baths" are said to be put in contact. During cross-polarization, neither the net energy nor the net spin polarisation are altered. However, during this contact time (typical ^{1}H→^{31}P contact time of 1-5 ms with a dipolar coupling strength of 200-300 Hz in phosphate groups in lipids [10]), energies and spin polarisations between the coupled nuclei can be exchanged. For every flip of a ^{1}H spin, a ^{31}P spin flips. Thus, the total energy is equilibrated between the two spin baths, causing an increase in magnetisation for the insensitive spin. Magnetisation cannot be transferred completely due to this equilibration limit, and some initial magnetisation is lost during the transfer due to ^{1}H spin relaxation, since the applied **B_1** field cannot support the same degree of magnetisation as the larger **B_0** field.

A simple implementation of a CP pulse sequence is shown in Figure 10b. Here, ^{1}H spins are initially excited by a 90° pulse. The magnetisation is transferred from the sensitive ^{1}H to the less sensitive ^{31}P nuclei via cross-polarization. During the contact, constant amplitude is attained at the ^{31}P channel, while the amplitude of the applied **B_1** field is ramped up from 70 to 100% on the ^{1}H channel. The slow ramping helps to fulfil the CP conditions for different spins by varying the **B_1** field and improves robustness towards pulse imperfections. Afterwards, an FID is recorded on the ^{31}P channel, while a decoupling sequence is applied on the proton channel.

1.12.3 ^{1}H 1D

1D proton spectra provide valuable information on both solute and solvent due to the abundance of ^{1}H spins in biomolecular and organic molecules. Additionally, ^{1}H spins are highly sensitive, giving rise to spectra with a high SNR, and have relatively short relaxation times, allowing for rapid data collection. The main disadvantage lies in low chemical shift dispersion, ranging for biological samples between 0 and 15 ppm, leading to a strong signal overlap.

The simplest experiment possible involves a direct excitation of ^{1}H spins with a 90° pulse, followed by immediate acquisition. In solution-state NMR, the most intense signal is often found to be the solvent. Since in biomolecular NMR spectroscopy samples are usually dissolved in aqueous buffers, the most dominant signal arises from the water itself. The broad and intense water signal causes major overlap and significantly reduces SNR by saturating the signal receiver. In this thesis, the WATERGATE W5 [11] pulse sequence (Figure 10c) was employed to suppress the water signal. This WATER suppression by GrAdient Tailored Excitation technique works by applying a 90° pulse to all spins, followed by gradient-based dephasing. All spins except the spins on transmitter resonance (i.e. water) are inverted by a composite pulse train of 180°. This is followed by a second gradient, which further dephases the water signal but re-phases all other signals.

1.12.4 Methyl TROSY

Solution-state NMR is ill-suited for studying larger molecular complexes due to their slow tumbling and corresponding fast T_2 relaxation. Until late 1990s, the molecular size limit for successful NMR bases analysis was considered to be 40-50 kDa, above which the resulting spectra became uninterpretable [12]. This size limit was significantly pushed forward by the development of the transverse relaxation optimised spectroscopy (TROSY), initially developed for the ^{1}H-^{15}N spin system encountered in the amide moieties of proteins [13]. The main principle behind this method is the cross correlation between two separate relaxation mechanisms, namely between the ^{1}H-^{15}N dipolar coupling and CSA of ^{15}N induced T_2 relaxation. This cross correlation creates both fast and slowly relaxing coherences, allowing for the specific selection of the slowly relaxing coherence while suppressing the fast relaxing ones. Although this coherence selection process discards a major part of the initial magnetisation, the resulting spectral quality is superior to that of non-TROSY sequences. Finally, the CSA size but not the dipolar coupling strength scales with the magnetic field. As a result, the completion of the destructive interference is field dependent, and ^{1}H-^{15}N-correlated TROSY spectra give rise to the highest SNR at 900 MHz.

Following this idea, the TROSY principle has been adapted to $^{13}C^{1}H_3$ spin systems [14]. Methyl groups are of particular interest to NMR spectroscopists since they possess multiple advantageous properties: 1) methyl group-bearing amino acids are found in hydrophobic regions of the protein, often in the core, and close to ligand binding regions [15]; 2) methyl resonances are well dispersed and 3) the initial magnetisation arises from three equivalent ^{1}H spins, increasing the amount of detectable signal. Finally, methyl groups are well suited for studying large complexes thanks to their increased rotational freedom, decoupling their movements from the overall slow tumbling rate of the molecular complex.

While the analogy of the amide TROSY principle cannot be applied to a methyl ($^{13}C^{1}H_3$) spin system due to a small CSA of the methyl carbon spin, the idea of the selection of a slowly relaxing coherence still holds. Methyl groups relax predominantly via dipolar interactions between the ^{1}H-^{1}H and ^{1}H-^{13}C spins, which can interfere constructively or destructively, creating fast and slowly relaxing components through cross-correlation. For isolated methyl groups, the standard HMQC sequence is a relaxation-optimised experiment and preserves slowly relaxing coherences throughout the pulse sequence [14]. In the HMQC pulse sequence (Figure 10d), initially, all magnetisation is split into a fast and slow relaxing coherence after the first 90° pulse. No mixing of these two components occurs during the experiment, preserving a large part of the initial magnetisation for the FID recording. For large, slowly tumbling molecules the TROSY enhancement dramatically improves the SNR in comparison with non-TROSY pulse sequences, even though 50% of the original magnetisation is discarded. Additionally, the fast and slow relaxing coherences give rise to signals with the same frequencies, adding to the final signal. This simplifies the pulse sequence since the fast-relaxing coherences do not need to be suppressed. Finally, the methyl TROSY effect is magnetic field-independent since it originates solely from dipolar interactions and not CSA.

The main requirement for methyl TROSY is to have complete deuteration of the protein (aside from the methyl group of interest): additional proton spins provide a ^{1}H-^{1}H dipolar coupled network, which mixes the slowly and fast relaxing coherence pathways and significantly enhances relaxation. Thus, the methyl group needs to be isolated from other spin systems, especially protons, and from other methyl groups. This requires the methyl probes to be diluted and labelling used sparingly throughout the protein.

Site-directed labelling of methyl groups in a perdeuterated background requires specific isotope labelling techniques. Overall, six amino acids carry methyl groups in their side chains: isoleucine (I), leucine (L), valine (V), methionine (M), alanine (A) and threonine (T). While isotope labelling of all these amino acids at the same time (ILVMAT labelling) is possible, it is often neither necessary due to limited additional information gain nor advisable due to peak overlap and loss of signal intensity because of enhanced relaxation [12].

Protein expression in *E.coli* provides an easy way to implement amino acid-specific labelling. The metabolism of the bacterial host has been extensively studied, providing knowledge of pathways to incorporate specifically labelled components. Additionally, *E.coli* are tolerant towards growth in minimal media, in which sources of carbon and nitrogen are precisely controlled, and even growth in a fully deuterated medium is supported. Side chain-specific labelling is then achieved by supplementing the medium with labelled amino acids or metabolic precursors thereof. For the production of ILVM methyl-labelled proteins, U-[^{2}D,^{12}C]-glucose is used as the main carbon source, while the bacteria are grown in D_2O as a solvent. To avoid scrambling into other amino acids, only methionine is added directly as ε_1-[$^{13}CH_3$]-methionine. Isoleucine labelling is achieved by adding α-ketobutyrate to the medium, while both leucine and valine are converted by *E.coli* from the same precursor, 2-acetolactate [12].

Chapter 2: ATPase cycle of p97

2.1 Evolutionary hierarchy of p97

Phosphate-loop (P-loop) NTPases are a class of proteins that process nucleotide triphosphates (NTPs) to generate energy for their cellular functions [16]. These enzymes share characteristic features, including an alternating α-helical and β-strand backbone (α/β) structure and conserved Walker A and B motifs responsible for nucleotide binding and hydrolysis, respectively (see below) [17]. P-loop NTPases are divided into two major subclasses: the kinase-GTPase and ASCE (additional strand, catalytic/conserved E) subclasses, out of which only the latter is relevant for this work [18].

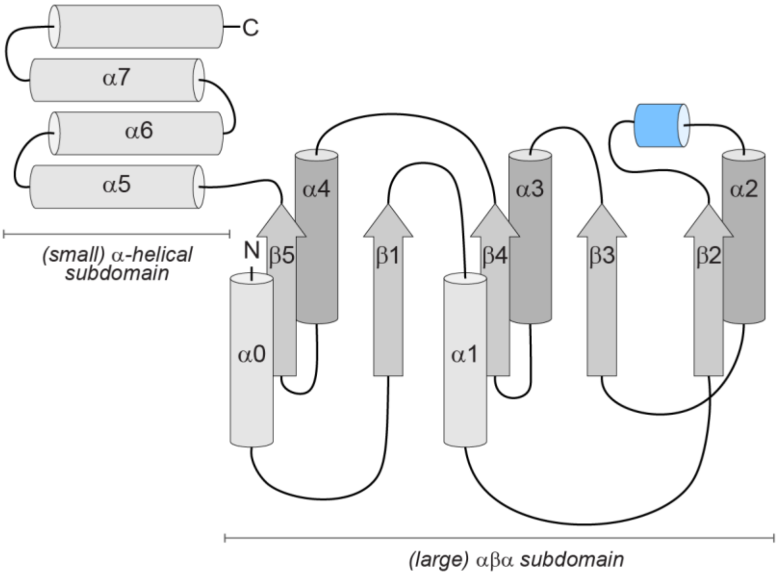


Figure 11. Schematic overview of the Classic clade of the AAA+ proteins. AAA+ proteins are built from from a large αβα "sandwich" subdomain and a small α-helical subdomain. The Classic clade (Clade 3) is defined by a short α-helical insertion between the β2 sheet and α2 helix (highlighted in blue). Figure adapted from ref. [19].

A major part of the ASCE subclass is the AAA+ ATPase Superfamily (extended family of **A**TPases **a**ssociated with various cellular **a**ctivities), which is a group of functionally diverse, essential enzymes involved in various cellular functions [20]. AAA+ proteins distinguish themselves from other NTPases by multiple insertions, including the α0 helix and the

C-terminal helical bundle (Figure 11) [20]. The main function of these molecular machines is to couple adenosine triphosphate (ATP) to adenosine diphosphate (ADP) hydrolysis with conformational changes, thus remodelling respective substrates. Considering the role of ATP as “fuel”, AAA+ proteins are occasionally referred to as molecular “motors”, ready to convert chemical into mechanical energy. AAA+ ATPases have an overall conserved architecture of their ATPase domains, built from an αβα “sandwich” topology and form oligomeric, ring- or spiral-shaped assemblies [20]. Moreover, they share several key structural motifs (which are discussed in detail below), leading to a conserved molecular mechanism for ATP hydrolysis [20].

Based on unique additional structural elements, the AAA+ ATPases are further subdivided into seven clades. Defined by an inserted short linker and an α-helix between the β2 sheet and α2 helix, the Classic clade (Clade 3) is the most diverse group of AAA+ proteins in terms of functionality (Figure 11) [18,20]. Enzymes of this clade are further divided into type I and type II ATPases, dependent on the number of ATPase domains in a single monomer. At least three distinct fusion events are responsible for the emergence of the second ATPase domain [21]. Specifically, in the larger Cdc48 family (incl. NSF, Cdc48/p97, Pex), proteins are built from a substrate recognising N-terminal domain and two homologous ATPase domains, the tandem domain structure resulting from a duplication event [21,22]. The work performed in this thesis focuses on one member of the larger Cdc48 family.

2.2 Biological function of p97

The transitional endoplasmatic reticulum ATPase (TER ATPase) is a molecular chaperone and a prominent member of the overall AAA+ superfamily. The associated gene VCP (Valosin-containing protein) is evolutionarily conserved, and homologues are found in all kingdoms of life [23]. In mammals, the enzyme is referred to as p97, in yeast as Cdc48, in worms as CDC48, VAT in archaea and TER94 in flies, but other names related to “VCP” can also be found in the literature [23,24]. In this thesis, all investigations are performed on the human homologue p97, which is the reason why, in the following, the enzyme is only referred to as such.

The ATPase p97 is one of the most abundant cellular proteins with a copy number of $1.38 \cdot 10^6$ or roughly 1% of total protein in human cancer cells [25,26]. Being a highly soluble protein, p97 is mostly localised in the cytosol, but a fraction can also be found in subcellular organelles, including the nucleus, mitochondria, the endoplasmatic reticulum (ER), Golgi apparatus and endosomes [27]. It is involved in multiple cellular processes, ranging from protein homeostasis and degradation to organelle assembly and DNA repair (Figure 12) [28]. The ATPase fulfils this multitude of functions by its ability to bind a variety of adaptor proteins, supporting interactions with even more substrates. Adaptor proteins, also called co-factors, contain conserved p97-binding domains such as UBX (Ubiquitin regulatory X), VIM (VCP-interacting motif) or SHP (short heterodimer partner) and mediate subsequent substrate binding (reviewed in ref. [28]). In these interactions, p97 works as a “segregase” or “unfoldase”, separating substrates from macromolecular assemblies and/or unfolding them for further processing. Most prominently, p97 is part of the ER-associated degradation (ERAD) system, retrieving misfolded proteins from the ER membrane [27]. Typical substrates of p97 are tagged by ubiquitin, marking them for downstream degradation by the proteasome system.

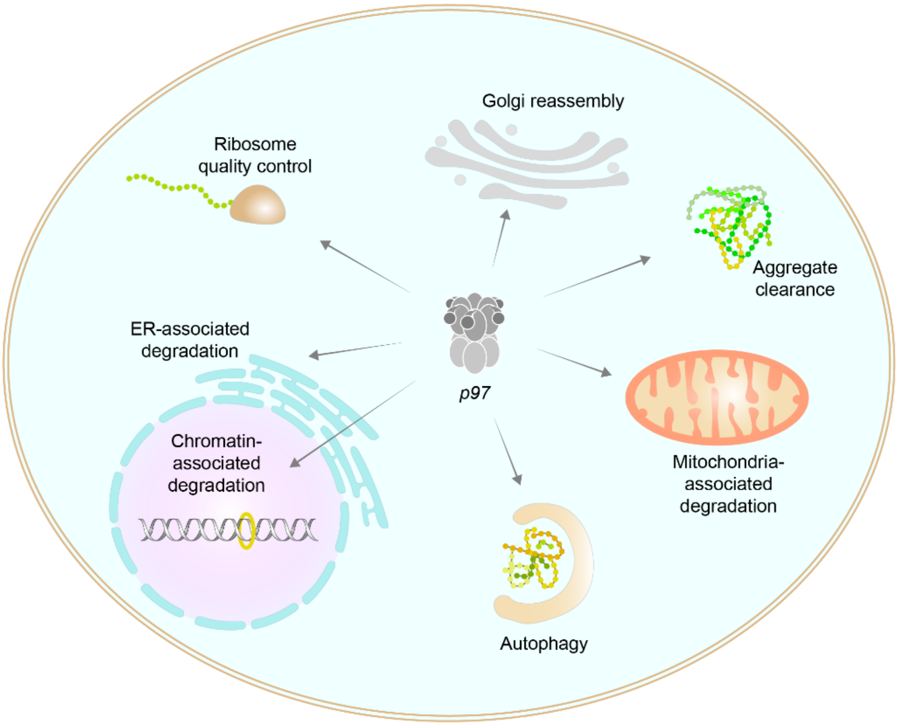


Figure 12. Overview of cellular pathways which p97 is involved in.

Considering the number of cellular pathways dependent on p97 activity, unsurprisingly, it is an essential protein. Defects in Cdc48 expressions in yeast create diverse effects, starting with cell cycle arrest and impaired ER membrane fusion and ending with aggregation of mitochondria and dysfunctional proteolysis [26]. In humans, several neurodegenerative diseases have been linked to mutations in p97, leading to a wide range of phenotypes [26]. Commonly, mutations in p97 have been associated with multisystem proteinopathy 1 (MSP1), but other conditions such as amyotrophic lateral sclerosis (ALS) and Parkinson's disease have been found to be instigated by defects in the VCP gene [26]. Moreover, mutations of p97 can affect pathologies of other diseases, such as Alzheimer's and Huntington's diseases [26]. The upregulation of p97 expression in cancer cells puts the ATPase into the spotlight as a potential anti-cancer target [29,30]. Consequently, several p97-inhibitors have been developed and structurally analysed (e.g. [31–33]).

Considering this biological and pathological relevance, p97 became a popular target of structural investigations. Additionally, p97 can function as a model AAA+ ATPase, and understanding its molecular mechanisms can shed light on the underlying mechanistic principles of many more proteins.

2.3 Domain organisation of p97

Despite its misleading name in reference to an originally reported precursor molecular weight of 97 kDa [34], (human) full-length p97 is comprised of 806 amino acids with a total molecular weight of 89 kDa in its monomeric form. Its classification as an AAA+ ATPase informs on further properties: p97 forms symmetric, planar ring-shaped hexameric assemblies with a total weight of 535 kDa with the nucleotide binding pockets located between neighbouring protomers (Figure 13). Moreover, as a tandem/type II ATPase, p97 is built from two homologous ATP processing domains, termed D1 and D2, connected by a 20 residues long D1-D2 linker. In total, p97 hexamers possess twelve nucleotide binding pockets, though not all binding sites might be occupied at the same time, depending on the nucleotide type [35].

A third domain, the N-terminal domain (NTD), is connected to the D1 domain via another 20 amino acids long flexible N-D1 linker and can be located either above ("up conformation") or coplanar with ("down conformation") the D1 domain (Figures 13 and 15). Finally, the D2 domain is capped by a short disordered C-terminal tail. The C-terminal tail functions as a binding partner to adaptor proteins bearing the p97-interacting PUB (peptide:*N*-glycanase and UBA or UBX-containing proteins) or PUL (PLAP, Ufd3p, and Lub1p) domains, although the majority of co-factors are known to bind the N-terminal domain [28]. Binding of adaptor proteins or substrate processing through the central pore also leads to symmetry breaking or adoption of other oligomeric states [36–38].

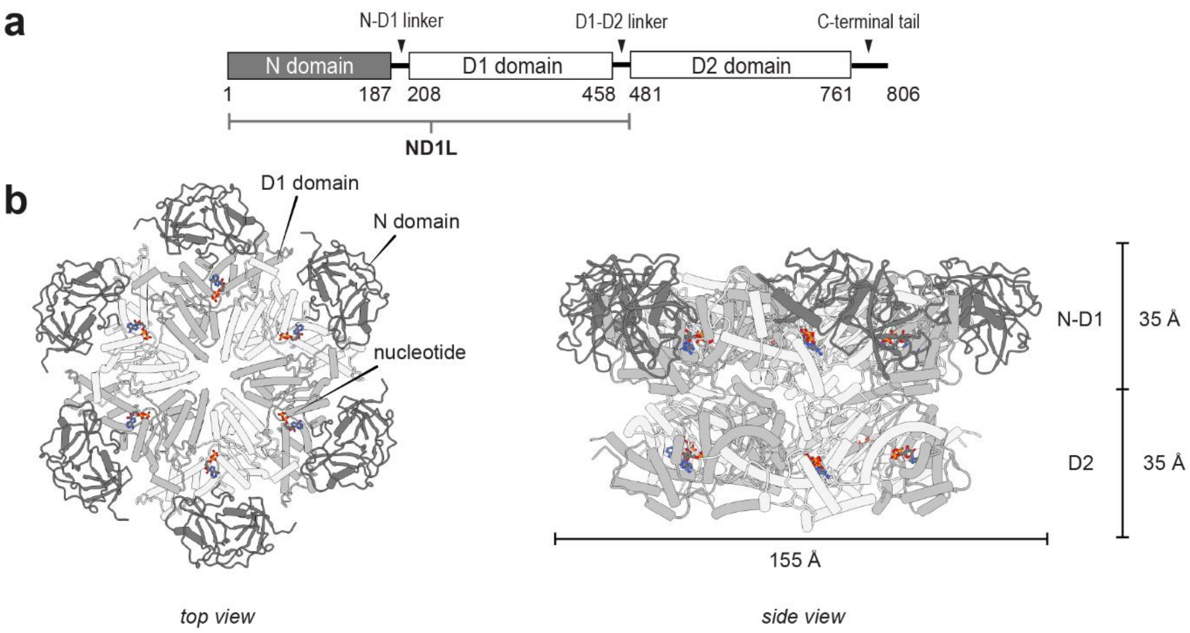


Figure 13. Domain organisation of p97. a) The ATPase is organised into three domains: N-terminal domain, D1 and D2 which are connected via flexible linkers. The ND1L construct is built from the NTD, D1 and the D1-D2 linker. **b**) Monomers assemble into ring-like hexamers with a central pore. Neighbouring monomers are shown in hues of grey. The nucleotide binding pocket is located at the interface of two neighbouring protomers. The two ATPase domains form a double ring structure. Top and side views of the full-length ADP-bound p97 are shown (PDB ID: 5ftk).

While both ATPase domains contribute to the total ATPase rate [39], they serve distinct functions. The D1 domain alone is sufficient for hexamerisation [40] and acts as an anchor for the up-down movement of the NTD, governing its positioning based on the nucleotide occupation [32]. Since the adaptor binding to the N-terminal domain is strongly dependent on the NTD conformation, the D1 nucleotide state is assumed to additionally regulate adaptor interactions [28]. On the other hand, the D2 domain is considered to be the driving force of substrate unfolding through pore opening and rotational movement between the two rings [32,36,41]. Additionally, nucleotide binding affinities of the D2 domain are lower [35], but nucleotide occupation in the D2 binding pocket has been reported to control dodecamer formation: a double hexamer tail-to-tail organisation of p97 mediated by the C-terminus [42]. In a recent report contrasting previous findings, double hexamers have been found to be ATPase and unfoldase active, raising the question of the purpose of the dodecamer formation [33,43,44]. Communication between the two domains is mediated by the D1-D2 linker and plays a critical role in D1 activity [45,46]. The N-D1 construct, comprised of the NTD and the D1 domain

alone, does not show any significant ATPase activity, but the ND1L construct (NTD + D1 + D1-D2-linker) has a similar ATPase rate as the full-length (FL) protein [47].
In this thesis, the focus lies on the ND1L construct. Comprising 480 residues and with a molecular weight of 53 kDa per monomer, the construct fulfils the basic functions of the full-length enzyme: it forms hexamers, it is ATPase active, and the NTD responds to nucleotide binding. Additionally, most known disease-related mutants affect the NTD-D1 interface [26,27]. Thus, the truncated construct is an effective, simplified tool to study p97.

2.4 Secondary structure elements of the D1 domain

As a member of the Classic Clade of the AAA+ ATPases, p97 shares multiple structural properties with the other members of the Clade. Both D domains can be further divided into two subdomains: a larger N-terminal α/β-structured RecA-like domain and a smaller C-terminal α-helical domain, with the nucleotide binding pocket located between the two subdomains, adjacent to the clockwise neighbouring protomer (Figure 14b). Moreover, both domains share several conserved motifs known to fulfil specific functions. Due to the nucleotide being bound between two protomers, some of the motifs contribute in *cis* (i.e. from the main monomer), while others are *trans*-acting (i.e. from the neighbouring monomer).

The Walker A motif, also known as P-loop due to its loop structure, has a conserved sequence of [G/A]XXXXGK[T/S] (with X being any residue) and is associated with the structuring of the nucleotide binding pocket [17,48]. The heavily conserved lysine is involved in Mg^{2+} and ATP binding, and mutations of the Walker A motif are known to lead to nucleotide binding deficiencies [16,35,41]. In the D1 domain of p97, the motif takes the form of 245-GPPGTGKT-252 (Figure 14a).

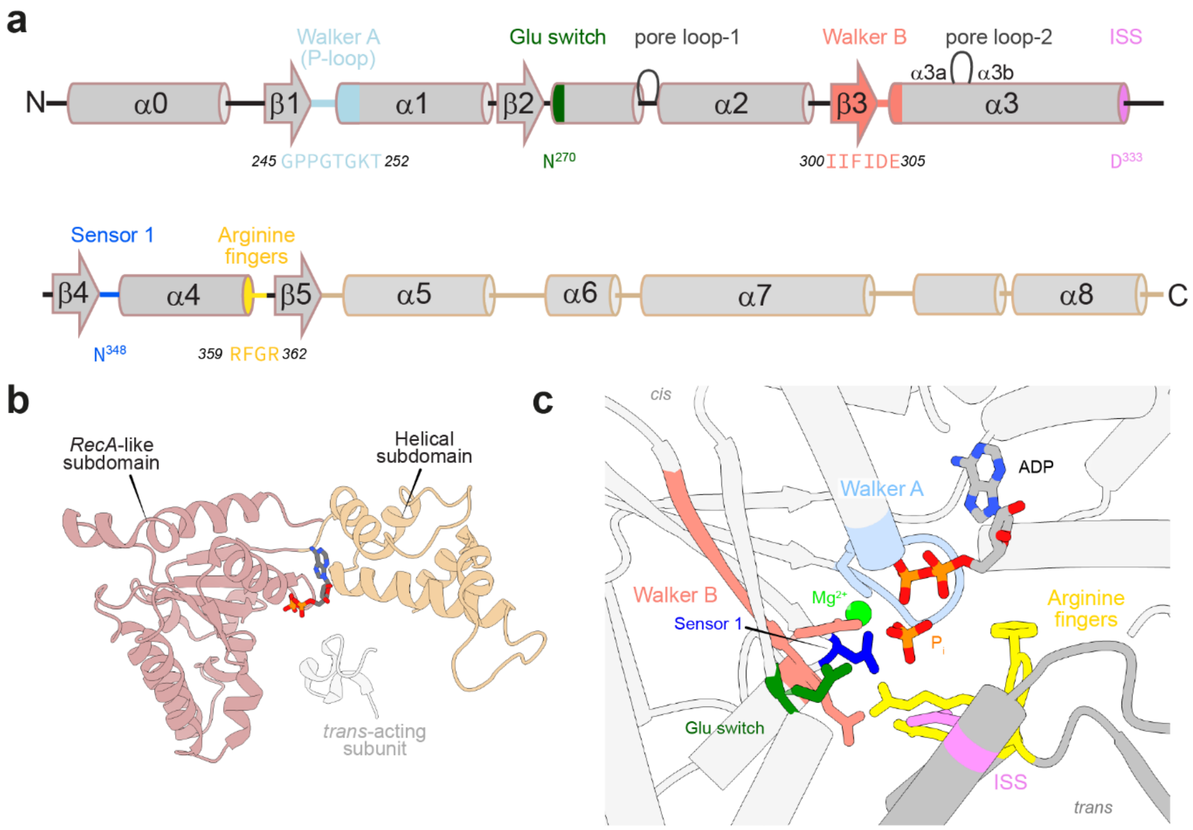


Figure 14. Subdomain organisation and conserved motifs of the D1 domain. **a**) Secondary structure organisation of the D1 domain. The structural element numbering is adapted from ref. [49]. Conserved motifs are highlighted in colour, and their sequences are shown below. **b**) The D1 domain can be subdivided into an N-terminal α/β RecA-like domain (brown) and an α-helical C-terminal domain (tan) with the nucleotide binding interface located between the two subdomains. In white, elements of the trans-acting protomer are shown. The colouring of the subdomains corresponds to the stroke colours of the secondary structure elements shown in **a**). **c**) Zoom in on conserved motifs surrounding the bound nucleotide. The motifs are coloured identically to the colours shown in **a**). Only the ISS residue and the arginine fingers act in trans. All the other motifs are found in the *cis*-acting subunit. PDB IDs: 5ftk, 8ooi.

The Walker B motif of the form hhhhD(D/E), where h is a hydrophobic amino acid, is involved in ATP hydrolysis by priming a water molecule for a nucleophilic attack on the terminal phosphate of ATP [50,51]. The conserved aspartate interacts with the bound nucleotide either via direct bonding to the Mg^{2+} ion or through hydrogen-bonded water [52–54]. Mutations of the Walker B motif have been associated with a deficiency in ATP hydrolysis, which is the reason why glutamate mutations have been often employed in structural studies of p97 [40,44,45,47,55,56]. Additionally, the glutamate residue supposedly acts as a "glutamate switch", regulating the transition between the active and inactive state via interactions with

N270 (in p97-D1) [57]. This proposed interaction, however, has not been observed for p97 [41]. In the D1 domain of p97, the Walker B motif takes the form of 300-IIFIDE-305 (Figure 14a).

Originally, the sensor-1 motif (N348 in D1) was assumed to *sense* the nucleotide and to detect its hydrolysis via hydrogen bond formation with the γ-phosphate [54,58,59]. Nowadays it is known that the sensor residue plays an active part in hydrolysis itself via coordination of the attacking water [41,60]. The naming of this motif also implies the existence of the sensor-2 motif, which is, however, absent in the AAA subfamily members of the AAA+ family [61]. Instead, the sensor-2 arginine has been replaced by a second *trans*-acting arginine finger (see below) [61].

Another conserved region between the β4 and β5 strands is termed the second region of homology (SRH) and is responsible for interprotomer communication [19,39]. This designation is used less frequently nowadays since the region includes two separate motifs: the sensor-1 residue and the Arginine or R-fingers. The latter is a term for *trans*-acting arginines, which interact with the nucleotide binding pocket of the neighbouring protomer, sensing the neighbouring nucleotides [19,62]. Additionally, Arg-fingers interact with the γ-phosphate of ATP and stabilise accumulated negative charges, thus supporting hydrolysis [50]. Instead of the sensor-2 arginine, p97 contains two arginine fingers, R359 and R362, separated by two residues. Typically, these two residues are proline and glycine, but p97 constitutes an exception from other AAA+ family members by replacing the proline residue with phenylalanine in D1. The role of the phenylalanine side chain is to act as a nucleotide sensor through rotamer flipping [41,63], and its side chain conformation is coupled to the motion of the arginine finger R359 [52].

Another communication channel in which the arginine fingers are involved is intersubunit signalling (ISS, Figure 14). The role of this motif, characterised by a D/E residue at the C-terminus of the α3-helix (D333 in p97-D1), is to sense the nucleotide state of the neighbouring protomer [62]. The network is built from interactions between the ISS residue and the arginine finger (D333 and R362), which in turn interacts with the glutamate residue of the Walker B motif (E305) [41].

The ISS is also tightly connected to a last set of conserved motifs, the pore loops. Being a special structural feature of the Classic Clade, some AAA+ ATPases contain insertions that are responsible for substrate interaction and threading through the central pore, which are accordingly named pore loops (Figure 14) [20,50]. Typically, these motifs are characterised by aromatic residues and are associated with substrate translocation (pore-1 loops) and substrate binding (pore-2 loops) [50]. In p97, the D1 pore-1 loop lacks aromatic residues, leading to weakened loop-substrate interactions in comparison to the D2 domain [36]. The structural conformation of the pore loops is coupled to the structural features of the ISS network, thus connecting the substrate translocation to nucleotide sensing [41].

2.5 ATPase cycle of p97

Over the course of the ATP hydrolysis cycle, p97-ND1L takes on four distinct nucleotide and protein states. A cartoon representation of the ATPase cycle considering only the D1 domain is shown in Figure 15a,b: the D1 domains form symmetric, hexameric ring structures, while the NTDs (bound to the N-terminus of the D1 domain) can switch positions between coplanar and above the D1 plane.

Full-length structures of uniformly bound nucleotides are presented in Figure 15c. Since the sequence of ATP hydrolysis events among the D1 and D2 domains is contested and no accepted hydrolysis model is established, a corresponding cartoon ATPase cycle for full-length p97 is omitted.

In the starting nucleotide-free state (apo), the N-terminal domain is located above the D1 plane, but it is highly flexible, leading to poor local resolutions in cryogenic electron microscopy (cryo-EM) studies [42]. Upon ATP(γS) binding, the NTD remains above the D1 plane but adopts a more rigid and structured conformation [32,42]. Hydrolysis of ATP leads to the formation of the post-hydrolysis ADP·P_i state and is coupled to an NTD power stroke downwards. The protein conformation of this state is characterised by a close similarity with the ADP-bound state. Further details are discussed in the next section.

Subsequent release of inorganic phosphate is the rate-determining step of the full ATPase cycle and leads to the creation of the ADP state [55], prominent for its locked NTD down position [44].

Finally, after slow ADP release, the nucleotide binding pocket is conformationally primed for the apo state, releasing the N-terminal domain from its locked position and enabling ATP binding, thus completing the cycle [42,47].

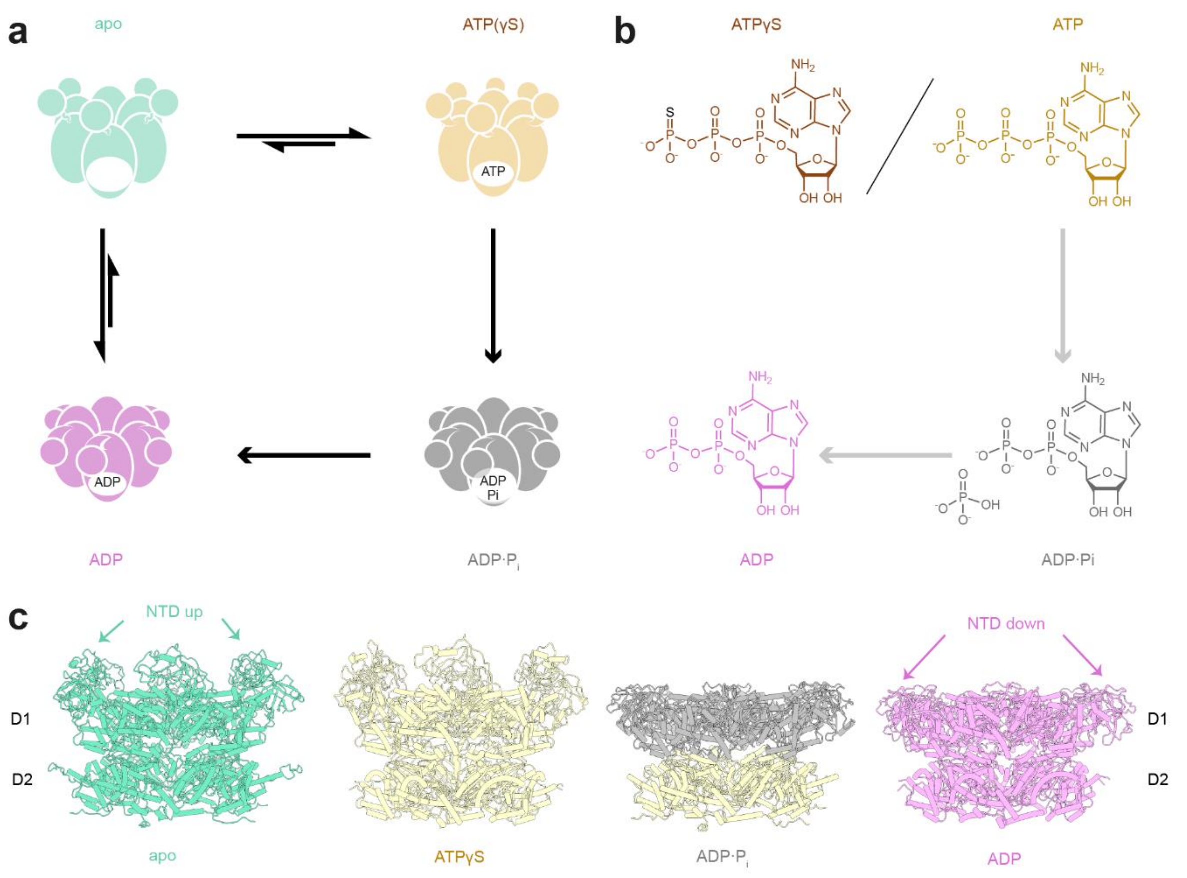


Figure 15. Overview of the p97 ATPase cycle. a) The ND1L ATPase cycle consists of four distinct states. In the apo and ATP(γS) bound states, the NTD is located above the D1 plane (“up”). After hydrolysis to ADP·P_i, the NTD performs a power stroke into the D1-coplaner “down” position. **b**) Structural formula of nucleotides corresponding to the different ATPase cycle states. **c**) Full-length p97 structures in apo, ATP-, ADP·P_i- and ADP bound states. PDB IDs: 7k59, 5ftn, 8ooi, 5ftk.

All four states can be caught by NMR measurements (Figure 16). Methyl probes spread throughout the protein act as conformoscopes, providing information on D1 domain structure, bound nucleotide type (e.g. I206 and L286) and the NTD position (e.g. V87, V116 and V154). Owing to the fact that p97 is involved in heat stress response [26,64], the hexameric assemblies have high thermal stability [65] and allow for NMR measurements of nucleotide bound p97 at 50°C. Only the apo state, more prone to monomerisation, is less stable and needs to be measured at 37°C. For static nucleotide states, ATPγS and ADP are provided under saturating conditions to enforce the complete formation of pre- and post-hydrolysis protein states, respectively [35,66]. In order to record the ADP·P_i state, p97 needs to be measured under constant ATP turnover, i.e. in the presence of an ATP regeneration system [55]. A comparison between the ADP·P_i state and ATPγS and ADP states is shown in Figure 17.

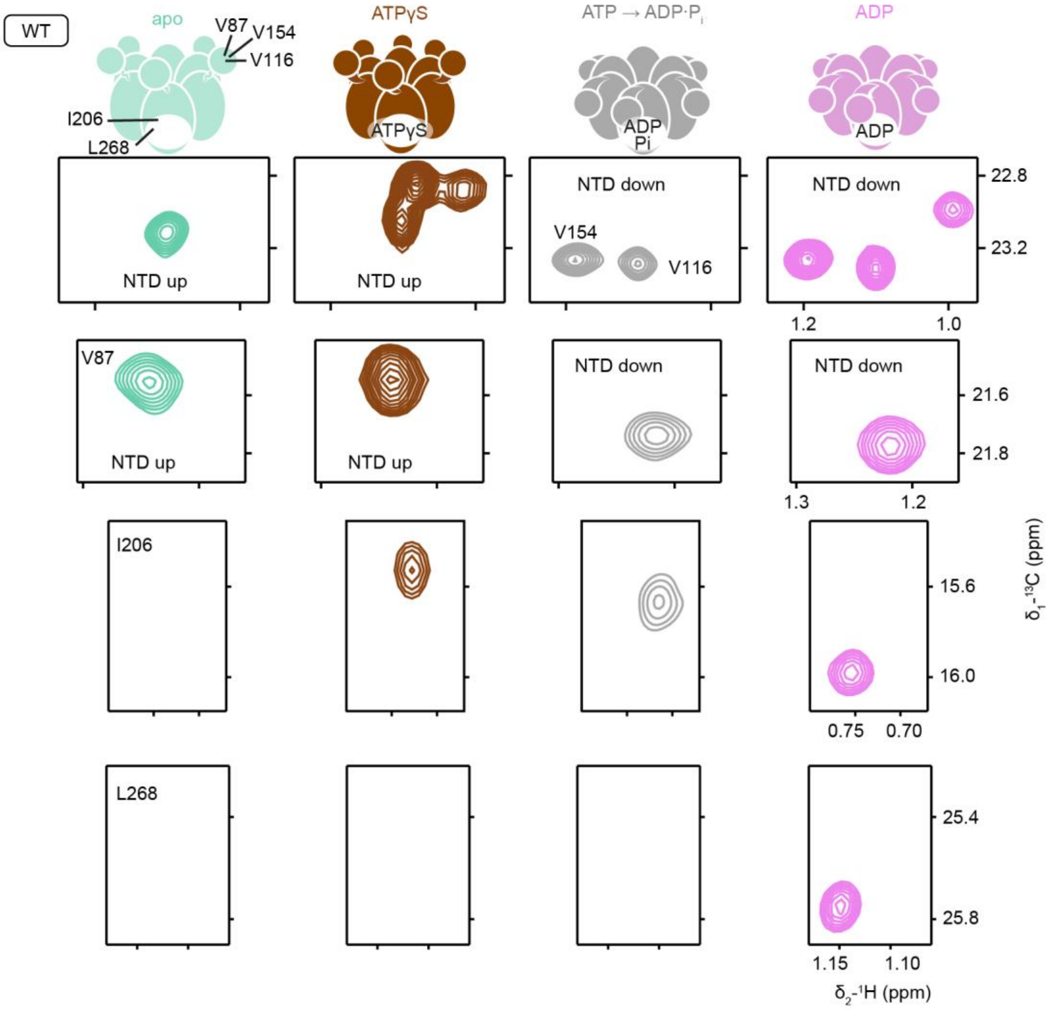


Figure 16. Behaviour of ND1L-wt in different nucleotide states. Methyl groups act as conformational probes in NMR measurements. ND1L-wt was recorded in distinct nucleotide states. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. Residues V154, V116 and V87 report on the NTD position, while residues I206 and L268 inform on the type of nucleotide bound in the binding pocket. In the apo and ATP(γS) bound states, the NTD is located above the D1 plane ("up"). After hydrolysis, the NTD performs a power stroke into the D1-coplaner "down" position.

Chemical shifts are highly sensitive to their environment and structural changes. In NMR, Chemical shift perturbations (CSP) are commonly analysed to map ligand binding sites and assess the strength of given interaction. Deviations in the peak position in one or more spectral dimensions from the reference point due to environmental changes (e.g. ligand binding) are referred to as CSPs. In the context of p97, in this thesis, nucleotide binding induced conformational change causes CSPs in the methyl probes, leading to different peak positions in different protein states.

2.6 The elusive ADP·P$_i$ state

In most studies discussing possible ATP hydrolysis models for p97/Cdc48, the existence of the ADP·P$_i$ state is neglected, and the ATP-bound protein state is directly followed by the ADP-bound one [24,54,63,67]. The issue is encountered in discussions of hydrolysis models of AAA+ ATPases in general, even when several mechanistic models are compared and the existence of a post-hydrolysis, pre-phosphate-release state is implied [20,22,50]. This observation is somewhat surprising, considering that the ADP·P$_i$ has been postulated as a reaction intermediate as early as 1970 [68]. Two arguments can be made to explain this discrepancy: 1) inorganic phosphate, as the hydrolysis product, is assumed to leave the binding pocket directly after hydrolysis, and 2) it is difficult to experimentally verify the presence of the cleaved phosphate ion and to resolve it in structural studies.

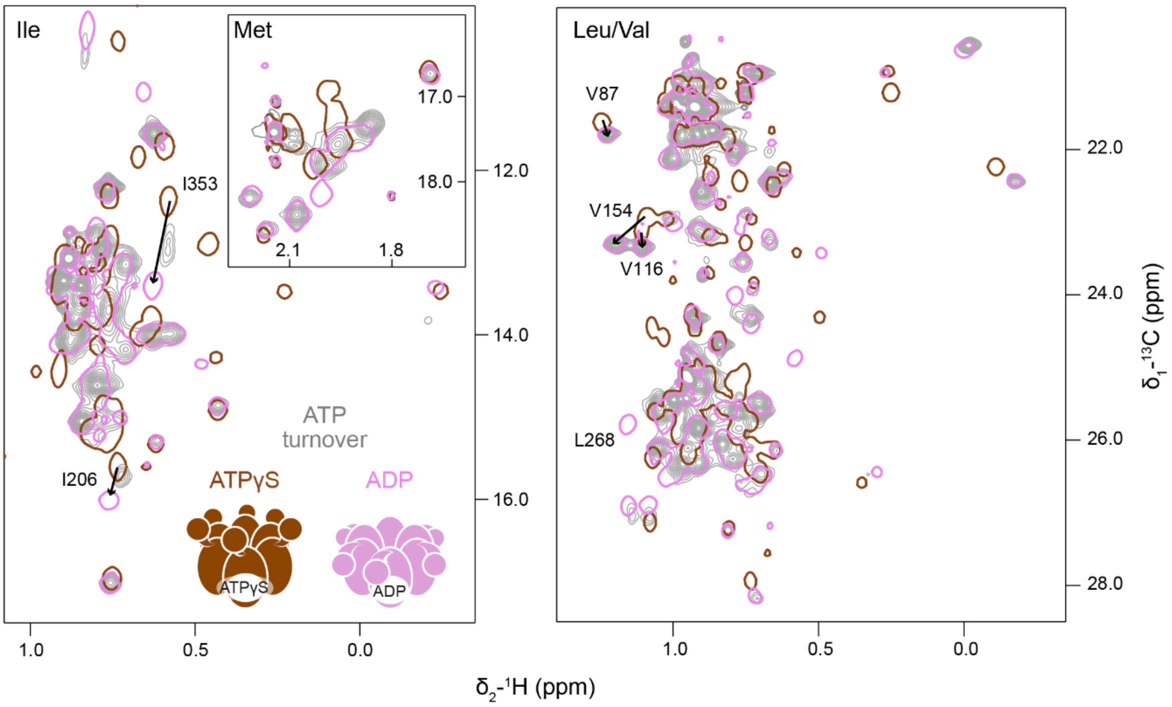


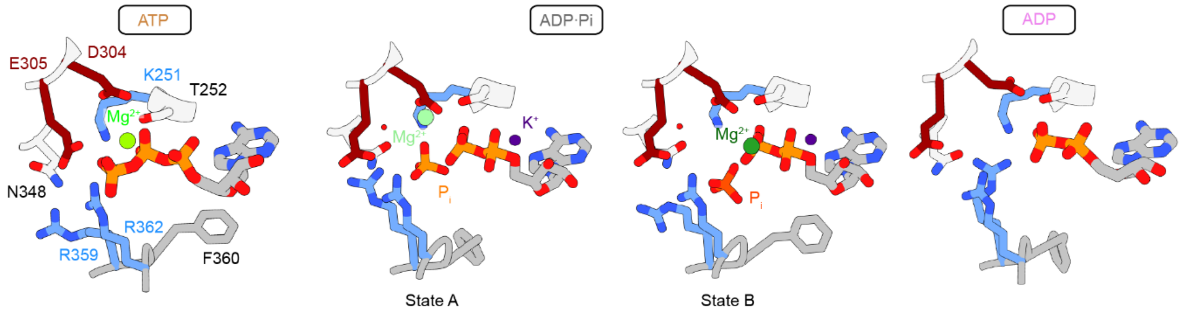
Figure 17. Methyl TROSY spectra of ND1L-wt under ATP turnover. Full spectral regions of isoleucines, methionines and valines and leucines are shown. Spectra of the ADP·P$_i$ state, measured under ATP turnover (grey) show high similarity with the ADP bound state (pink). Reference spectra of static nucleotide states (ATPγS: brown, ADP: pink) are shown in single contours.

Such arguments make sense, considering the wealth of published p97 structures with bound ATP and ADP or nucleotide analogous but never reporting an additional bound phosphate [24].

However, this does not necessarily mean that the ADP·P_i state has never been observed: the inorganic phosphate is a small molecule, and its identification requires a high-resolution structure. Moreover, its location(s) inside the binding pocket and the distance to ADP are unknown in advance, which is additionally aggravated by fluctuations and dynamic behaviour, rendering an unambiguous identification tedious.

Still, the disregard for the ADP·P_i state is surprising, considering its proven existence in several cases: Indeed, multiple ADP·P_i structures have been previously published, including Hsc70 [69], F-actin [70], myosin [71], adenylate kinase [72] and F_1-ATPase [73]. The resolution of these structures ranged from 2.0 Å to 3.4 Å. All these structures have in common that they were not produced during active nucleotide turnover and thus might not represent authentic post-hydrolysis states. Additionally, none of the mentioned proteins belong to AAA+ ATPases. Moreover, published structures with bound inorganic phosphate or the phosphate mimic sulphate alone (without additional nucleotide) prove that phosphate ions binding can lead to stable protein states [41,74].

Previously, the detection of a post-hydrolysis state of p97 has been reported, which differs from the regular ADP-bound state (Figure 17) [55]. This state could be measured by solution-state NMR only in the presence of an ATP regeneration system, i.e. under constant ATP turnover and in the absence of ADP. However, the location of the bound inorganic phosphate and general binding modalities could not be determined from NMR experiments alone.



State A

State B

Figure 18. Comparison of binding pockets in different nucleotide states. Zoom-in on the p97-D1 nucleotide binding pocket in different nucleotide states. Two states have been determined for the ADP·P_i state, termed State A and B. They differ from each other in terms of Mg^{2+} and P_i locations and in rotamer positions of *trans*-acting R359 and F360. PDB IDs ATP: 7lmy, ADP·P_i: 8ooi, ADP: 5ftk. This figure was previously published in ref. [52].

In a follow-up study, by combining Molecular Dynamics (MD) simulations with a high resolution (2.6 Å) cryo-EM map, the binding locations of Mg^{2+} and P_i could be established

(Figure 18). In this study, two ADP·P_i states (termed State A and State B) were observed, which differ from each other in relation to Mg^{2+} and P_i positions and in sidechain rotamers of *trans*-acting R359 and F360. Indeed, the two states could potentially represent a natural transition from the immediate post-hydrolysis ADP·P_i state to the ADP state via phosphate release.

2.7 The choice of the model system

As a tandem AAA+ ATPase, p97 tightly regulates its ATPase activity via interprotomer and interdomain coordination. Although p97 is generally described as a homo-hexamer, several structures have been published displaying different assemblies of p97 with lack of symmetries upon binding of a co-factor and substrate [36,37,75]. Neighbouring nucleotide binding pockets of these asymmetric structures can differ in the type of bound nucleotide [36], reminiscent of the staircase models found for many AAA+ ATPases [50]. ATP hydrolysis is generally assumed to propagate around the hexamer counter-clockwise, requiring a tight interprotomer regulation, but no single mechanistic model of hexamer engagement has yet been proven for p97 [22].
Still, other structures of co-factor bound p97 provide hexameric symmetries [76,77], continuing the debate of whether the hexameric form of p97 is biologically relevant or not. In its co-factor free form, wild-type p97 forms homohexameric ring assemblies, which change conformations depending on the nucleotide state [32]. These hexamers are able to continuously hydrolyse ATP without any associated purpose or related "work", i.e. idle. The signalling pathways of this coordinated ATP hydrolysis are not expected to differ fundamentally from the ones during substrate processing, given how fine-tuned ATP processing appears to be. This means that the study of the idle ATP hydrolysis by p97 can yield relevant biological information regarding its modus operandi during its segregase and unfoldase activities or during regulatory "hibernation" periods [42].
In this thesis, we study the simplified system of p97, lacking any translocating substrate or bound adaptor and co-factor proteins. Instead, the focus lies on nucleotide processing under substrate-free ("idle") conditions and structural features of the nucleotide binding pocket.

Aim of the Thesis

The aim of this thesis is to study the elusive ADP·P_i state, which has been recently observed *in situ* as a post-hydrolysis state of p97. Using solution and solid-state NMR in combination with biochemical and –physical tools, the ADP·P_i state of p97 will be analysed with a focus on the following questions:

-What are the general characteristics of the ADP·P_i state? Is it symmetric in all nucleotide binding pockets? Is it dynamic or static? How does its protein conformation differ from the other states of the ATPase cycle? What type of chemical environment does the ADP·P_i nucleotide experience in comparison to the other nucleotide states?

-Which residues contribute to the longevity of the state? What roles do they perform in overall the ATPase cycle?

-How is the inorganic phosphate released from the nucleotide binding pocket after hydrolysis? What conformational changes occur upon product release?

-How do individual protomers coordinate ATP hydrolysis among each other?

In order to answer these questions, the ADP·P_i state will be investigated both from the side of the protein and from the side of the nucleotide.

In Chapter 3, the nucleotide side will be considered. Using solid-state ^{31}P NMR, nucleotides bound to the protein in ATP and ADP states as well as under turnover of ATP will be probed, analysed and compared.

Then, in Chapter 4, several ATP analogues will be tested for biological relevance and practical usefulness. Additionally, gathered spectra will be used to support chemical shift assignments of native nucleotides.

Finally, in Chapter 5, the protein side will be investigated. Starting from point mutants cloned via site-directed mutagenesis, methyl TROSY NMR experiments of different nucleotide states in combination with biophysical and –chemical techniques will be used to characterise the constructs. This will allow us to analyse potential mutant defects and determine the functions of the respective residues. Finally, kinetic experiments with the ATP analogue ATPγS will be employed to study interprotomer cooperativity.

Chapter 3: ^{31}P NMR investigations of p97 nucleotide states

3.1 Introduction

^{31}P is an NMR-active nucleus with a spin quantum number of ½ and 100% natural abundance. In combination with its relatively high gyromagnetic ratio (~40% of ^{1}H) and a wide chemical shift dispersion (from -180 ppm to 250 ppm), this nucleus is attractive for NMR spectroscopists. Similar to ^{13}C, NMR experiments with ^{31}P suffer from long T_1 relaxation times (e.g. 3-30 s in organophosphates [4]), but the overall receptivity (i.e. the ease of detection by NMR, considering spin properties and natural abundance) of the phosphorus nuclei is significantly stronger than that of ^{13}C spins by a factor of ~390 [3].

Moreover, the natural presence of phosphorus in nucleic acid phosphate backbones, lipids and nucleotides raises interest in this element for biomolecular NMR spectroscopists. For DNA and RNA structural calculations, ^{31}P NMR can provide valuable information about phosphodiester bond angles since these strongly affect *J*-couplings and chemical shifts [78]. On the other hand, phosphoryl transfer reactions are principal to cellular metabolism, making their mechanistic understanding a crucial necessity for biochemical and medicinal applications [79,80]. Mechanistically, these reactions can follow a multitude of reaction pathways [79,80]. However, there is reason to believe that the catalytic approach of enzymes to ATP hydrolysis might be similar [81].

^{31}P NMR is a useful to answer mechanistic questions concerning phosphoryl transfer. By retrieving spectroscopic observables of the target molecule (chemical shifts, couplings), ^{31}P NMR can contribute to mechanistic analysis.

For p97, the exact mechanism of ATP hydrolysis is still being investigated. While the intermediate nucleotide state after ATP hydrolysis, the ADP·P_i state, was postulated in 1970 [68], capturing it remained a challenge. Although several ADP·P_i state enzyme structures have been published, these ADP·P_i states were achieved through enzyme incubation with exogenous phosphate [70,73] or a phosphate-containing species [72] or through lengthy incubation with ATP [69,71,82]. For an uncertain reason, the capture of a native, post-hydrolysis ADP·P_i state of p97 remained impossible, raising suspicions regarding its relevance in the ATPase cycle of AAA+ proteins in the first place.

Due to the transient nature of ADP·P_i state, standard methods of structural biology such as X-ray crystallography and cryo-EM can be ill-suited to capture it: conformational changes during the ATP hydrolysis cycle interfere with crystal growth, while cryo-EM has difficulties resolving short-lived and/or dynamic conformations. Still, even though NMR can, in principle, detect and analyse low-populated conformations, no previous study of this nature has been reported.

Naturally, for the ADP·P_i state to be regarded as such, both the nucleotide identity and a distinct protein conformation (different from other states) need to be confirmed. Previously [55], an optimised ATP regeneration system has been introduced to p97 measurements in order to mimic a cellular environment where the ATP levels are high while ADP concentration levels remain minimal [83,84]. With this setup, in solution-state NMR measurements, a novel protein conformation has been detected which is distinctly different from the static nucleotide states (ATPγS and ADP) but shares overall similarities to the ADP state (Figure 17). This novel conformation could fit the expectation of a post-hydrolysis state, but necessary additional experiments were required to identify the nucleotide.

To fulfil this task, solution-state ^{31}P NMR is the preferred method. However, due to the large size of the p97 hexamer and the resulting line broadening of signals originating from the bound nucleotide, such an approach is impractical. Instead, solid-state ^{31}P NMR is the method of choice in this thesis. Employing different pulse programs, it is possible to differentiate between the free nucleotides in solution (directly pulsed experiments) and the nucleotides bound to the sedimented protein (cross-polarization-based experiments).

In this chapter, ^{31}P NMR measurements are used to study nucleotide states of p97. First, reference measurements of free nucleotides in solution are performed in order to better understand some of the later encountered effects. Then, solid-state ^{31}P NMR measurements of nucleotides bound to p97 in the absence or presence of ATP turnover are reported, and the signals are assigned and discussed.

3.2 Free nucleotide measurements

3.2.1 Introduction

Before spectra of nucleotides bound to p97 can be analysed, it is important to perform reference measurements of free nucleotides in solution: Measurements of nucleotides alone (e.g. ADP or ATPγS) require a simple buffer setup containing low salt concentrations. On the other hand, measurements under ATP turnover employ a complicated ATP regeneration system which requires additional ionic species (Mg^{2+}, KCl) and, most importantly, the reaction fuel, phosphoenolpyruvate (PEP), which is converted to inorganic phosphate during the ATPase reaction. Thus, ionic strength and Mg^{2+}-induced effects are expected to play a role.

ATP hydrolysis leads to the release of a proton. Thus, during ATP turnover, the NMR sample is slowly acidified while PEP is consumed and inorganic phosphate is released. At a pH of 7.5, ADP and ATP are almost fully deprotonated, although multiple ionic species (including Mg^{2+}-complexes) can co-exist in solution. With a pKa value of ~6.5 for $HATP^{3-}$, a mixture of ATP^{4-}, $HATP^{3-}$ and the $MgATP^{2-}$ complex is expected to undergo a fast exchange during the measurement [85,86]. pH-dependence studies have shown that the deprotonation of phosphates leads to a downfield shift, implying that interactions that stabilise the deprotonated form of the nucleotides lead to an overall increase in the chemical shift values [87,88]. While the buffer system in our measurements is supposed to minimise the effect of acidification during the measurement, it is worthwhile to investigate effects that can still be encountered during real-time measurements of enzymatic ATP hydrolysis.

Finally, for reaction control during ATP turnover, free molecules present in the solution need to be monitored and their resonances assigned unambiguously. Thus, to prepare for the more complicated solid-state ^{31}P NMR measurements of the protein, these effects are first tested on a simple setup of pure nucleotides in a minimal buffer system, and the individual chemical shifts are assigned.

3.2.2 Chemical shift assignment and the influence of the ionic strength

At first, two sets of experiments were performed: ATP and ADP were dissolved in two different buffer systems, termed high and low salt buffers, and measured using directly pulsed ^{31}P NMR experiments (Figure 19). The buffer compositions of the low and high salt buffers are given in

Table 2. The low salt buffer composition was chosen for its simplicity, while the additional salts in the high salt buffer (NaCl, KCl) more accurately represent the actual measurement conditions of the ATP turnover setup (see Methods, Table 17).

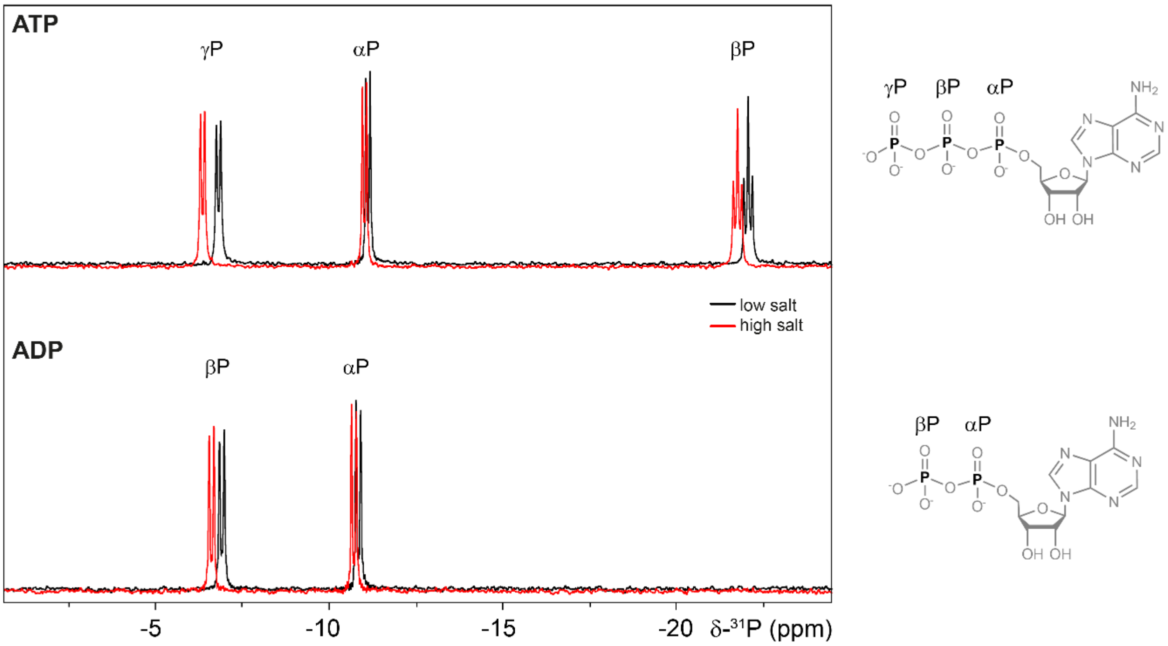


Figure 19. Higher ionic strength induces a downfield shift. Directly pulsed ^{31}P measurements of ATP and ADP in low and high salt buffers (left). ATP and ADP were dissolved in low (black) or high (red) salt buffers in the absence of Mg^{2+}. The increased ionic strength of the high salt buffer leads to a downfield shift of the peaks. On the right, the phosphate group nomenclature is shown. Figure adapted from ref. [78].

In ATP spectra three, in ADP spectra two groups of coupled peaks were observed, corresponding to the given number of phosphates. The peak splitting of ~15.5 Hz (ATP) or ~18 Hz (ADP), respectively, is attributed to $^{2}J_{P\text{-}P}$ couplings. Peaks were assigned based on well-matching literature values [89]. It is worth pointing out that the order of αP and βP peaks flips from ATP to ADP, with the βP-ADP peak having a chemical shift of approximately -6 ppm, while its counterpart of ATP has a chemical shift of -19 ppm.

Table 2. Composition of high and low salt buffers.

Buffer	Composition
Low salt	5 mM nucleotide, 10 mM HEPES, 1 mM DSS, 5% D_2O, pH 7.5
High salt	5 mM nucleotide, 10 mM HEPES, 1 mM DSS, 5% D_2O, pH 7.5 25 mM NaCl, 50 mM KCl,

Finally, an increasing ionic strength induces a downfield shift on all phosphate peaks. However, the shift is non-uniform, and the αP shifts of both nucleotide species seem to be less affected than those of other phosphates. This ionic strength-induced downfield shift has been reported [90] and was also observed by our computational study [78]. The downfield shift of phosphate peaks with increasing ionic strength has been attributed to a decrease in pKa values: Additional monovalent metal ions compete with H^+ ions for interactions with the negatively charged nucleotides, stabilising the deprotonated state [91].

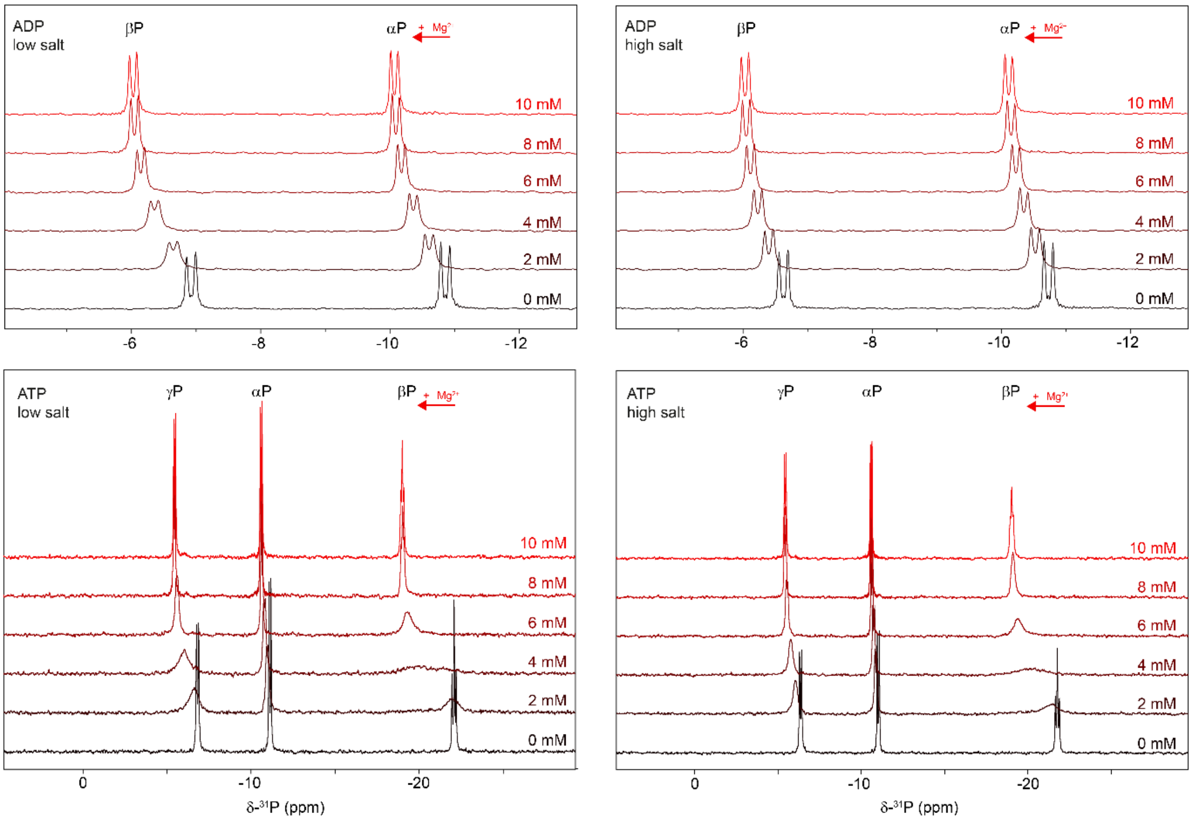


Figure 20. Magnesium titrations of nucleotides in high and low salt buffers. Directly pulsed ^{31}P measurements of Mg^{2+} titrations with 5 mM ADP (top) or ATP (bottom) in low (left) and high (right) salt buffers. $MgCl_2$ concentrations range from 0 to 10 mM in 2 mM steps. Mg^{2+} induces a downfield shift on the phosphate peaks, which is much more pronounced for βP-ATP. Additionally, for intermediate Mg^{2+} concentrations (between the two extrema), line broadening affects the line width which is much more pronounced for ATP in general, and specifically βP- and γP-resonances.

3.2.3 Effect of the Mg^{2+} concentration on chemical shifts

To test the influence of the Mg^{2+} concentration on the nucleotide spectrum, $MgCl_2$ titrations were performed in both buffer systems, with concentrations ranging from 0 to 10 mM in 2 mM steps (Figure 20). Noticeably, peaks of both ATP and ADP shift downfield with a higher Mg^{2+} concentration, implying a strong magnesium ion binding (K_D~100 µM [86]). Overall, this downfield shift is much more pronounced for ATP, for which the maximum downfield shift difference Δδ (between 0 and 10 mM Mg^{2+}) of 1.36 ppm (γP) and 3.09 ppm (βP) is far larger than the one experienced by ADP resonances (0.79 ppm for αP, 0.89 ppm for βP). The only exception to this rule is the moderate downfield shift of the αP-ATP resonance, which experiences a maximum shift of 0.51 ppm, suggesting a non-uniform binding mode of magnesium to the nucleotide. Again, the downfield shift can be explained by a stabilisation of the deprotonated species via Mg^{2+}-complexation, leading to a decrease in pKa values.

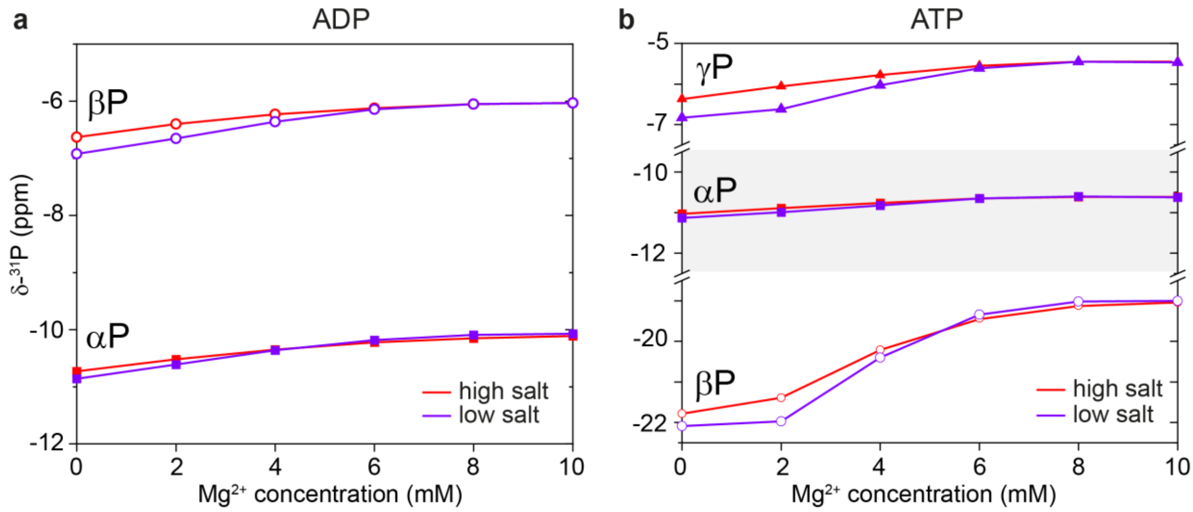


Figure 21. Increasing Mg^{2+} concentration induces a downfield shift on nucleotide resonances. Analysis of nucleotide chemical shifts depending on the solution Mg^{2+} concentration. Both ADP (**a**) and ATP (**b**) peaks shift downfield with increasing Mg^{2+} concentrations. This downfield shift is much stronger pronounced for γP- and especially βP-ATP resonances. Under substochiometric conditions (below 5 mM) peak chemical shifts are affected by ionic strength, but the chemical shifts equalise under stochiometric conditions and the buffer dependency disappears.

The overall Mg^{2+}-induced downfield shift is more pronounced for nucleotides in the low salt buffer (Figure 21). This is rationalised by the observation the ionic strength affects the chemical shifts only below substoichiometric conditions (below 5 mM $MgCl_2$). In contrast, above conditions of equal stoichiometry (Mg^{2+}:nucleotide = 1:1), nucleotide chemical shifts of the two different buffers equalise, overruling the effect of the ionic strength. It should be pointed out that downfield shift continues even above saturating conditions (most prominently visible for βP-ATP), seemingly approaching a final value asymptotically: Mg^{2+}-free and complexed forms of the nucleotides exist in an equilibrium, and an increase in magnesium concentration further pushes the equilibrium towards the Mg^{2+}-bound form. This explanation is supported by the fact that ionic strength does not affect the chemical shift above stoichiometric conditions (c.f. low and high salt buffers).

Under non-saturating conditions, a fast exchange between Mg^{2+}-bound and Mg^{2+}-free nucleotide forms is expected to exist. The peak chemical shifts at 0 mM Mg^{2+} can be used as a reference for the Mg^{2+}-free form, while the chemical shifts at 10 mM Mg^{2+} represent the metal-bound forms, explaining the continuous peak shift throughout the titration and the line broadening in intermediate conditions. Previous studies reported that titrations of further equivalents of magnesium ions lead to a reversed trend of upfield shifting [90]. This reversal, however, is of no relevance for the protein measurements as the latter are conducted under substoichiometric conditions.

3.2.4 Effect of Mg^{2+} concentration on line broadening

The formation of Mg^{2+}-nucleotide complexes and the subsequent exchange with the metal-free forms is a major source of line broadening, which is visible even to the naked eye (Figure 20). Upon closer inspection, it becomes apparent that the broadening only appears in intermediate regimes, and disappearing at the extrema, such that the line width in both 0 mM and 10 mM Mg^{2+} spectra is 7-9 Hz for ADP and 7-14 Hz for ATP (Figure 22). In intermediate regimes, line broadening affects all peaks and nucleotide species, although not to the same extent: in general, ADP resonances are only mildly broadened to maximum line width of 17.6 Hz (βP-ADP with 2 mM $MgCl_2$) and the broadening completely disappears above the Mg^{2+}-ADP ratio of 1:1. In contrast, βP and γP-resonances of ATP completely lose their multiplet structure due to the large broadening, leading to a maximum peak width of 560 Hz for βP-ATP. Noticeably, both resonances are far from equally affected, as the maximum broadening of γP-ATP barely reaches 100 Hz. This is not surprising, however, since the difference in the chemical shifts of Mg^{2+}-bound and free form is over twice as large for βP than γP, thus having a larger exchange-driven broadening. The only exception to the ATP resonance broadening is presented by the αP-resonances: these are affected by line broadening only to a similar extent as ADP resonances, agreeing with the observation of a limited chemical shift dependence of this resonance on the Mg^{2+} concentration.

Additionally, the broadening of αP-ATP returns to a minimal value already at a Mg^{2+}-ATP ratio of 1:1, while βP and γP-resonances show higher line widths to much higher metal ion concentrations and clearly remain worse resolved than the same peaks at 0 mM. This observation can be explained by assuming an intermediate exchange between several different species, incl. $MgATP^{2-}$, $HATP^{3-}$ and possible nucleotide dimers and trimers mediated by metal binding [88]. Obviously, these exchanges cannot occur in the absence of metal ions. A possible exchange between mono- and bi-metal forms, however, appears to be only of minor relevance: Mg_2ATP complex formations have been measured to have a stability constant of below one per mille of the monovalent complex [92]. Based on published complex stability values, these considerations should apply to ADP as well, leading to a slight broadening with 10 mM $MgCl_2$ [88]. Potentially, a lack of resolution prevents this observation for ADP under the present conditions but is detectable for ATP due to the large chemical shift differences.

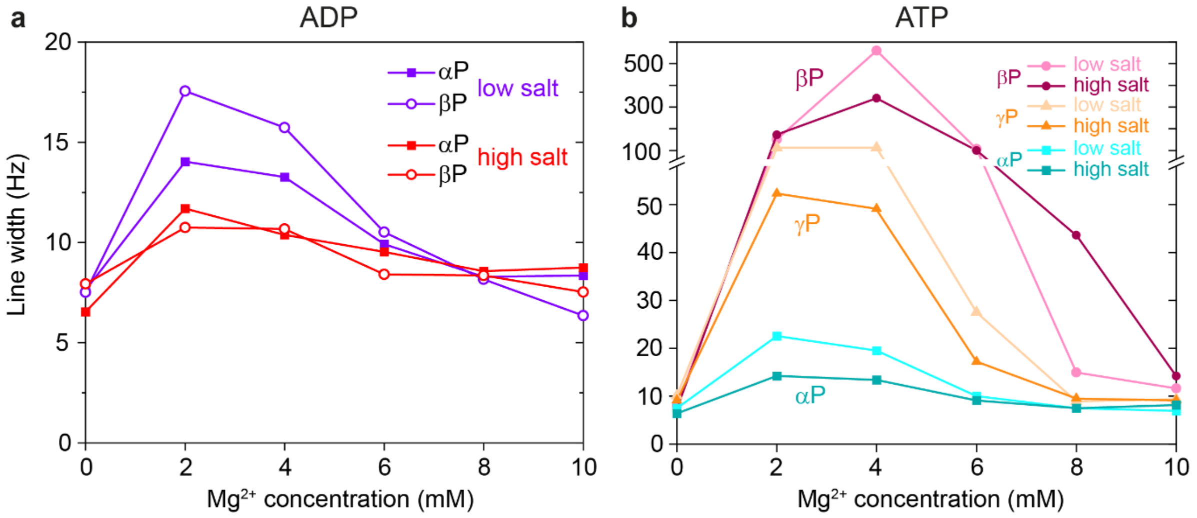


Figure 22. Mg^{2+} ions induce line broadening. Analysis of peak line widths of ADP (**a**) and ATP (**b**) depending on the Mg^{2+} concentration. Peaks are the narrowest at both extremes while increasing for intermediate concentrations. ATP is much more affected by line broadening, especially γP and βP-resonances. Line broadening is decreased in the high salt buffer. Line widths (full width at half maximum) were determined from peak fitting using MNova v11. Spectra were processed with an exponential window function with 5 Hz. In cases where the multiplett structure is no longer observable, the line width is given for the whole peak.

Finally, it can be documented that the broadening is generally reduced in the high salt buffer, which can be explained by competing Na^+/K^+ electrostatic interactions with the nucleotide phosphates, thus reducing the strength of interactions with magnesium ions. Again, in the superstoichiometric regime, the difference in line widths in the two buffer systems vanishes as the Mg^{2+} binding overwhelms any other ionic interactions.

3.2.5 Real-time reaction control

With the results of the aforementioned peak assignments, ATP turnover can be directly followed using ^{31}P NMR. For solid-state NMR measurements, the sample is sedimented into a MAS rotor, but the protein remains surrounded by the reaction buffer, in which small molecules can freely move around. Thus, using directly pulsed ^{31}P 1D NMR experiments, the composition of the molecules, specifically free (non-bound) nucleotides, can be determined in real time. Over the course of the whole measurement, a depletion of PEP and an increase in inorganic phosphate is observed (Figure 23). As long as PEP is present in the solution, ADP is regenerated to ATP. Thus, its presence remains below the limit of detection. After PEP is depleted, ADP starts to emerge, creating a mixture of nucleotides in solution. Eventually, all ATP is used up, and only ADP remains present. The minor adenosine monophosphate (AMP) peak is attributed to a slow degradation of ADP due to apyrase contamination [93].

Table 3. Comparison of line widths of free ATP under solution and solid-state conditions. Except for the αP-resonance in the solution-state experiment ([a]), no multiplett structure was resolved. Line widths are given at full width half maximum of the unresolved peaks. [a]: Line width is given for the full resonance, even though the multiplett structure is partially resolved.

Measurement / line width [Hz]	$\Delta\nu_\alpha$	$\Delta\nu_\beta$	$\Delta\nu_\gamma$
Solid-state NMR (7°C)	54.4	186	56.4
Solution-state NMR (25°C)	32.3[a]	167	51.6

The reaction system was set up with 10 mM of ATP and 4 mM $MgCl_2$, amounting to a Mg^{2+}-ATP ratio of 2.5. The ratio corresponds to the free nucleotide measurement of ATP in the high salt buffer with 2 mM Mg^{2+}. This can partially explain the overall broadened lines and the broad βP-ATP resonance around -20 ppm in particular. A quantitative line width comparison between the solid and solution-state NMR measurements indicates that the line widths of γP and βP resonances remain similar throughout both measurements, but the αP peak is significantly narrower under solution-state conditions (Table 3). The origin of the additional broadening of the αP-resonance under solid-state conditions is uncertain, as it could potentially be attributed to temperature (solution- and solid-state measurements were performed at different temperatures) or viscosity effects.

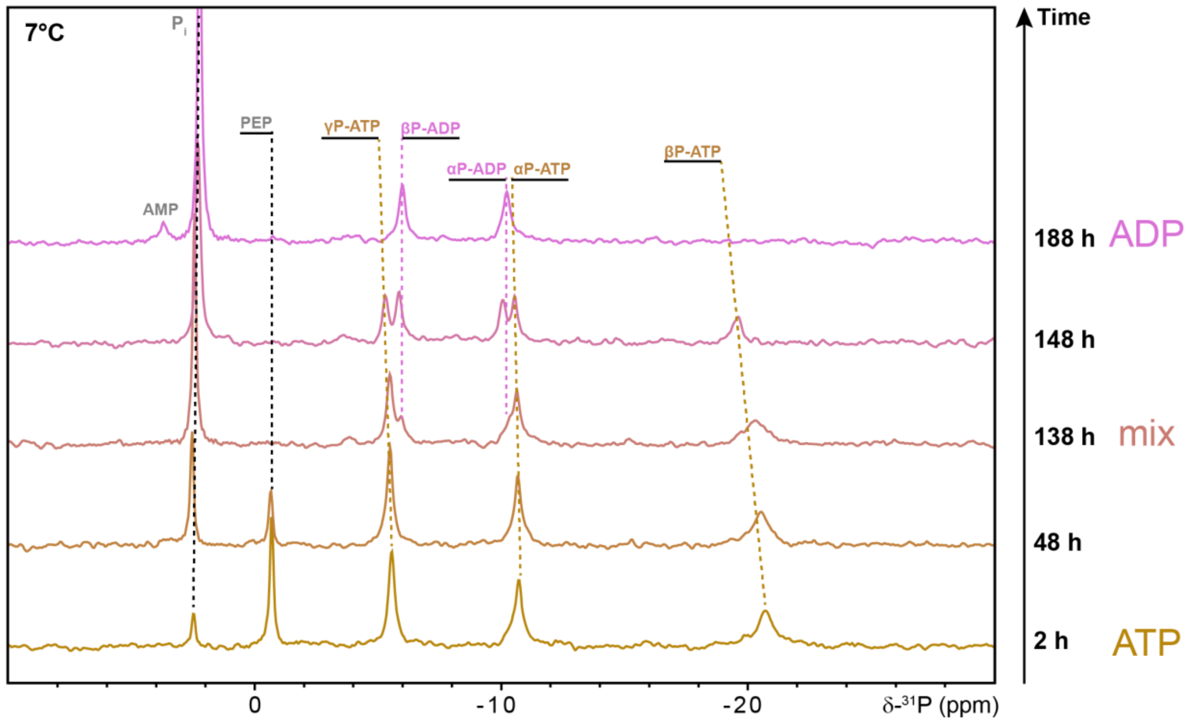


Figure 23. ^{31}P NMR used as real-time reaction control. Directly pulsed ^{31}P 1D NMR measurements were used for reaction control during ATP turnover. Over time (from bottom to top), phosphoenolpyruvate (PEP) is depleted, while inorganic phosphate (P_i) is built up. At the beginning of the reaction, while PEP is still present in solution, only ATP is detected, while over the course of the measurement a nucleotide mixture emerges, until only ADP is left at the end. The measurement was recorded at 45 kHz MAS with the temperature internally calibrated to 7°C. The nucleotide concentration was 10 mM, the $MgCl_2$ concentration 4 mM, resulting in a Mg^{2+}-ATP ratio of 2.5. Figure adapted from ref. [55].

Another effect to notice is the continuous downfield shift of resonances during the measurement. Since during ATP hydrolysis by p97 one inorganic phosphate ion is released per cleaved ATP, the P_i concentration in solution increases over time. As stated previously, decreasing pH (i.e. pushing the equilibrium towards the protonated form; Figure 24b) should lead to an upfield shift, but the opposite is observed. This can be explained by an effective increase in Mg^{2+} concentration owing to the loss of Mg-chelating PEP. The stability constant of the Mg^{2+}-PEP complex is approximately 2.5 times as large as that of Mg^{2+}-P_i [86], leading to an increase of the free magnesium ions over the course of the measurement. This also suggests that the initial Mg^{2+} concentration experienced by free ATP is actually lower than 4 mM since several other species (PEP, P_i, ribose-5-phosphate) bind magnesium ions, although generally with higher K_D values [86].

3.2.6 Modes of Mg-nucleotide binding

The modes of coordination of the Mg^{2+} ion by ATP have been a long-standing question. In their computational study, Buelens and coworkers [94] analysed two different chelation geometries for the $MgATP^{2-}$ complex, which differ by the number of phosphates involved in the metal binding, namely two or three (termed C2 and C3; Figure 24a), leading to a folded or extended phosphate ester tail. Although a strong dependence of the βP and γP, but not of the αP chemical shifts on the Mg^{2+} concentration implies a preferential bidentate metal binding with the two terminal phosphates, this conclusion can be misleading, as pointed out by Jaffe and Cohn already in 1978 [89]. Recent QM/MM calculations show that ^{31}P NMR shifts strongly react to changes in bond lengths and angles [78], preventing an easy conclusion from the lack of strong ^{31}P chemical shift dependence of the αP-resonance on Mg^{2+} concentration.

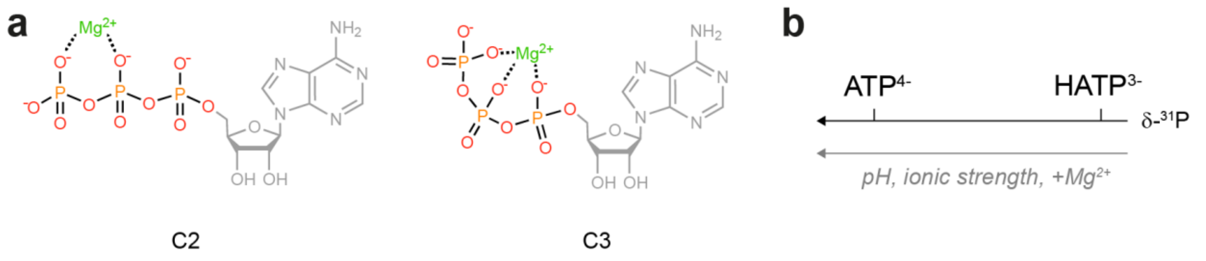


Figure 24. Physical and conformational parameters affecting the ATP chemical shift. a) Possible chelation geometries of the $MgATP^{2-}$ complex. Mg^{2+} ions can coordinate ATP phosphates in two different geometries. The bidentate coodination of the βP- and γP-phosphates, termed C2 (left), is characterised by an elongated phosphate ester tail. In the tridentate binding mode (C3, right) all three phosphate groups are involved, requiring a folded phosphodiester backbone. **b**) At a neutral pH, ATP exists in an equilibrium between the monoprotonated $HATP^{3-}$ and the fully deprotonated ATP^{4-} forms. An increase in the pH value, the ionic strength or the Mg^{2+} concentration leads to the stabilisation of the deprotonated state, thus increasing the chemical shift.

Conversely, it is reasonable to assume that the two ATP conformations would produce vastly different chemical shifts. However, in the spectra, there is no indication of the presence of additional species, implying that the exchange between the conformations is either very fast and not hindered by a high energy barrier, or one of the conformations is simply not present in the solution.

The distribution of bi- and tridentate coordination was investigated quite early using various approaches. Some studies lacked direct experimental proof and worked instead with substituting bivalent metal ions like Cd^{2+} or Mn^{2+} and drew conclusions from measurements with thiosubstituted nucleotide analogues [88], but these modified systems can behave differently from the native Mg^{2+}-ATP complex [95,96]. First direct experimental evidence for both bi- and tridentate magnesium binding was provided using NMR of ^{17}O-labelled ATP, estimating the two coordination geometries to be populated to an approximately equivalent extent [97]. Considering that the geometries are approximately isoenergetic [94], this implies that the geometry exchange must be fast and thus not hindered by a high energy barrier, which is consistent with the corrections to the computational approach introduced in the aforementioned paper [94].

Still, another valid question is of the relevance of the tridentate conformation for enzymatic reactions, as the vast majority of published PDB structures of P-loop ATPases bind magnesium in the C2 conformation [94].

3.3 Measurements of the bound nucleotides

3.3.1 Introduction

^{31}P chemical shifts are not only very sensitive to the geometric and electronic structure of the nucleotide [78] but also to the immediate surrounding environment. It is thus reasonable to assume that a nucleotide bound in a protein active site would have different chemical shifts than the one moving freely in a solution. Nonetheless, building on the foundation of results we acquired in the previous section, we can turn to the actual task: the measurement of the bound nucleotides.

The approach to the task is, in principle, straightforward: the protein needs to be sedimented into the solid-state MAS rotor, and the bound nucleotide selectively detected, while real-time reaction control is performed by measuring free nucleotides in solution. This can be achieved by interlacing two types of experiments and repeatedly measuring them. The first one is a cross-polarization-based ^{1}H-^{31}P transfer, allowing the selective detection of the bound nucleotides. This dipolar-interactions-based transfer is efficient for immobilised and solid structures but very inefficient for flexible, mobile moieties. The freely rotating nucleotides in the surrounding solution, which make up the overwhelmingly large part of all ^{31}P nuclei in the whole rotor, become invisible in this NMR experiment. Since the CP transfer depends on dipolar (i.e. through space) interactions, the lack of directly bound protons to the nucleotide is not an issue, although it can affect the transfer efficiency. Additionally to ribose and base protons, inside the binding pocket protons from amino acid side chains are plentiful and can be used as the original source of ^{1}H magnetisation.

The second type of experiments is again a directly pulsed ^{31}P 1D, detecting all nuclei present in the rotor. This is used to monitor the reaction and nucleotide identities in real time, as is demonstrated in Figure 23. For optimal results, both types of experiments are interlaced and repeated until a sufficient signal-to-noise ratio is achieved or the reaction is finished.

From the solution-state NMR experiments, it was known that the ADP·P_i state shared many similarities with the ADP state. Prior to measuring the bound nucleotide during ATP turnover, it was essential to measure references of the static states, ATP and ADP. Since the ATP state is not stable (ATP is immediately hydrolysed), a p97 disease-associated mutant (R95G), which has its rate-determining step changed from product release to hydrolysis [55], was used instead. Furthermore, solution-state experiments observed the ADP·P_i state both in measurements of the full-length and the truncated ND1L constructs, but for solid-state NMR, initially, only the ND1L construct was chosen to reduce spectral complexity. Measurements of the full-length construct will be shown and discussed at the end of the Chapter.

Prior to the measurement, answers to several key questions were unknown: solid-state NMR spectra suffer from strong line broadening in comparison to solution-state NMR spectra, raising the question of whether the resolution would even be sufficient to separate the different phosphate resonances. *A priori,* it was not known how many signals one would encounter in the first place. While two or three peaks for the ADP or ATP states, respectively, are to be expected, it remained uncertain how many signals an ADP·P_i state might produce. Given the nature of this state, an ADP-like spectrum with an additional peak for the cleaved inorganic

phosphate appears to be reasonable, but on the other hand, if the ADP·P_i state shares similarities with a “barely cleaved” ATP, an ATP-like spectrum could also be expected.
This also raises the question of assignments. While the measurements of the free nucleotides in the previous section present a solid foundation for subsequent experiments and provide a basis for peak assignment, it is unknown in advance how the resonances might be affected by the surrounding protein environment. Finally, the advantage of the CP-based selection method is also its biggest weakness. While directly pulsed experiments can capture flexible moieties and cross-polarization-based experiments work well with rigid moieties, neither of them provide high signal-to-noise for molecules falling in the intermediate mobility regime. In case the ADP·P_i state is too dynamic, there is a possibility that the nucleotide signal would be too weak for detection.

3.3.2 Measurement of the reference states

The simplest reference state to measure is one of p97-ND1L in the presence of ADP (Figure 25). This spectrum was recorded in a setup involving only the protein, buffer system and ADP. In this ^{31}P NMR spectrum, two peaks are detected, presumably corresponding to the αP and βP-resonances. In comparison to the free nucleotide reference, these peaks are shifted downfield, but the peak separation of ~4.2 ppm remains identical in both cases. As mentioned previously, a chemical shift perturbation of the nucleotide phosphates upon binding to the protein is expected. The assignment of these two peaks to the corresponding phosphate groups is rather straightforward since this spectrum shares high similarities with the free nucleotide reference. Noteworthy, in the protein context, the linewidth seems to vary between the two resonances, with the αP-resonance being substantially broader and thus less intense.

The other necessary reference involves an ATP-bound state. Since ATP is readily hydrolysed, it is impossible to simply supply p97 with ATP and record a spectrum. Instead, a disease-associated mutant R95G was employed. Solution-state NMR measurements in the presence of the regeneration system showed that for this mutation, the rate-determining step was shifted from product release to hydrolysis, showing pre-hydrolysis state spectra [55]. However, since R95G additionally leads to an increased ATPase activity [98], this mutation was cloned into the ND1L-E305Q background to limit the ATPase activity. This is necessary to extend the possible measurement time to record insensitive ^{31}P spectra before the regeneration system is depleted.

The spectrum of the ATP-bound reference shows three peaks, which presumably correspond to the respective phosphate groups. As is the case with the ADP-bound spectrum, the peaks differ significantly in intensity and line width. The most upfield resonance is assigned to the β-phosphate, following the logic of the free nucleotide reference, despite the chemical shift difference of ~4.6 ppm. The presumed αP-resonance is assigned following the same argument, showing a downfield chemical shift from the free nucleotide reference of ~4.9 ppm.

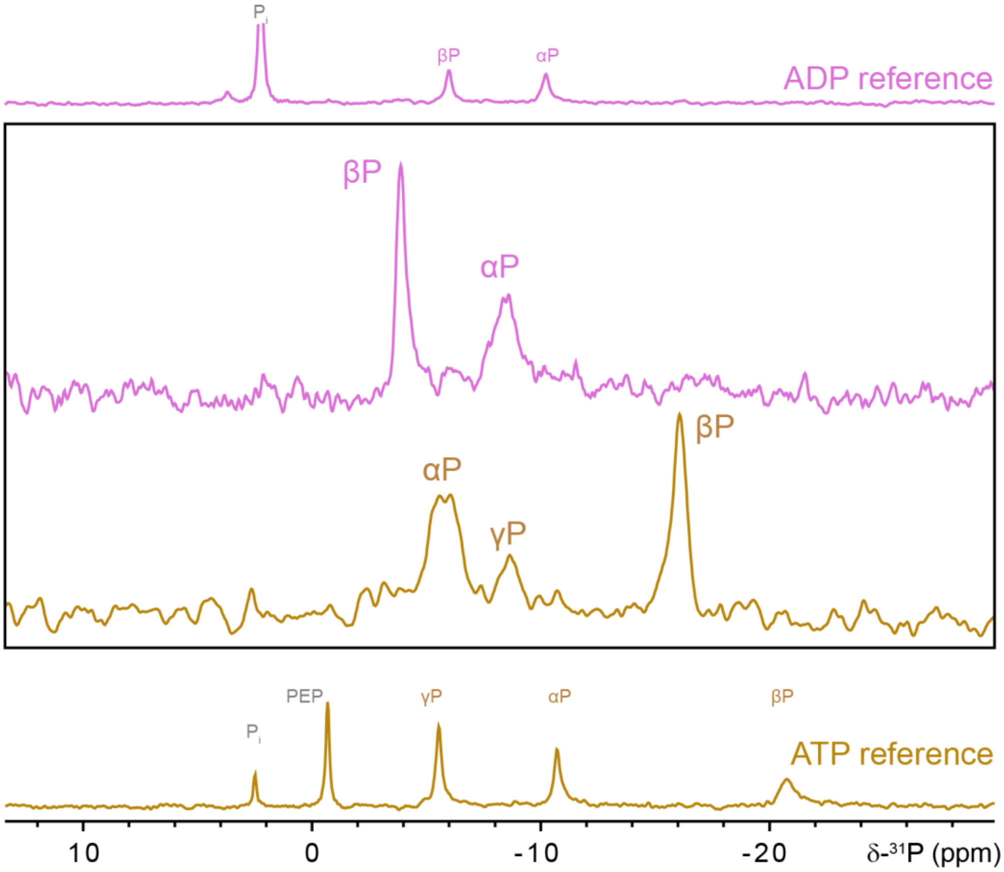


Figure 25. ^{1}H-^{31}P Cross-polarization based ^{31}P NMR spectra of p97 reference states. Reference states of ND1L constructs were recorded in the presence of ADP (pink) or ATP (brown). ADP bound to ND1L-E305Q produces two unequal resonances which are shifted in comparison to the free nucleotide reference (top). The ATP bound state was measured using a mutant (R95G) whose rate-determining step is changed to hydrolysis in the ND1L-E305Q background in order to slow down hydrolysis activity. Three groups of signals of varying intensity are detected which are assigned to the three phosphate groups of ATP. (Peak assignment is discussed in Section 4.3.4.) The free nucleotide (directly pulsed ^{31}P) spectra were recorded in ~2 h. The CP-based spectra were recorded in 16.3 h (ADP) and 33 h (ATP).

This argument is supported by the observation that the peak widths of the αP and βP resonances show similarities to the respective resonances of the bound ADP: both βP-peaks are sharp and narrow, while the αP-peaks are less intense and broad. Interestingly, the αP-resonance shows signs of a possible peak splitting, indicating the potential presence of multiple peaks. Finally, the last peak is assigned to γP simply by the process of elimination. The behaviour of this peak strictly differs from that of the other resonances, as its protein-bound resonance is shifted upfield from the corresponding free nucleotide reference. Additionally, the low intensity of this peak could be explained by an increased flexibility of the terminal phosphate, leading to an inefficient CP-transfer.

The overall lower signal-to-noise ratio of the ATP-bound spectrum, despite being recorded for twice as long as the ADP-bound spectrum, hints at the more dynamic nature of that state or can even indicate a mixture of several states (e.g. ATP and ADP·P_i). This is discussed below.

Finally, it is worth emphasising the general simplicity of both spectra. Although the peaks of the ATP spectrum indicate the potential presence of peak shoulders and multiple peaks, overall, only a minimal number of peaks is detected. Considering that these ^{31}P NMR spectra show bound nucleotides of six distinct binding pockets, the conclusion must be that the ND1L-hexamer is symmetric in nature – at least in the ADP state. These results fit those of the solution-state NMR protein spectra, in which a homogeneous protein state is observed.

3.3.3 Measurement of the ADP·P_i state

For the measurement of the ADP·P_i state, the Walker B mutant E305Q mutant was chosen in order to reduce the ATPase activity and prolong the possible measurement time. In solution-state NMR measurements, the ND1L-E305Q mutant produced spectra that showed a very high resemblance to wild-type references, including those measured in the presence of the ATP regeneration system [55]. This emphasises that under ATP turnover conditions, product release remains the rate-limiting step for the Walker B mutant, producing the ADP·P_i state, although this mutation should affect the local environment of the binding pocket.

In total, two measurements under ATP turnover were performed with similar results, one recorded at 7 °C and one at 20 °C. The real-time reaction control at 7°C is shown in Figure 23 and at 20°C in Figure 26a. As expected, the reaction progresses faster at elevated temperatures, with PEP already depleted after 6 h. In contrast, at 7°C the ATPase activity was reduced to such a degree that the rotor needed to be heated up to 30°C for 24 h in order to measure the end state after reaction fuel depletion (time period highlighted in grey in Figure 26b). Ratios of nucleotides (ATP/ADP in relation to total nucleotide concentration) from integrated peaks in the solution were plotted, splitting the measurements into time periods of different molecular composition: in region I, only ATP is present, in region II, PEP is used up and ADP start to emerge and in region III all ATP is depleted and only ADP remains. CP spectra recorded during those three time periods are shown in Figure 27. It is apparent that the spectra recorded in the presence of PEP, i.e. in the absence of ADP (region I, pure ADP·P$_i$ state), are dissimilar to the ones recording in the exclusive presence of ADP (region III, pure ADP state). In the intermediate region during which both ATP and ADP are present in solution, a mixture of both ADP·P$_i$ and ADP states is detected (region II).

Comparison with the reference spectra reveals certain features of the ADP·P$_i$ state (Figure 28): in the ADP·P$_i$ state, only two major peaks are detected (at -5.0 and -15.8 ppm) with chemical shifts and peak intensities and widths showing resemblance to the ATP-bound reference. However, a third obvious peak is missing despite the overall good signal-to-noise ratio. Instead, a broad signal between -2 and -12 ppm is detected, showing several peak maxima. Based on these observations, the two main peaks have been assigned to the αP and βP-phosphates, while the broad signal is tentatively assigned to the cleaved P$_i$ (see Section 4.3.4). At least six distinct signals were identified and fitted using Mnova, with line widths varying from 170 to 500 Hz (see insert in Figure 28). Thus, to explain the many different chemical shifts of the inorganic phosphate, large heterogeneity must be assumed, originating from different chemical or magnetic environments. Interestingly, the overall symmetry of the hexamer does not seem to be broken, as the presence of two main peaks still implies homogeneous binding among all six protomers. Thus, the environment heterogeneity of the cleaved phosphate does not disturb the rest of the protein.

Another point to notice is the aforementioned resonance flip that occurs between the ATP and ADP peaks. In the ATP-bound state, the order of the chemical shifts is αP>γP>βP, while in the ADP state the order is βP>αP in the order of increasing shielding. The order of those chemical shifts can indicate whether the local protein environment of this intermediate state is more similar to the pre-hydrolysis (ATP) or to the post-hydrolysis (ADP) states. For that reason, proof of assignment is required.

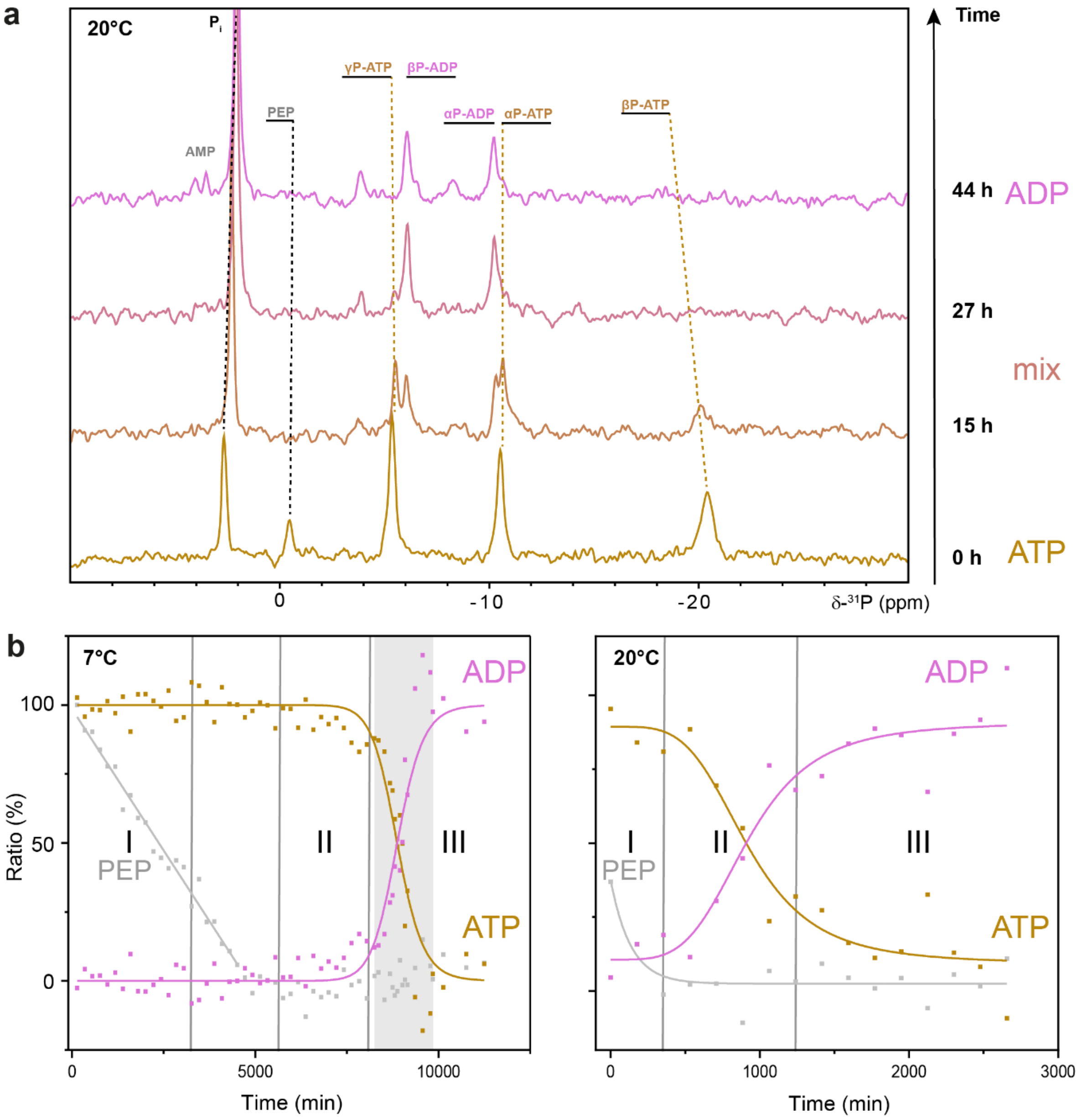


Figure 26. Real-time reaction control using ^{31}P NMR. a) Directly pulsed ^{31}P 1D NMR measurements were used for reaction control during ATP turnover at 20°C. An identical measurement performed at 7°C is shown in Figure 23. At the beginning of the reaction (bottom), while PEP is still present in solution, only ATP is detected, while over the course of the measurement a nucleotide mixture emerges, until only ADP remains left at the end (top). The measurement was recorded at 45 kHz MAS. **b**) Relative abundances of nucleotides and PEP have been determined from peak integration, fitted and plotted for the 7°C (left) and 20°C (right) measurements. In order to speed up the reaction at 7°C, the rotor was heated up to 30°C for ~24 h, as indicated by the grey box, and afterwards cooled to 7°C again. CP-based spectra acquired during time periods indicated by roman numeral are shown in Figure 27. Figure adapted from ref. [55].

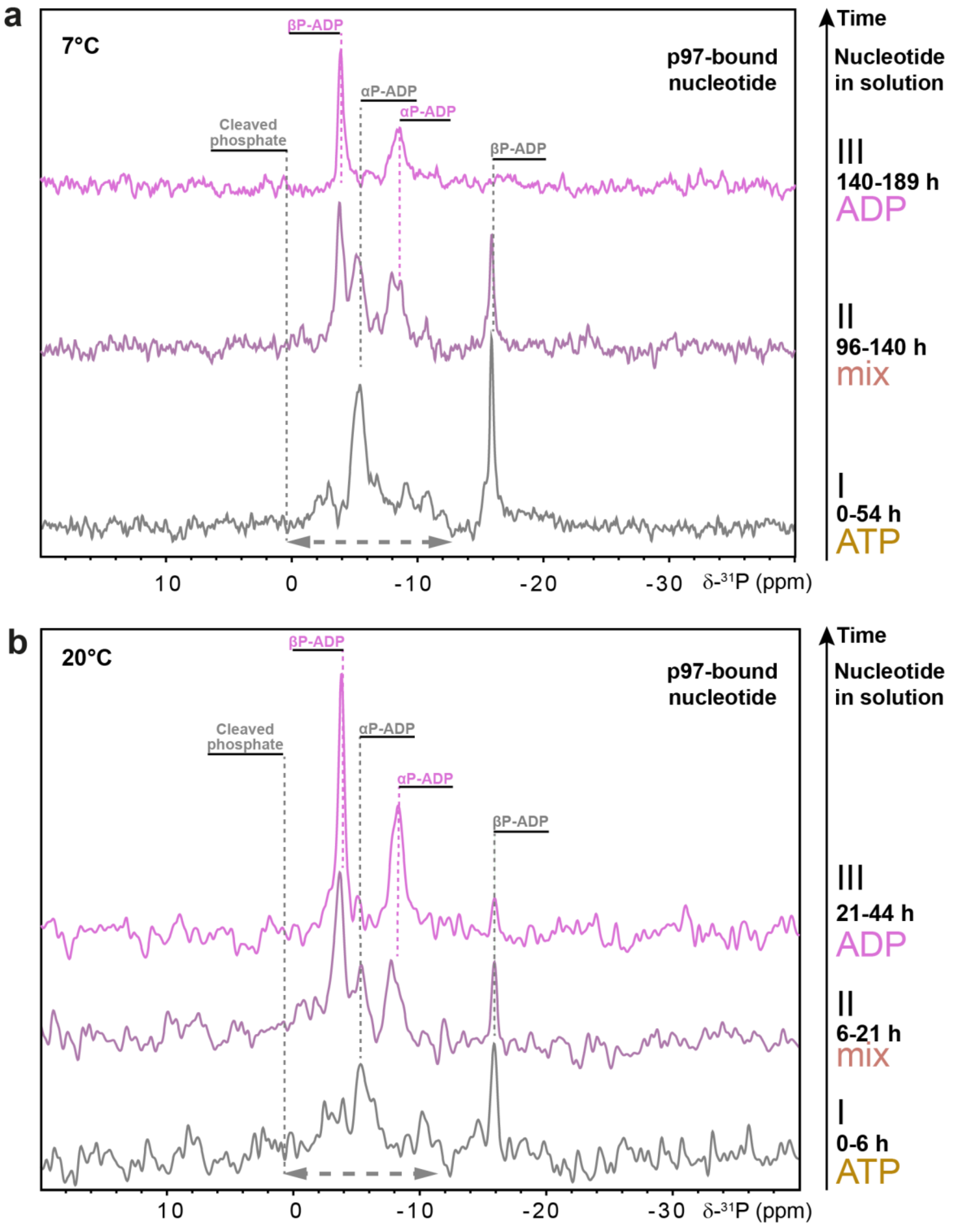


Figure 27. Cross-polarization ^{31}P NMR spectra of bound nucleotides during ATP turnover. Combined cross-polarization spectra recorded over the course of the whole rection at 7°C (**a**) or 20°C (**b**). Reaction and time progresses from bottom to top, as indicated by time stamps on the right. Roman numerals on the right indicate the corresponding time periods as they are shown in Figure 26. In the beginning, in the presence of PEP, the ADP·P$_i$ state is captured (region I), while at the end of the reaction when only ADP remains in solution, only bound ADP is seen in the CP-spectra (region III). In the intermediate regions, mixed spectra are recorded, proving the presence of both states (region II). The spectra were recorded at 45 kHz MAS. Figure adapted from ref. [55].

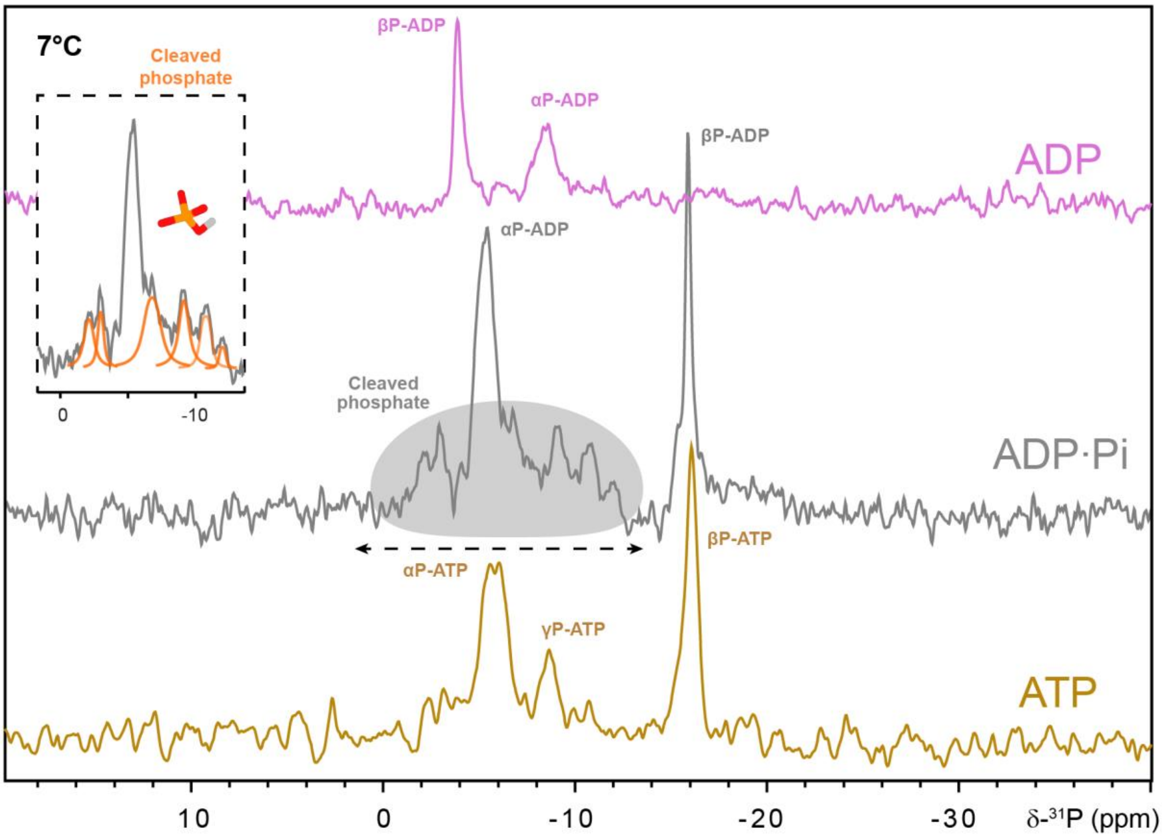


Figure 28. Comparison of ^{1}H-^{31}P CP spectra acquired for different nucleotide states. Spectra showing protein bound nucleotides were recorded for all three states: ATP-bound by a mutant which captures the pre-hydrolysis state in the presence of the regeneration system (bottom), ADP·P$_i$-bound by ND1L-E305Q during ATP turnover (middle) or ADP-bound by ND1L-E305Q in the presence of ADP alone. The ADP·P$_i$ state shares similarities with the ATP-bound reference in terms of peak distribution. Additionally, multiple small peaks are visible in the range between -2 and -12 ppm, belonging to the cleaved inorganic phosphate. The large dispersion of P$_i$ chemical shifts and the multitude of peaks hints at heterogeneity in the chemical environment experienced by the cleaved P$_i$. The ADP·P$_i$ spectrum was recorded with 131k scans in 37.3 h. Insert: Line fitting of the P$_i$ peaks using Mnova. Line widths vary from 170 to 500 Hz.

3.4 Solid-state ^{31}P NMR measurements on full-length p97

Eventually, the goal is to transfer the methodology to connect the protein state and ^{31}P NMR chemical shifts to the D2 domain of p97. While experiments on the truncated ND1L construct help to reduce the spectral complexity and peak overlap, neither do they present any information on the properties of the second ATPase domain nor do they prove biological relevance. Under solution-state NMR conditions, FL-wt and ND1L-wt both behave very similarly, producing the same protein states in the presence of identical nucleotides (see Figure 29) [55].

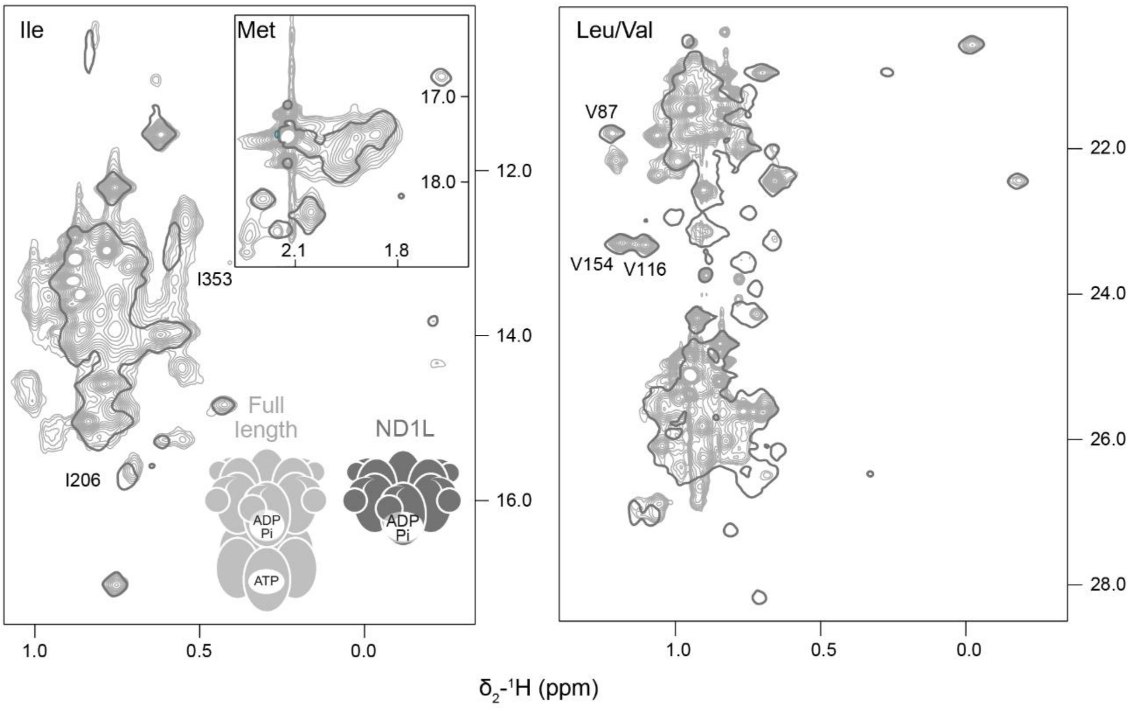


Figure 29. p97-ND1L protein spectra are not affected by the lack of the D2 domain. Methyl TROSY spectra of ND1L-wt (dark, single contours) and FL-wt (light grey) in the presence of ATP and the regeneration system produce both ADP·P_i states. The great overlap of the ND1L and the full-length spectra indicates that the ND1L construct is not affected in its protein state by the lack of the D2 domain and is a biologically relevant system to be studied. The same conclusions can be drawn from spectra recorded in the presence of ADP and ATPγS [55]. The additional peaks of the D2 domain have not been assigned.

Following this observation, solid-state ^{31}P NMR experiments were repeated on full-length constructs. However, since FL-E305Q shows an elevated ATPase rate (see Section 5.6.3), in order to decrease the ATPase activity to a manageable extent, an additional Walker B mutation in the D2 domain was introduced. Thus, the results under active nucleotide turnover were collected on the double mutant FL-E305Q-E578Q, while the static ADP state was collected on the full-length wild type. ADP-bound FL-wt produces a spectrum of four distinct peaks grouped in pairs of two (Figure 30). These peaks correspond to the αP and βP resonances of bound ADP in D1 and D2, respectively. Perfect peak overlap with the ND1L-acquired spectrum confirms that the lack of the D2 domain does not affect the binding in the D1 domain. Additionally, since only two additional peaks are detected, uniform nucleotide binding in all binding pockets can be assumed, as has been observed previously [32]. The successful resolution of D1 and D2 peaks of the ADP state encourages further, more complicated measurements on this system.

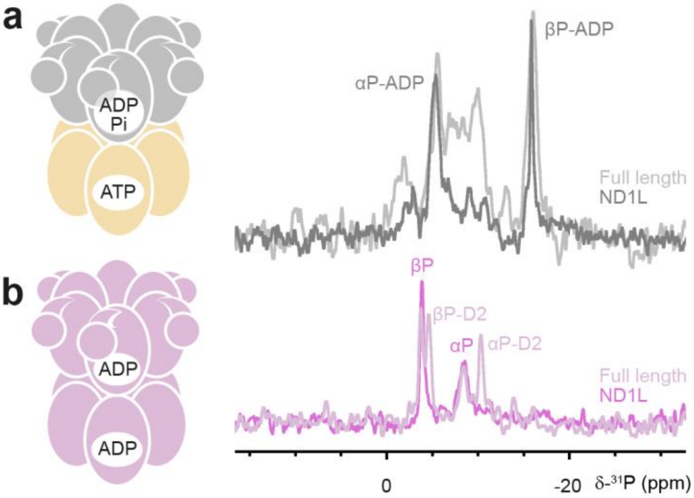


Figure 30. ^{1}H-^{31}P-CP spectra of full-length p97 overlap perfectly with ND1L-acquired spectra. ^{1}H-^{31}P cross-polarization spectra were acquired in the presence of the ATP regeneration system on FL-E305Q-E578Q (**a**) or ADP on FL-wt (**b**). For reference, an overlay of ND1L-E305Q based spectra are shown in lighter hues. On the left, a schematic representation of the hexamer is shown. Overall, nucleotide binding in D1 is not affected by the presence or absence of the D2 domain, as can be seen by the perfect overlap of the ND1L spectra with the full-length spectra. **a**) In the presence of the ATP regeneration system, multiple peaks are visible which are absent in the ND1L spectrum. The perfect overlap with the latter confirms ADP·P_i state binding in the D1 domain, however, the identity of the additional peaks originating from the D2 domain is unclear. In the cryo-EM reconstruction of p97 in the presence of the ATP regeneration system, ATP was detected in D2 [52]. **b**) In the presence of ADP, only two groups of two peaks are visible, corresponding to αP and βP bound by the D1 and D2 domains, respectively. The lack of additional signals confirms the hexameric symmetry of the state. The full-length ADP measurement was recorded at 35 kHz MAS; all other spectra were recorded at 45 kHz MAS. This figure was previously published in ref. [52].

Remarkably, the αP peak of the ADP bound in D2 is significantly narrower than its D1 counterpart.

In the presence of the ATP regeneration system, the full-length spectrum produces additional peaks not observed in the ND1L spectrum. However, the perfect overlap of both spectra can be interpreted that even in the full-length context, D1 is still found in the ADP·P_i state. The additional peaks are poorly resolved, and the overall limited signal-to-noise ratio prevents a successful spectrum subtraction. Thus, the chemical shifts and even the number of peaks originating from the D2 domain remain ambiguous, and the type of nucleotide state cannot be assigned. Based on the limited peak dispersion alone, both ATP and ADP·P_i states are conceivable, but certainly not the apo state. In the cryo-EM reconstruction of the full-length wild type under identical conditions, an ATP state was found in D2, suggesting that the new peaks rather originate from bound ATP [52].

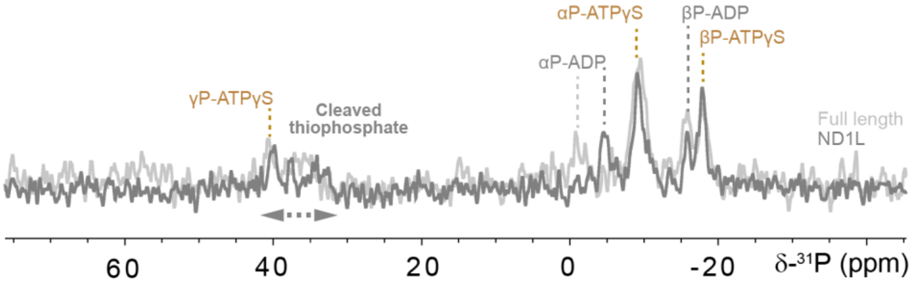


Figure 31. ^{1}H-^{31}P-CP spectra in the presence of ATPγS of full-length p97 overlap perfectly with ND1L-acquired spectra. A ^{1}H-^{31}P cross-polarization spectrum was acquired in the presence of ATPγS and the regeneration system on FL-E305Q-E578Q. For reference, the ND1L-E305Q spectrum is shown in a lighter hue. The state of nucleotide binding in D2 is unclear due to the lack of additional signals. All measurements were recorded at 45 kHz MAS. Peak assignment is discussed in Chapter 4.

In contrast, no additional peaks were observed in FL-E305Q-E578Q spectra recorded in the presence of ATPγS and a regeneration system, and the spectrum generally overlaps very well with the ND1L one (Figure 31). While this confirms ATPγS and ADP·SP_i binding in D1, not much information can be gain regarding the D2 domain. The absence of additional peaks can imply a complete lack of nucleotide binding in D2 (apo state), a highly dynamic binding pocket environment or a perfect peak overlap with the peaks from D1 (i.e. binding of identical species in D2). Based on the extremely slow ATPase rate of the reaction and literature reports of ATPγS binding in both domains [32,43], a significant spectral overlap with the D1 peaks is the most likely interpretation. Thus, slow hydrolysis of ATPγS to ADP·SP$_i$ is most likely found in D2 as well. Non-uniform nucleotide binding in the D2 domain, mixing ATPγS with the apo state, is possible as well. A previous study has shown that only 3-4 D2 binding pockets can bind ATPγS simultaneously [35]. The implication of the downfield shift of the αP-ADP·SP$_i$ peak is, however, unclear.

Overall, measurements of the full-length construct confirm the biological relevance of experiments on the ND1L construct and strengthen those results. Aside from the ADP state, however, interpretation of the full-length spectra is difficult and partially ambiguous.

3.5 Discussion

3.5.1 Characterisation of the ADP·P$_i$ state

In this work, solid-state ^{31}P NMR spectra of nucleotides bound to p97 were recorded. The goal of this endeavour was to characterise the ADP·P$_i$ state from the side of the nucleotide, as major characteristics of the protein side have been studied beforehand [55]. Overall, in the ADP·P$_i$ state, the protein conformation shares major similarities with the one in the ADP state, emphasising its post-hydrolysis nature. Methyl TROSY spectra alone cannot prove the binding of an additional inorganic phosphate, but such a signal should be detectable by ^{31}P NMR.

Prior to observing the ADP·P$_i$ state, we focused on obtaining reference spectra of ADP and ATP bound states, respectively. The static ADP bound state produces spectra with two well-resolved peaks, which are downfield shifted from the free nucleotide reference by ~2 ppm. Intuitively, the peaks were assigned to the αP and βP resonances without assuming peak order flipping. This follows the assignments of previous studies [99,100], where the upfield peak was assigned to the αP phosphate. No universal correlation can be found between the chemical shifts of free and bound nucleotides: in this work, a general downfield shift upon binding is observed (see Section 4.3.3), while another study reported increased peak dispersion [101]. Moreover, binding to different domains of the same protein produces different chemical shifts, as is seen for the full-length construct.

In order to prove peak identity, we used the thiosubstituted ADP analogue *Sp*-ADPαS and observed the bound nucleotide with a strongly shifted αS resonance but a βP chemical shift identical to native ADP (see Chapter 4). Moreover, QM/MM calculations of ^{31}P chemical shifts of ADP bound to p97 provided very strong agreement with the experiment [78]. These two independent experimental approaches support our intuitive assignment and establish a pathway for the analysis of subsequent experiments.

However, more information can be extracted from the ADP-bound spectra. For once, the observation of only two peaks agrees with the experimental evidence of a symmetric homohexamer [32]. The clear resolution of the additional D2-peaks in the context of the full-length protein suggests that a broken hexameric symmetry, as, for instance, in a stairway arrangement observed for adaptor bound p97 [36,102], would produce multiple distinguishable or significantly broadened resonances. Even more, studies have reported additional peaks with distinct chemical shifts [99] or at least visible as peak shoulders [100] originating from secondary binding states. Thus, in our setup, peak resolution is not a limiting factor.

The other observation pertains to the peak width. The αP resonance of ADP is significantly broadened in comparison to the βP resonance, leading to different peak intensities. The same features were reported for ADP bound to crystalline Ras [100], but it appears not to be a universally occurring property: ADP bound to DnaB or to BmrA produces different peak ratios [99,103]. The lack of such broadening of the αP-ADP resonance bound to p97-D2 proves that the origin of this effect is not an experimental artefact but rather a dynamic feature of the D1 binding pocket. In general, it is difficult to infer conclusions from this observation since the amount of the available initial magnetisation (i.e. number of ^{1}H spins in close proximity) and molecular rigidity have an influence on the peak intensity and width.

A similar distribution of peak intensities is observed for the reference ATP-bound state. Although three distinct peaks are detected, the γP-ATP peak is clearly the weakest and the βP peak the sharpest. Moreover, the peak order of the αP and γP resonances is flipped in comparison to the free nucleotide references, and the spectrum generally shares more similarities with the ADP·P$_i$ state spectrum rather than being a distinct state. This becomes even more obvious when the p97-R95G-E305Q ATP bound state is compared to the pre-hydrolysis states of DnaB, BmrA or Ras, which display three peaks of similar intensity and whose peak assignments can be directly traced from the free nucleotide references [99,103,104]. However, other examples show that spectra of bound ATP analogues may provide unexpected chemical shifts and peak broadening: DnaB-bound AMP-PCP shows three well-resolved peaks but with highly varying line widths [103]; DnaB-bound AMP-PNP instead produces a multitude of peaks of varying intensities [105]; and upon cooling, AMP-PNP bound to Ras displays a second peak for the γP resonance only [104].

Naturally, an ADP·P$_i$ state requires the binding of an ADP molecule with its two phosphate groups and, additionally, a third (inorganic) phosphate still bound by the protein, whose exact location is not known *a priori*. However, as will be discussed in Chapter 4, the release of the cleaved phosphate itself can still influence the state of the protein and affect the resulting spectra. The main feature of the ADP·P$_i$ bound spectrum is the detection of two main peaks and a broad distribution of weak, overlapping resonances. The latter feature can only be explained by a heterogeneous environment of a third ^{31}P-containing species, which, in this case, is necessarily the inorganic phosphate.

Still, the claim of the detection of a new, previously unobservable state requires supporting proof, which is provided in the following.

First of all, the conditions under which the ATP turnover measurement was performed produced a solution-state NMR protein spectrum which was completely different from ADP and especially from ATP(γS)-bound states. The only major difference between the solution and solid-state conditions is the temperature (7°C vs. 50°C), but the same features were observed by ^{31}P solid-state NMR at 20°C as well (Figure 27), supporting the idea that the temperature has a limited influence on the spectral appearance.

Secondly, spectra recorded in the presence of ATPγS display a mixture of two states: one which strongly resembles a bound ATPγS and one with chemical shifts very similar to the ADP·P_i spectrum (see Chapter 4, Figure 46a). Finally, a range of weak peaks appears between 30 and 38 ppm, while the broad base between -2 and -12 ppm is absent. This thiosubstitution-induced downfield shift can only be rationalised by a γSP or a cleaved SP_i phosphate group. Since a distinct γSP peak is visible in the spectra, it must be assumed that the broad peak distribution of the cleaved P_i has indeed been downfield shifted.

Thirdly, in the cryo-EM reconstruction of full-length p97 under ATP turnover in the D1 domain bound ADP, additional, unaccounted cryo-EM densities have been detected. These findings were consistent with MD simulations of the ADP·P_i state, which saw two stable positions of both the Mg^{2+} ion and the cleaved phosphate. Upon combining the results from these two methods, two stable post-hydrolysis states have been identified [52]. It cannot be ruled out that further stable intermediate states exist. Certainly, the inorganic phosphate needs to be transported out of the binding pocket at some point after hydrolysis. This process, discussed in detail in Chapter 4, produces additional (short-lived) states in which several positively charged residues interact with the negatively charged inorganic phosphate along its pathway through the protein interior. All these distinct and heterogeneous environments fit very well with the detection of a broad peak distribution of the cleaved phosphate.

The overall similarity between the ATP-bound and the ADP·P_i state raises the important question of whether both spectra represent the same state and have been misinterpreted. Indeed, it could be argued that the weak γP peak of the ATP-bound state is an artefact of the low signal-to-noise ratio and distracts from the somewhat broadened peak base around the αP resonance – a primary feature of the ADP·P_i state. Several observations can be made to counter this argument: First of all, the double mutant ND1L-R95G-E305Q has been selected for this reference measurement based on the pre-hydrolysis state-like spectra it produced under solution-state NMR conditions in the presence of the ATP regeneration system [55]. Secondly, a comparison with another pre-hydrolysis state, the ATPγS-bound spectrum, reveals a strong similarity shared between the respective αP and βP peaks in terms of chemical shifts and line widths (see Figure 46, Chapter 4). Only the terminal phosphate resonance is shifted to ~40 ppm due to the thiosubstitution, and no weak peak is observed between the main αP and βP resonances. These findings show that in the context of p97, αP and βP phosphate peaks of pre-hydrolysis nucleotides can indeed produce the aforementioned characteristics. Additionally, the thiosubstitution-induced downfield shift can only affect the γ-phosphate, so the absence of the small peak in the ATPγS-bound spectrum is logical. Finally, QM/MM calculations of the ATP-bound state generally agree with the experiment in terms of chemical shifts, although the calculations could technically also support a swapped assignment of the αP and γP chemical shifts [78]. This flipped assignment, however, would not be consistent with the relative peak intensities and peak separation of the αP and βP resonances in the ATP, ATPγS- and ADP·P_i bound spectra.

Nonetheless, the resemblance between the ADP·P_i and ATP bound spectra is striking, considering how the ADP-bound spectrum is distinctly different in appearance. This is actually surprising since, by nature, the ADP·P_i state consists of a cleaved phosphate and an ADP molecule. A possible argument to reconcile these observations is to discuss the role of Mg^{2+} in the binding pocket. Metal ions are needed to counteract the negative phosphate charges, especially of triple phosphates such as ATP. Published structures of ATPγS-bound p97 always show the presence of Mg^{2+}, while the metal ion is missing in ADP-bound structures (e.g. [32,36,43,75]). Thus, Mg^{2+} is expected to leave the binding pocket before ADP but after hydrolysis. Both stable ADP·P_i states (A and B) identified by the cryo-EM reconstruction and MD simulation provide locations for both the cleaved P_i and the Mg^{2+} ion [52]. Expectedly, immediately after ATP hydrolysis, the metal ion is still located inside the binding pocket and thus interacts with the newly formed ADP molecule and the cleaved P_i at the same time. Positioning of amino acid sidechains and Mg^{2+} specifically for interactions with ATP and for ATP hydrolysis will still be in place after immediate hydrolysis, partially explaining the similarity between the ^{31}P spectra of ATP and ADP·P_i bound states. After or prior to P_i release from the binding pocket, the metal ion leaves the active site so that only ADP remains. Under these conditions, the side chain positioning and the phosphodiester angle also change, affecting ADP phosphate shifts.

Another observation to point out is the presence of additional peaks in the full-length spectrum recorded under ATP turnover (Figure 30). As stated before, these additional, poorly resolved peaks must originate from the D2 domain, although their identity is unknown. A cryo-EM reconstruction of full-length wild-type p97 recorded in the presence of the ATP regeneration system detected an ADP·P$_i$ state in the D1 binding pocket, but ATP bound by the D2 domain [52]. This result aligns with our solid-state ^{31}P NMR spectra since the ND1L and full-length spectra show great overlap, suggesting ADP·P$_i$ binding in D1. The chemical shift range of the additional peaks falls between -1 and -13 ppm, in which both αP and γP resonances of a presumed ATP molecule would be expected.

Of course, a caveat needs to be made at this point. ^{31}P NMR spectra have been recorded on p97-Walker B mutants, and the presence of these mutations can affect both the bound nucleotide types and chemical shifts in comparison to the wild-type protein. For that reason, the additional peaks seen in the FL-E305Q-E578Q spectra can only be tentatively assigned to ATP since the cryo-EM data was recorded on FL-wt. Still, both ND1L-wt and ND1L-E305Q produce very similar protein spectra under solution-state NMR conditions [55], leading to the assumption that the introduction of the Walker B mutation does not substantially affect the protein conformation and the bound nucleotide found in ^{31}P NMR experiments.

To the best of our knowledge, our work presents the first detection of a post-hydrolysis ADP·P$_i$ state that unlike previous studies [69–73,82] includes proof that the captured state is authentic. An alternative, long-practised approach to measuring intermediate states is the usage of ATP mimics. While nucleotide analogues such as AMP-PNP, AMP-PCP and ATPγS mimic the pre-hydrolysis state by preventing or slowing down hydrolysis, analogues such as ADP·AlF$_x^-$ or ADP·VO$_4^-$ supposedly mimic a transitory hydrolysis state. Finally, post-hydrolysis states are mimicked with ADP, or ADP in combination with inorganic phosphate or sulphate [99].

Thus, using the hydrolysis transition mimic ADP·AlF$_4^-$ on DnaB, two narrow peaks were detected by ^{31}P NMR at -6.1 and -7.4 ppm and were assigned with a peak order opposite to the ADP state (i.e. identically to the peak order assigned in this work) [103]. However, the poor peak dispersion is substantially smaller than that of the ADP state. In contrast, the ADP·VO$_4^-$ bound to BmrA produced a spectrum with two types of bound ADP with varying βP shifts [99]. The nucleotide resonances were assigned following the usual ADP state assignment. Contrasting, ATP bound to a hydrolysis-deficient mutant of BmrA produced three major peaks with a peak dispersion similar to the ATP-bound state of p97 described in this thesis, but a flipped assignment of the αP and γP resonances [99].

All these examples lack evidence of the heterogeneous P_i environment, highlighting the importance of our work.

3.5.2 Challenges of solid-state ^{31}P NMR measurements

At this point, it can be beneficial to summarise the challenges and limitations of solid-state ^{31}P NMR measurements encountered during this work.

Phosphorus measurements benefit from the natural advantage of having a monoisotopic, NMR-active, spin ½ ^{31}P nucleus to work with, significantly simplifying the preparations of NMR-visible species. With their generally great chemical shift dispersion (~500 ppm), ^{31}P nuclei are also well suited for solid-state NMR measurements, where peak resolution is often an issue. The downside to this dispersion is that offset effects play a larger role, and transmitter frequencies must be selected carefully according to the region of interest.

Unfortunately, ^{31}P NMR measurements of biological samples can still suffer from poor peak dispersion even under solution-state NMR conditions since the inherent limitation is the overall similar chemical environment of phosphate groups, like nucleic acids or lipids [106,107]. Similar difficulties have been encountered in this work, especially in the context of assigning αP and γP resonances. ^{31}P chemical shifts of nucleotides depend on many factors, including bond lengths, bond angles and torsion angles, rendering assignments of protein-bound nucleotides ambiguous: for instance, we observed shifts perturbations of 2-6 ppm in comparison to the free nucleotide reference upon protein binding (Table 5). Thus, unless the phosphate group has been chemically modified (thiosubstitution) or is bound in a distinct matter (βP), peak assignments from 1D spectra require additional input.

This input could be, in principle, provided by multi-dimensional or filtered NMR experiments. However, ^{31}P nuclei are relatively insensitive with long relaxation times, and in our experimental setup even a single 1D experiment requires many hundreds or thousands of scans, i.e. many days of measurement time to obtain a decent signal-to-noise spectrum. Thus, multi-dimensional experiments (or those requiring additional magnetisation transfer steps) are highly impractical.

Also, there are many further limitations. Spectral signal-to-noise is directly proportional to the number of identical spins inside the sample tube – or the MAS rotor. Since the sample volume inside the rotor is limited, a higher compression rate (i.e. a sedimentation force) is necessary.

Ultimately, the amount of protein inside the rotor determines the number of ^{31}P nuclei. However, every 320 kDa ND1L hexamer is only binding six nucleotides (12 or 18 ^{31}P nuclei). One of the possible solutions to increase the number of nuclei inside the RF coil is to increase the rotor size. However, the increased rotor size comes at the cost of a reduced spinning speed, which (aside from broader lines) introduces overlap with spinning side bands. On an 800 MHz spectrometer, a 45 kHz spinning speed (1.3 mm rotor) corresponds to ~139 ppm. Thus, spinning side bands fall outside of the spectral width. 17.5 kHz spinning speed (3.2 mm rotor), however, corresponds to 54 ppm. Peaks with a chemical shift of -10 ppm produce side bands at 44 ppm, directly in the region of the thiosubstituted phosphates, which, in turn, produce spinning side bands between -10 and -20 ppm. In summary, while spectra recorded in 3.2 mm rotors provide higher signal-to-noise, their interpretation is substantially more complicated due to spinning side bands and increased line widths.

A separate issue encountered during this work is the time limit imposed by the ATP regeneration system. ATP turnover measurements require constant ATP regeneration using PEP as fuel. Once all PEP is converted to pyruvate, the nucleotide composition in solution changes and that particular measurement cannot be continued. Over the course of this work, we tried several times unsuccessfully to restart the ATPase reaction by adding new PEP into the packed rotor but never achieved the desired result. A possible explanation for the failure of this approach is that the diffusion of PEP inside the protein sediment is too slow to regenerate all ADP at once, a requirement for a symmetric homohexamer.

There are two general ways to increase the measurement time under ATP turnover. The first one is to reduce the ATPase rate. An estimation of the PEP consumption rate from an ATPase assay is a necessary prerequisite when deciding on the nucleotide (analogue) or construct, but it may not directly reflect the true rates observed inside the rotor: kinetic processes in gels or sediments can be different from those in free solutions. Using slow hydrolysing constructs such as Walker B mutants is often the only way to reduce the ATPase rate so that PEP can last for a sufficiently long time, even though, of course, measurements on the wild-type protein are preferred. Moreover, the addition of protein-specific inhibitors can also help, such as thiosubstituted nucleotide analogues (ATPγS) or adjusting buffer compositions, e.g. via higher salt concentrations in case of p97 [56]. Finally, reduced temperatures are advantageous for two reasons: for the reduced ATPase rate and for decreased protein mobility, improving the efficiency of dipolar-based magnetisation transfer. This is especially important for capturing the transient, highly dynamic ADP·P_i state. This temperature effect can be seen in Figure 27, where the CP spectrum recorded at 20°C has a significantly lower signal-to-noise than that recorded at 7°C (although the number of scans is also partially reduced).

The second way to prolong measurement time is by increasing the PEP concentration. This concentration cannot be indefinitely increased, as inorganic phosphate is an inhibitor of the pyruvate kinase [108,109], thus rendering extra PEP (>200 mM) useless. Moreover, additional PEP increases the ionic strength inside the rotor. While this has the benefit of slowing down the ATPase rate, it also lowers the spectral signal-to-noise and can even cause problems with probe matching.

Thus, it is important to optimise the reaction setup before the actual measurement and to constantly refrigerate all components (incl. the ultracentrifuge and filling tools) during setup and filling. One should consider the rotor filling time itself: sedimentation speed depends on molecule size and takes approx. 20-48 h for ND1L hexamers and 12-15 h for full-length hexamers at 130,000 g. During rotor filling, the ATPase reaction is already running, consuming PEP and decreasing the possible measurement time. Concentrating the protein beforehand to lower volumes decreases the sedimentation time but also limits the total reaction time due to an absolutely smaller amount of PEP in the solution. Similarly, a higher protein mass and density in the rotor (e.g. in a 3.2 mm rotor) can decrease diffusion rates, possibly slowing down hydrolysis rates. However, at the same time, the additional protein mass decreases the amount of liquid volume, decreasing the total amount of PEP while at the same time increasing the overall ATPase rate due to increased protein mass. The influence of these effects is difficult to predict and control, but one should be aware of them.

3.5.3 Alternative approaches to support peak assignment

In order to support peak assignment of the ADP·P_i state, it is sufficient to prove the identity of either αP or βP phosphate.

For once, instead of using chemically modified nucleotides (Chapter 4), one could use isotope-labelled ATP and show specific contacts: ^{13}C-labelled ATP could be used to show contacts between the ribose C5’ and the α-phosphate. Either a ^{13}C-filtered HCP 1D experiment could be employed [106], or a ^{13}C-^{31}P-TEDOR setup could be used to determine the distance between the ribose carbons and the respective phosphorus nuclei. The feasibility of the latter approach has been shown during the assignment of a small RNA via solid-state NMR techniques [110].

Another procedure involves α or β-^{17}O-labelling of the phosphate groups. Depending on the labelling scheme, the splitting pattern in the ^{31}P spectrum alone could provide sufficient information to assign the peaks [97,111].

Commonly, ^{31}P-^{31}P-2Ds in the form of DARR or PDSD are recorded to assign peaks of bound nucleotides [100,101]. Provided sufficient signal intensity, observed contacts between βP and the cleaved P_i would be sufficient for peak identification.

Finally, an alternative approach would involve a magnetisation transfer from a specific protein sidechain (e.g. an arginine or lysine) to the phosphate groups. The feasibility of this approach has been shown via a ^{15}N{^{31}P}-REDOR experiment, showing arginine side chains-RNA phosphate backbone contacts [112], while in another study, assigned lysine and arginine side chains were correlated to the phosphates of bound ATP via NHHP and CHHP experiments [113].

The main limitation of all aforementioned methods lies in the requirement of ample initial magnetisation. Considering that the ADP·P_i spectrum was acquired in ~130,000 scans, it appears unlikely that any pulse sequence requiring additional magnetisation transfer steps would ultimately provide sufficient signal-to-noise.

Chapter 4: Usefulness of ATP analogues in structural studies

4.1 Introduction

As discussed in Chapter 2, four major protein and/or nucleotide states form the ATP hydrolysis cycle: apo state (nucleotide-free), pre-hydrolysis ATP bound, post-hydrolysis ADP·P_i bound and ADP bound. Major differences between these states regarding the position of the NTDs have been discussed in the introduction as well.

In general, protein and nucleotide conformational states can be inaccessible for studies under native conditions, e.g. due to short lifetimes of the states of interest or due to their transient nature. In the case of p97, investigations of ATP-bound wild type protein are difficult, since the nucleotide is quickly hydrolysed to ADP upon binding. Thus, to probe otherwise inaccessible biological states, the usage of nucleotide analogues (also referred to as mimics) can be helpful.

A nucleotide analogue is only as useful as it is biologically relevant. Although many different analogues have been used to study ATPases (ATPγS, AMP-PNP, AMP-PCP, ADP-BeF_x, ADP-AlF_x) [99], it remains controversial whether the usage of such analogues is pertinent since faulty conclusions may be drawn by creating artificial protein states. Additionally, as is the case with the thiosubstituted nucleotides, analogues can suffer from decreased chemical or thermal stability.

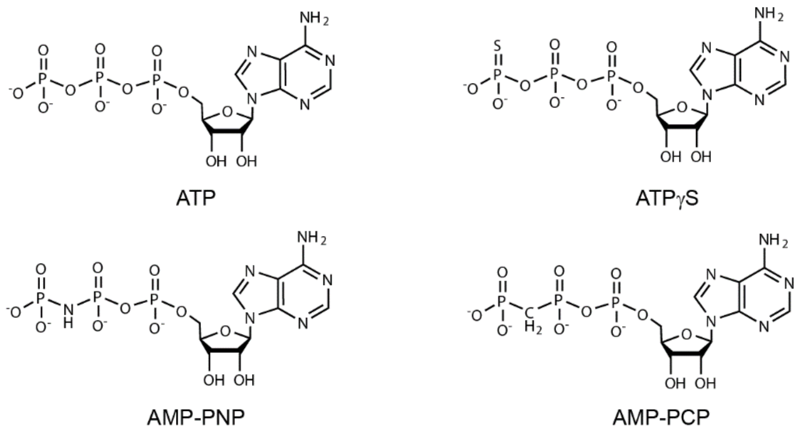
Nucleotide analogues can also be used to support studies performed on native nucleotides. In Chapter 4, peak assignments have been driven by comparison to literature values and reference spectra. However, for structural interpretations, an unambiguous proof of peak identity is required. This can be achieved by using thiosubstituted ATP analogues. This approach can yield three advantages: the thiosubstitution of an oxygen 1) confers chirality to the phosphate group, thus supporting mechanistic studies [114], 2) shifts phosphate resonances downfield by approximately 40 ppm in comparison to unsubstituted chemical shifts, facilitating peak assignment [89] and 3) often creates a hydrolysis resistant nucleotide analogue which can help to stabilise intermediate protein states [99].

In this chapter, first, we compare the ability of the three nucleotide analogues (Figure 32) to stabilise the pre-hydrolysis ATP-like state and the resistance of these molecules to thermal degradation or autohydrolysis at elevated temperatures employed during solution-state NMR experiments.

In the second part, complementary measurements with thiosubstituted ATP analogues are analysed to support NMR peak assignments presented in Chapter 3. Here, we first study the effect of nucleotide binding to p97 in solution and then perform solid-state ^{31}P NMR measurements.

4.2 Hydrolysis-resistant ATP mimics in structural studies of p97

A principal feature of the p97 ATP-bound state is the positioning of the N-terminal domain above the D1 plane, which has been observed for the wild-type protein bound to ATPγS [32]. However, ATPγS is not strictly non-hydrolysable but merely hydrolysis-resistant, i.e. it slows down the hydrolysis rate of the respective ATPase [99]. Its general usefulness as an ATP mimic can be attributed to its minimal chemical modification in comparison to ATP: the exchange of non-bridging oxygen by a sulphur atom. Both elements are members of the chalcogene group, and both elements are electronegative compared to phosphorus.


ATP

ATPγS

AMP-PNP

AMP-PCP

Figure 32. Chemical structures of hydrolysis-resistant ATP analogues. ATP is chemically modified to create a hydrolysis-resistant analogue. With ATPγS, a non-bridging oxygen of the γ-phosphate is replaced with a sulphur atom. In the cases of AMP-PNP and AMP-PCP, β,γ-phosphate cleavage is prevented by replacing the bridging oxygen with an imido (AMP-PNP) or methylene (AMP-PCP) group.

The inability of two mutants (ND1l-D304N and ND1L-F360P, see Section 5.2.3) to respond to ATPγS binding in a wild-type-like way led to the search for additional hydrolysis-resistant nucleotide analogues to control for a possible unusual ATPγS-processing by p97. Favoured non-hydrolysable nucleotide analogues aside from ATPγS are AMP-PCP and AMP-PNP (Figure 32) [99], and even multiple p97 structures have been reported using specifically AMP-PNP [75,115,116].

4.2.1 Thermal stability of hydrolysis-resistant nucleotide analogues

Hydrolysis of these nucleotide analogues can be directly followed in ^{1}H 1Ds (Figure 33). Peaks of AMP-PNP are sufficiently separated from the resonances of its hydrolysis product AMP-PNH$_2$, both for ribose (~4.9 ppm) and H8 (~8.7 ppm) protons. In contrast, ATPγS and ADP ribose peaks are well separated around 4.9 ppm but lack resolution at 8 ppm. This resolution allows to follow the hydrolysis of both nucleotide analogues in solution as they slowly degrade at 40 and 50°C. In contrast, AMP-PCP appears to be non-hydrolysable since no additional peaks from the likely hydrolysis product, AMP-PCH$_3$, are detected over the course of the experiment at 50°C.

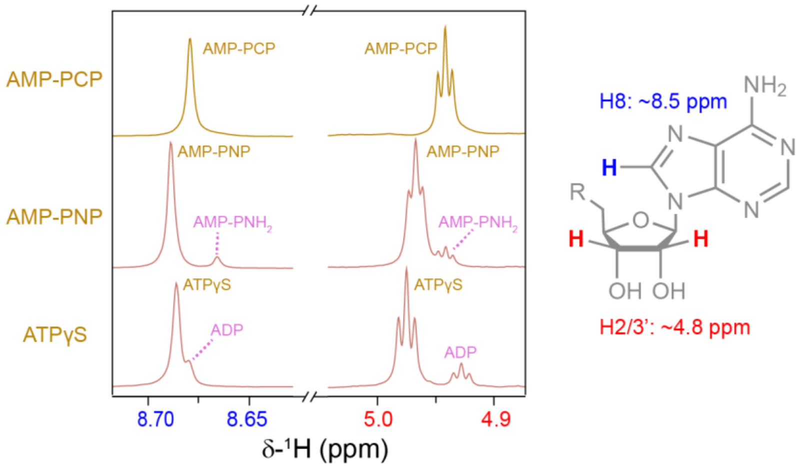


Figure 33. Reaction control of hydrolysis-resistant ATP analogues. All nucleotide analogues can be identified in ^{1}H 1Ds by H8-adenine and ribose signals around 8.7 (left) and 4.9 ppm (right), respectively. AMP-PCP shows no signs of hydrolysis (neither enzymatic nor thermal) and gives rise to a single peak over the course of the whole measurement. AMP-PNP is not thermally stable and degrades over time at elevated temperatures. Peak separation between the ATP mimic and its hydrolysis product, AMP-PNH$_2$, is great for the H8 proton and sufficient for the H2/3' protons. Peaks of ATPγS and its hydrolysis product ADP are well separated around 4.8 ppm, but lack an effective resolution at 8.7 ppm. All presented spectra lack a consistent internal calibration, so that the reported chemical shifts are relative. Spectra were measured at 50°C in the absence of the ATP regeneration system.

However, ultimately, the origin of hydrolysis cannot be determined from this data alone. Whether ATPγS is hydrolysed by p97 or simply undergoes autohydrolysis at the elevated temperatures of the experiment is discussed in greater details in Section 5.6.4.

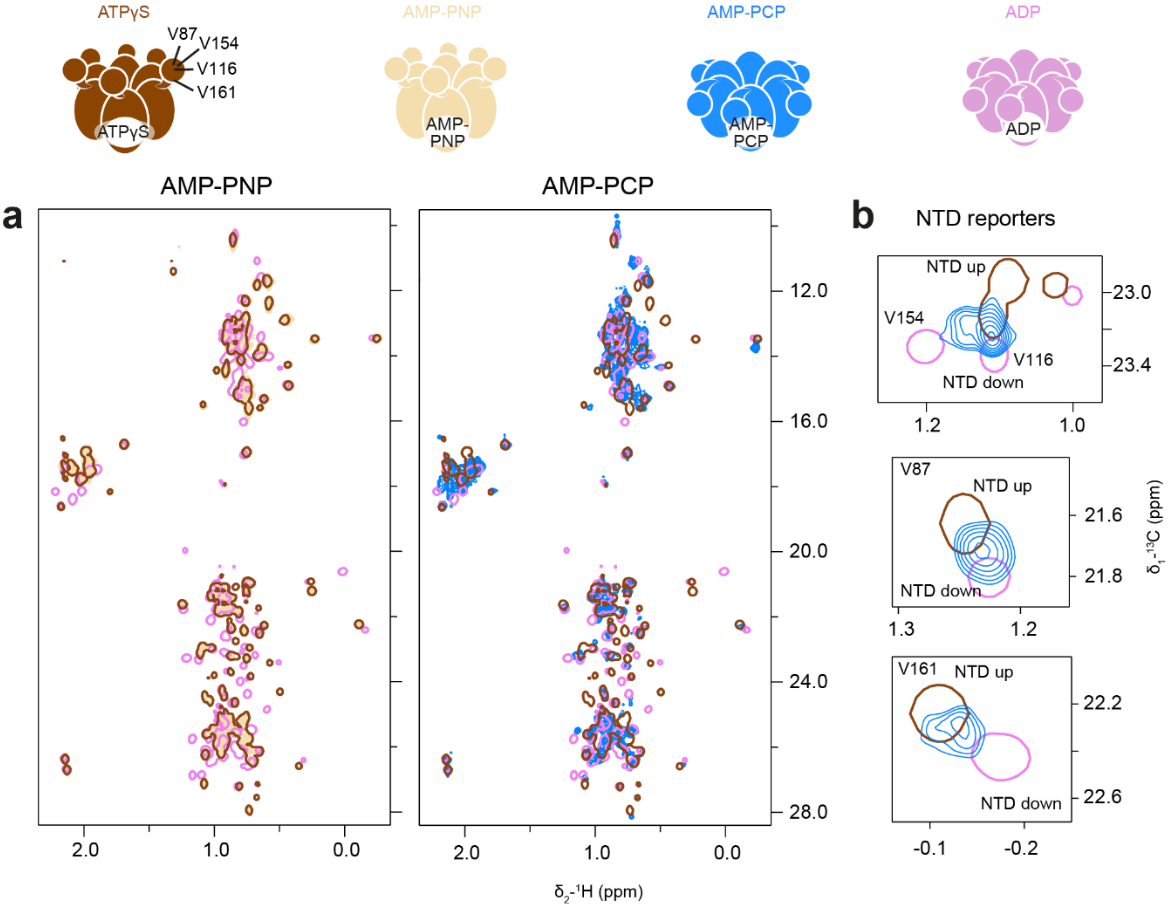


Figure 34. AMP-PNP stabilises the pre-hydrolysis conformation of p97, while AMP-PCP does not. Methyl TROSY spectra of ND1L-wt in the presence of hydrolysis-resistant nucleotide analogues. **a**) Spectra of p97 bound to AMP-PNP are highly similar to those measured in the presence of ATPγS, while binding to AMP-PCP induces an ADP-like state in p97. **b**) Zoom in on NTD residues V87, V116, V154 and V161 in the AMP-PCP bound state indicate an intermediate state of the N-terminal domain similar to a fast exchange between the up and down positions. All spectra were recorded at 50°C with a 5 mM nucleotide concentration, but the Mg^{2+} concentration varies from 0 mM (ADP) to 2 mM (AMP-PNP, AMP-PCP) to 5 mM (ATPγS). Spectra were partially redacted to hide spectral noise.

4.2.2 Behaviour of p97 in the presence of nucleotide analogues

In Figure 34, methyl TROSY spectra of ND1L-wt in the presence of AMP-PNP and AMP-PCP are shown. The enzyme seems to recognise AMP-PNP as a pre-hydrolysis type nucleotide, as the spectrum overlaps almost perfectly with the one recorded in the presence of ATPγS. In contrast, the protein misinterprets AMP-PCP for an ADP or ADP·P_i nucleotide since the spectrum recorded in its presence shares strong similarities with the ADP state.

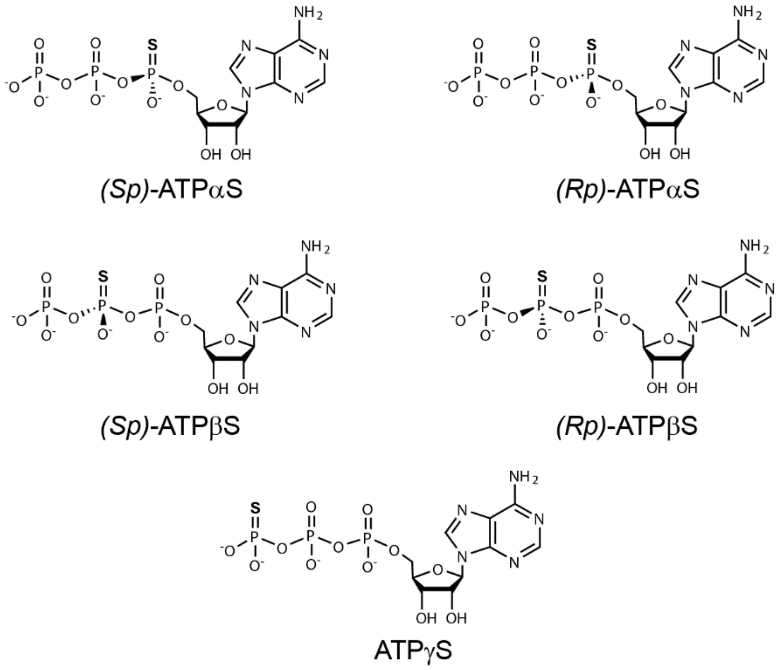
At 50°C, the N-terminal domain in the AMP-PCP bound state seems to indicate increased mobility reminiscent of fast exchange, as the respective peak positions indicate an intermediate state between up and down (Figure 34b). In summary, AMP-PNP was selected as a second ATP mimic alongside ATPγS, while AMP-PCP was disregarded in the following work. This is somewhat unfortunate, as the thermal stability of the latter would have been of great use for measurements at 50°C.

4.3 Nucleotide assignment using thiosubstituted ATP analogues

4.3.1 Limitations of thiosubstituted nucleotide analogues

ATP, with its three phosphate groups, has three bridging and seven terminal oxygen atoms. Usually, a terminal oxygen atom is replaced by a sulphur atom, forming three general types of thiosubstituted ATP: ATPαS, ATPβS and ATPγS (Figure 35). The most commonly used thiosubstituted ATP analogue is ATPγS which is a slowly hydrolysable nucleotide. In general, it is considered to be a useful ATP mimic, able to take multiple conformations and adapt to the nucleotide-binding pocket [99]. In p97, the binding of ATPγS stabilises the protein in its pre-hydrolysis/ATP-like state, enabling crystallisation, cryo-EM and NMR studies (e.g. [32,54,66]).

The other two types are less commonly used, and, to the best of my knowledge, the effects of their binding have not been published for any AAA+ ATPase. A substantial problem arises from the fact that thiosubstitutions at αP or βP positions introduce a new *P*-chirality centre (which is lacking for ATPγS). Since oxygen is considered to be a hard atom, but sulphur is soft, the interaction with metal ions can be affected [79,114]. In highly specialised environments such as protein binding pockets, where metal ion location is mostly predetermined by amino acid side chain positioning, a new *P*-chirality can affect nucleotide binding modes or even completely inhibit biological functionality. Thus, compatibility with p97 has to be checked for both *Sp*- and *Rp*-enantiomers. Additionally, for measurements involving the ATP regeneration system, compatibility with the pyruvate kinase has to be verified since it needs to process the respective ADP analogues (and their diastereomers) – a problem not encountered with ATPγS. As mentioned above, no data has been published on the binding effects of neither ATPαS nor ATPβS on p97 or any AAA+ ATPase, for that matter. Myosin, a P-loop ATPase, has been shown to hydrolyse both αP-diastereomers of ATPαS [117], but *Rp*-ADPαS is a known inhibitor of rabbit muscle pyruvate kinase [114,118–120]. Additionally, studies disagree on whether *Sp*-ADPαS is a non-competitive inhibitor or just a slow substrate of the pyruvate kinase [114,118–120], but since the *Rp*-diastereomer is certainly inhibitive, the *Sp*-diastereomer was chosen for further studies. The pyruvate kinase is reported to be able to preferentially produce the *Sp*-ATPβS diastereomer from the non-*P*-chiral ADPβS in a ratio of 80:20 [119], further complicating the reaction nucleotide composition. Moreover, neither diastereomer of ATPβS was commercially available, thus requiring an initial regeneration of ADPβS into its ATP analogue.



(Sp)-ATPαS

(Rp)-ATPαS

(Sp)-ATPβS

(Rp)-ATPβS

ATPγS

Figure 35. Chemical structures of thiosubstituted ATP analogues. Thiosubstitution at the αP and βP positions adds a P-chirality to the nucleotide, forming *Sp*- and *Rp*-ATPα/βS analogues. In contrast, ATPγS is not P-chiral.

The utility of both *Sp*-ATPαS and a pyruvate kinase (PK)-regenerated ADPβS was initially tested via an ATPase assay but remained inconclusive (Figure 36). In the *Sp*-ATPαS containing assay, NADH was consumed equally fast in all samples, including the protein-free control sample. This implies that either *Sp*-ATPαS is not a substrate of p97 or *Sp*-ADPαS is not a substrate of the pyruvate kinase (or both). From the assay, it is impossible to determine whether the ATP analogue was hydrolysed but remained in its *Sp*-ADPαS form because it is not a substrate of the pyruvate kinase or whether it was not hydrolysed in the first place. On the other hand, the ATPase assay involving ATPβS failed owing to unclear reasons. In the initial step, ADPβS was incubated with the regeneration system, but the NADH concentration kept decreasing during the equilibration step, never reaching a plateau with all ATPβS formed. The later addition of p97 to the buffer did not affect the rate of NADH consumption, similar to the *Sp*-ATPαS samples. No conclusion can be drawn from this experiment, as it remains unclear whether ADPβS is not a substrate of the pyruvate kinase or whether p97 could not hydrolyse the regenerated ATP analogue.

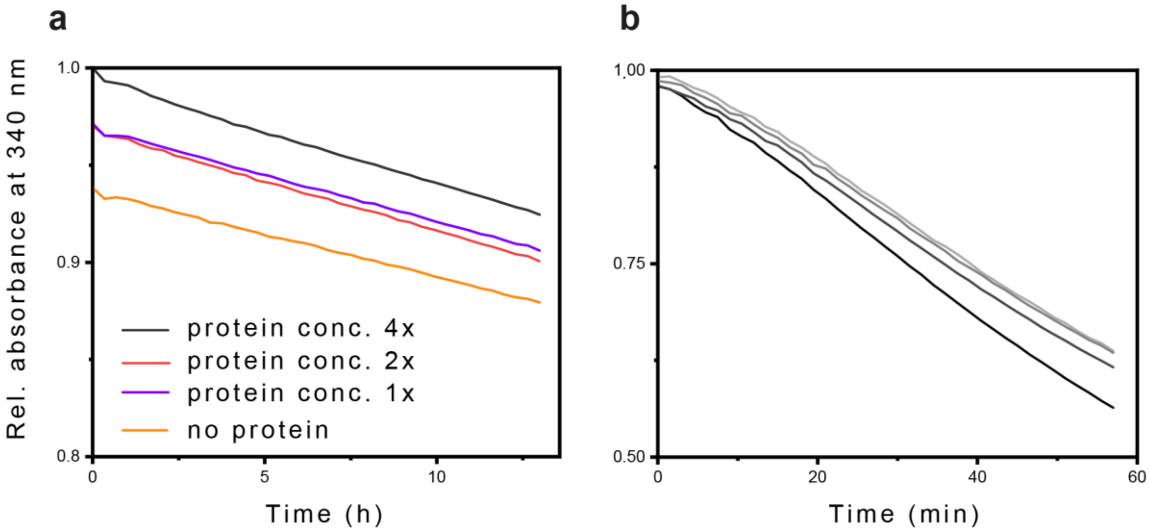


Figure 36. ATPase assay of thiosubstituted nucleotide analogues. a) ATPase assay of ND1L-wt with *Sp*-ATPαS at 27°C. NADH is consumed equally fast in all samples, regardless of the protein concentration. **b**) NADH consumption during regeneration of ADPβS at 37°C. Consumption of NADH continues equally fast in all samples. Subsequent addition of p97 did not affect the rate of NADH consumption (not shown).

4.3.2 Analysis of thiosubstituted nucleotide binding by solution-state NMR

Reaction control using ^{1}H 1Ds

Prior to the analysis of the protein NMR spectra, it is a prerequisite to be able to monitor nucleotide presence in solution as a form of reaction control. ^{1}H 1Ds are sufficient to identify nucleotide turnover over time – or the lack thereof.

In Figure 37, regions from ^{1}H 1Ds of thiosubstituted nucleotides are presented. Although no conclusion can be drawn from the spectra alone regarding the cause of nucleotide hydrolysis (thermal degradation, protein-catalysed or autohydrolysis), assigning resonances to specific nucleotides is the first step to understanding their interaction with p97.

Sp-ATPαS

In the absence of PEP, *Sp*-ATPαS slowly transforms to a different nucleotide – presumably *Sp*-ADPαS (whether p97 is involved in hydrolysis or it is simply a product of autohydrolysis is discussed below). The peaks of these entities are well separated for both H8 and H2/3’ protons and the left peaks lose intensity over time while the right peaks become stronger.

ATPβS

The same analysis could only be partially performed for ADPβS, as the reaction setup at 4°C contained a higher percentage of water, preventing analysis of peaks located near the water resonance. However, after incubating ADPβS with the regeneration system (Figure 37b, top and middle), a second peak starts to emerge, which is assigned to *Sp*-ATPβS. The assumption that the produced species is the *Sp*-diasteromer is based on previous reports showing a preference of the pyruvate kinase to produce *Sp*-ATPβS from ADPβS [118,119]. Unfortunately, it appears that ATPβS acts as an inhibitor to the pyruvate kinase, as the nucleotide regeneration halted after processing only approximately 50% of the initial ADP analogue (~4 mM). Eventually, the reaction ran (almost) to completion after the reaction volume had been diluted fourfold, thus reducing the concentration of the inhibitory *Sp*-ATPβS (Figure 37b, bottom). At the end of the reaction, the nucleotide ratios were determined to be 83% *Sp*-ATPβS, 12% *Rp*-ATPβS and 5% ADPβS, with all peaks being reasonably well separated from each other. Considering that, to a minor extent, the *Rp*-diastereomer is produced as well, it remains unclear which diastereomer acts as an effective inhibitor.

ATPγS

Finally, the reaction control for ATPγS requires more complicated analysis due to signal overlap and the presence of multiple species at the same time in solution: in the absence of a regeneration system, ATPγS and its hydrolysis product, ADP, can be easily monitored using peaks of the ribose H2/3’ protons, while the resolution is lacking for the H8 peaks (Figure 37c, top). However, in the presence of the regeneration system, a mixture of nucleotides emerges since the hydrolysis product ADP is converted back to ATP. In this case, ATPγS and ATP can be told apart using the H8-resonances around 8.7 ppm, while the respective ribose peaks (~4.8 ppm) overlap almost completely (Figure 37c, bottom).

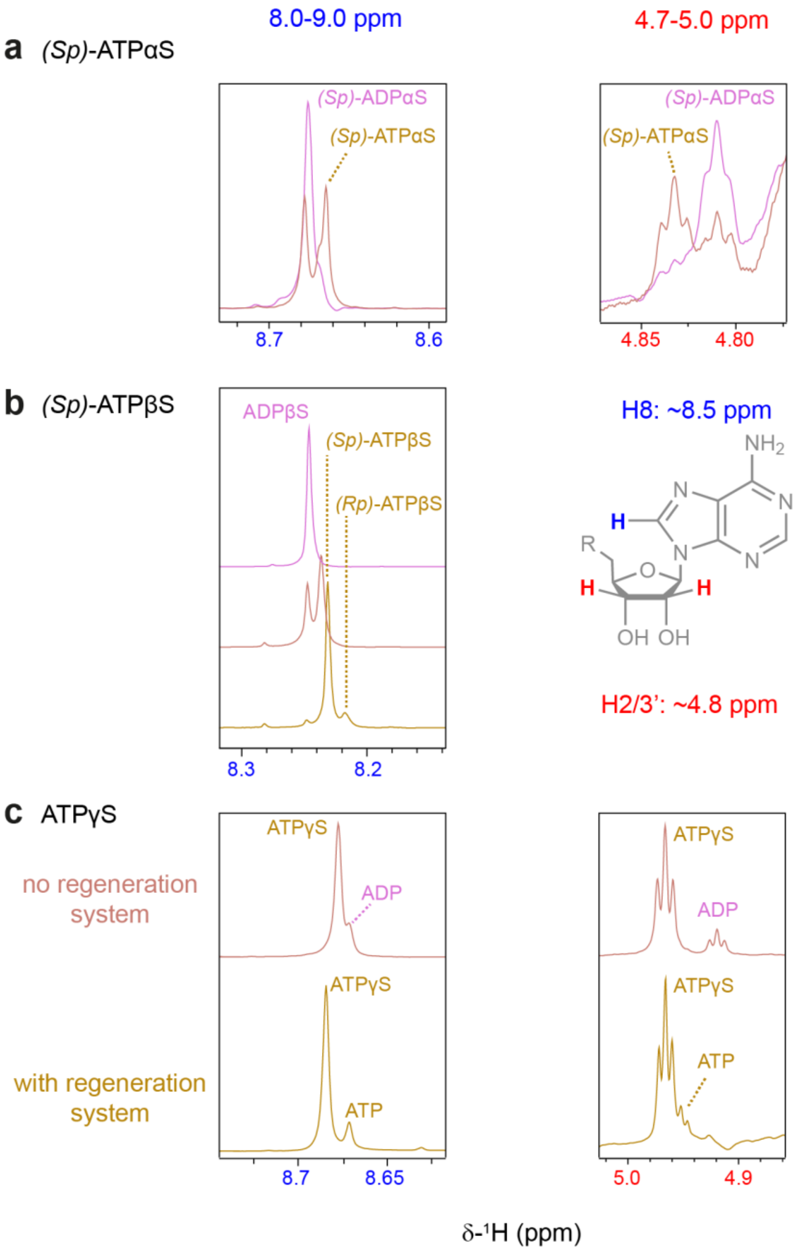


Figure 37. Reaction control of thiosubstituted nucleotide analogues. All thiosubstituted nucleotide analogues can be identified in ^{1}H 1Ds by adenine and ribose signals around 8.7 and 4.9 ppm, respectively. **a**) *Sp*-ATPαS is slowly hydrolysed to *Sp*-ADPαS at 37°C over the course of ~4 h in the absence of a regeneration system (but in the presence of ND1L-E305Q). This conversion can be monitored using resonances at 8.7 ppm (left) and 4.8 ppm (right). **b**) ADPβS is regenerated to *Sp*-ATPβS by the pyruvate kinase at 4°C over the course of ~48 h, but the reaction stops after converting ~50% of the ADP analogue (initital concentration: 8.7 mM). After diluting the reaction volume fourfold, regeneration continues. The final nucleotide ratios present in solution: 83% *Sp*-ATPβS, 12% *Rp*-ATPβS and 5% ADPβS. Peak dispersion at 4 ppm remains unknown, since the large amount of H_2O in the sample prevented an analysis of resonances in its proximity. **c**) Top: Peaks of ATPγS and its hydrolysis product ADP are well separated around 4.8 ppm, but lack resolution at 8.7 ppm. Bottom: In the presence of a regeneration system both ATPγS and ATP are present. Here, peaks are well dispersed around 8.7 ppm, but overlap at 4.8 ppm. All presented spectra lack a consistent internal calibration, so that the reported chemical shifts are relative.

Measurements of the protein state upon binding thiosubstituted ATP analogues

The motivation to perform experiments with thiosubstituted nucleotide analogues was to obtain evidence to support assignments of ADP·P_i state resonances. The first question to tackle is, thus, whether p97 shows an ADP·P_i-like or ATPγS-like state in methyl TROSY spectra in the presence of these analogues and a regeneration system.

Sp-ATPαS

In the case of *Sp*-ATPαS, ND1L-E305Q overall adopts an ADP·P_i-like conformation (Figure 38a, left). Two observations have to be made: 1) Methyl probes which are part of the NTD indicate a mixed conformation, with both up- and down positions being detected (Figure 38b). However, it is unclear whether this mixed positioning originates from the slow exchange due to NTD mobility or because of mixed hexamers (with some NTDs locked in the up-conformation and some NTDs in the down-conformation). 2) Key probes in proximity to the binding pocket vary in their identification of the bound nucleotide, e.g., the peak position of I206 is similar to that of the ATPγS or ADP·P_i state, but the L268 is similar to the ADP state (Figure 38c). After PEP is depleted and all *Sp*-ATPαS is converted to *Sp*-ADPαS, p97 assumes an ADP-like state (Figure 38a, right). Most importantly, I206 now shifts to its ADP position, thus being clearly able to distinguish between the ATP- and ADP-analogue.

This leads to several conclusions. First of all, the change of the I206 peak position after the depletion of PEP implies that the regeneration system was constantly converting *Sp*-ADPαS to *Sp*-ATPαS, and a possible pyruvate kinase inhibition did not have an obvious effect.

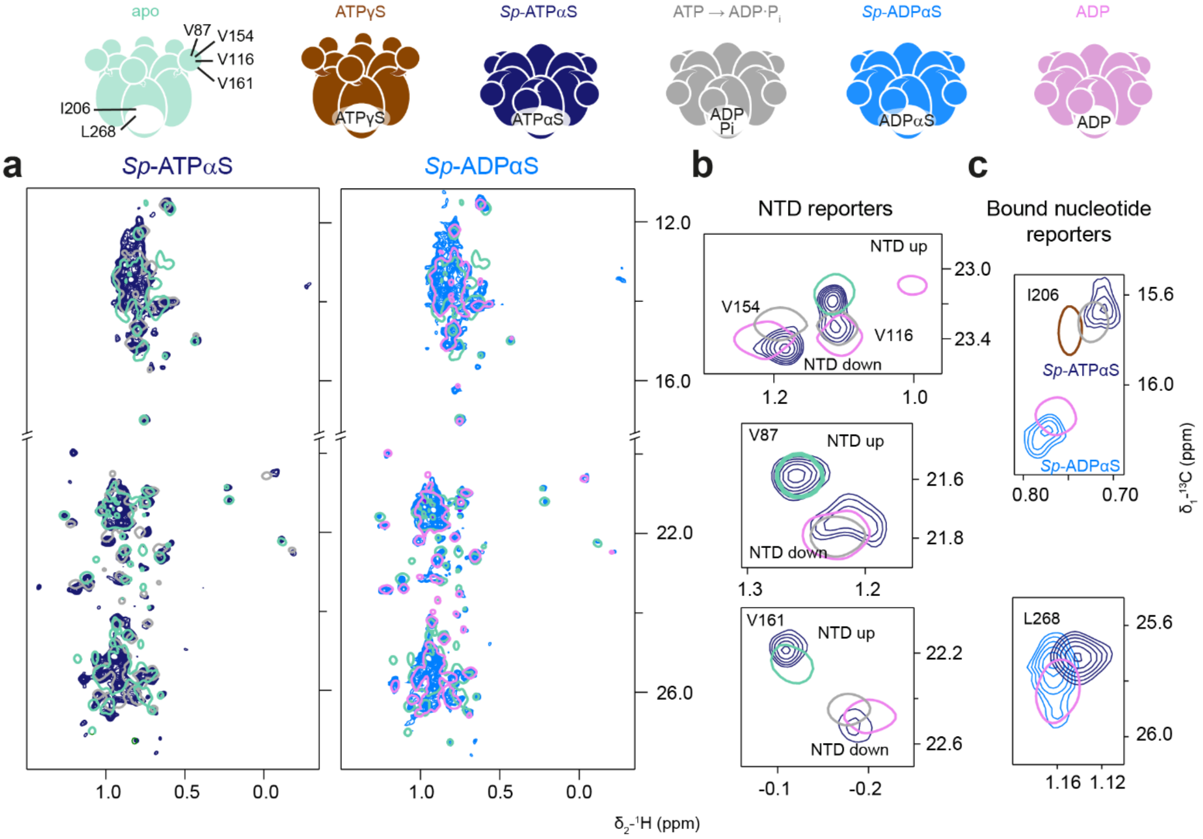


Figure 38. ND1L-p97 adopts an ADP·Pi-like conformation in the presence of *Sp*-ATPαS. Methyl TROSY spectra of ND1L-E305Q in the presence of *Sp*-ATPαS and the regeneration system show a general similarity to the ADP·Pi state (**a**), but the NTD displays mixed up- and down-conformations (**b**) due to the presence of apo-state protein. Similarly, *Sp*-ADPαS-bound p97 displays a mixture of ADP- and apo-state conformations (**a**). **c**) Key residues close to the binding pocket interpret the bound nucleotide either as an ADP·Pi (I206) or ADP analogue (L268). This mixed nucleotide interpretation does not originate from mixed bound species, as I206 is clearly able to differentiate between bound *Sp*-ATPαS and *Sp*-ADPαS. The *Sp*-ATPαS spectra were recorded at 37°C, the ADP·Pi reference at 50°C and all other spectra at 40°C.

Secondly, both I206 and L268 show differences between the ATP- and ADP-analogue binding and the detection of the L268 peak in the ADP·Pi state indicates a non-native binding of the ATP analogue. This is supported by the observation that the NTD is positioned in a mixed conformation, which could be attributed to a misinterpretation of the protein regarding the nature of the bound nucleotide. Another interpretation of the mixed NTD positioning would be to assume a slow exchange between nucleotide bound and -free (i.e. the apo) state. In the apo state, the N-terminal domain is located above the D1 plane, similarly to the ATPγS-bound conformation. The ATPase could possess lower affinities for the nucleotide analogue.

Overall, *Sp*-ATPαS is clearly hydrolysed and regenerated while PEP is present. Under these conditions, the observed protein state shares similarities with the ADP·P_i state formed in the presence of native ATP.

ATPβS

A similar picture emerges in the measurement of ND1L-E305Q with ATPβS in the presence of the regeneration system. Although the spectra show high similarities to the ADP·P_i state, again, the NTD is showing a mixed up-down-conformation, which, similarly to *Sp*-ATPαS, could indicate a partial apo state (Figure 39). Additionally, methyl probes close to the binding pocket indicate the presence of an ATPγS/ADP·P_i (I206) or ADP (L268) analogous conformation, with the latter even showing two peaks (Figure 39c). Since no significant PEP consumption was observed over the course of the measurement, it is unknown how these residues would react to the binding of pure ADPβS. Moreover, since both ATPβS diastereomers are present in solution (although in vastly different concentrations) and the binding affinities are unknown, it is possible that both diastereomers are bound to a similar extent, which would explain the observed peak splitting of L268. Conversely, hydrolysis of ATPβS inside the binding pocket and the formation of a stable ADPβS·P_i state (incapable of product release) could explain the lack of PEP consumption but would require I206 to misinterpret the type of bound nucleotide. Either way, it appears that, similarly to *Sp*-ATPαS, this thiosubstituted analogue is bound non-natively by p97, and the protein conformation is ambiguous with some resemblance to the ADP·P_i state.

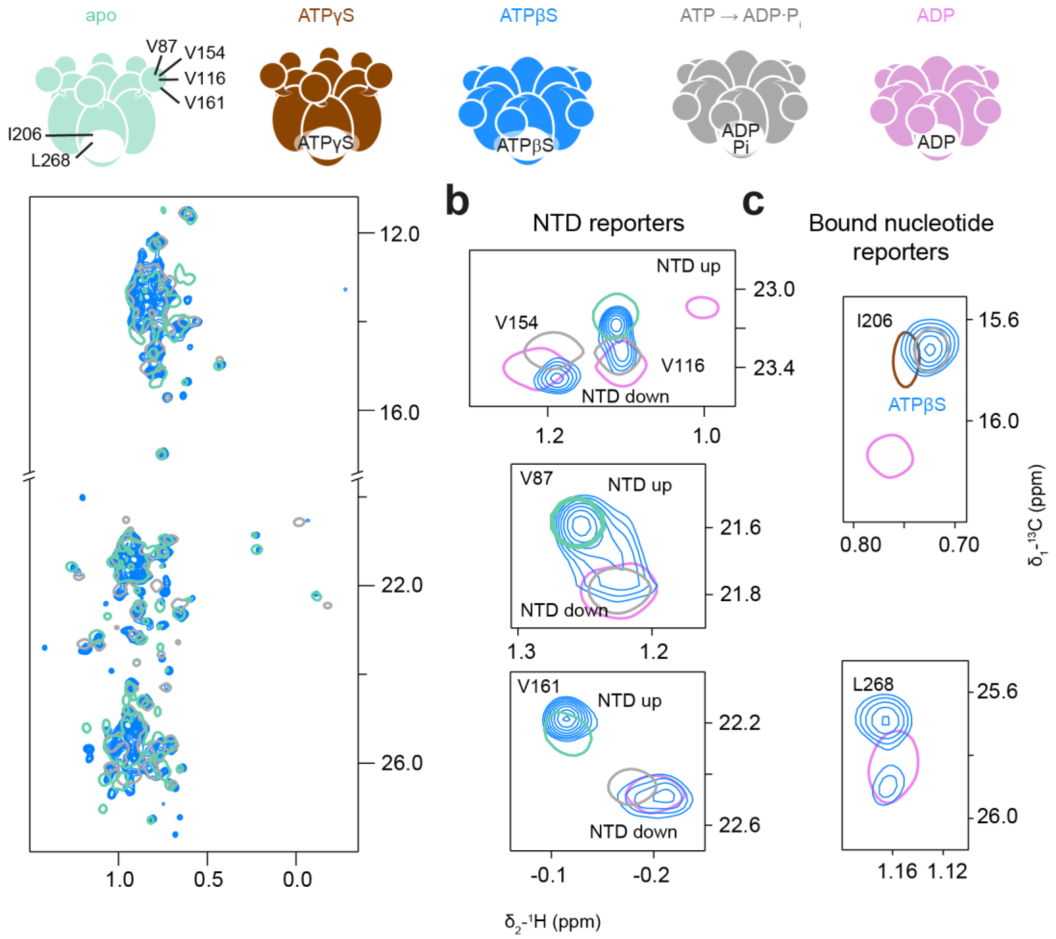


Figure 39. ND1L-p97 adopts an ADP·P$_i$-like conformation in the presence of ATPβS. Methyl TROSY spectra of ND1L-E305Q in the presence of *Sp*-ATPβS and the regeneration system show a general similarity to the ADP·P$_i$ state (**a**), but the NTD shows mixed up- and down-conformations (**b**) possibly due to presence of apo-state protein. **c**) Additionally, key residues close to the binding pocket interpret the bound nucleotide either as an ADP·P$_i$ (I206) or ADP analogue (L268). The *Sp*-ATPβS spectra were recorded at 37°C with a 2 mM nucleotide concentration, the ADP·P$_i$ reference at 50°C and all other spectra at 40°C.

ATPγS

In contrast, the effect ATPγS binding has on p97 has been studied extensively (e.g. [32,63,66]): as a slowly hydrolysable analogue of ATP, binding of ATPγS mimics the pre-hydrolysis state with the N-terminal domain located above the D1 ring. The resulting methyl TROSY NMR spectra differ significantly from the ones of the ADP state and are used as a reference throughout this thesis (see Section 2.5). In combination with a regeneration system and ATPγS, however, ND1L-E305Q adopts a regular (homohexameric) ADP·P$_i$ state (Figure 40). More precisely, the resulting state should be rather described as ADP·SP$_i$ state, considering that an inorganic thiophosphate is cleaved off. The implication of this observation is that in the ATPase cycle involving ATPγS, the rate-determining step remains the product release. This would make ATPγS a biologically relevant analogue of ATP and an excellent tool of structural studies, considering the reduced PEP consumption during the measurement. A major caveat of this observation is the fact that the hydrolysis product of ATPγS, ADP, is regenerated to regular ATP, leading to a nucleotide mixture over time. Although binding affinities of ATPγS are reported to be even higher than those of ATP [47], the effect ATP plays in the formation of this ADP·SP$_i$ state is not understood. This is further discussed in detail in Section 5.6.4.

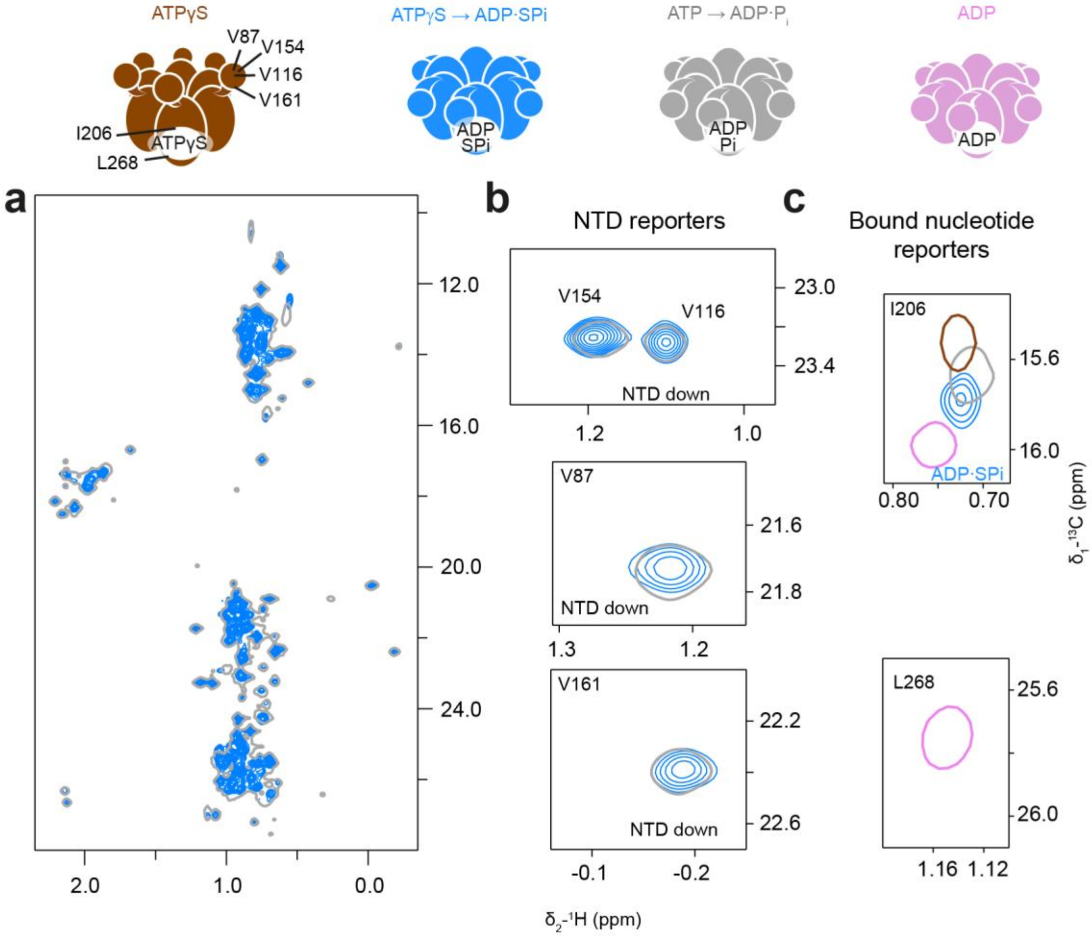


Figure 40. ND1L-p97 is found in the ADP·P$_i$ state in the presence of ATPγS and the regeneration system. Methyl TROSY spectra of ND1L-wt in the presence of ATPγS and the regeneration system show a highly similar ADP·P$_i$ state (**a**,**b**). **c**) A key residue close to the binding pocket (I206) detects a bound nucleotide which is highly similar to ADP·P$_i$. All spectra were recorded at 50°C.

4.3.3 Analysis of thiosubstituted nucleotide binding by solid-state NMR

Reaction control using ^{31}P 1Ds

NMR experiments involving nucleotide analogues plus the regeneration system require reaction control. The assignment of ^{31}P resonances was straightforward, following literature values [89] and previously established rules:

Sp-ATPαS

Any peak with a chemical shift around 30-40 ppm is downfield shifted due to the thiosubstitution. Thus, in the case of the measurement with *Sp*-ATPαS, the peak at 43.5 ppm is assigned to the αSP phosphate, the one -21.8 ppm to the βP phosphate and the peak at -5.8 ppm, by process of elimination, to γP phosphate (Figure 41a). Overall, chemical shifts of neighbouring phosphates are weakly perturbed by the thiosubstitution as they fall very close to those of pure ATP [89], although the reason for a strong broadening of the αSP peak is unclear. After all PEP is depleted, *Sp*-ADPαS starts to emerge with a βP resonance at -6.59 ppm. The αSP peak of *Sp*-ADPαS seems to overlap with the one of *Sp*-ATPαS, as only one apparent peak is visible around 43.5 ppm. The loss of *Sp*-ATPαS can be instead monitored by the loss of signal at -21.8 ppm. Based on the fact that the *Sp*-ADPαS signal only starts to rapidly build up after the depletion of PEP, it seems logical that the regeneration system is, to some degree, compatible with the αS-analogues. In this combination, however, the setup is still partially deficient since *Sp*-ADPαS is clearly present in minor quantities already at the beginning of the reaction. Considering that the *Sp*-ATPαS hydrolysis to *Sp*-ADPαS is happening at 7°C, a thermal degradation can be excluded, suggesting a direct hydrolysis by p97.

ATPβS

In contrast, no significant PEP consumption is detected over the course of the ATPβS measurement, and the spectra remain identical over time (Figure 41b). The αP and γP peaks are weakly (~1 ppm) shifted upfield from the ATP reference, while a range of peaks between 30 and 39 ppm is assigned to the βSP nucleus. Based on the nucleotide ratios determined in ^{1}H 1Ds, it is known that the solution contains a mixture of *Sp*-ATPβS, *Rp*-ATPβS and ADPβS, though this multitude of species is not obviously visible in the ^{31}P 1D. Based on literature values [89], chemical shifts of the two ATPβS diastereomers differ by 0.2-0.4 ppm – a difference too small to be resolved with the resolution of the solid-state NMR setup.

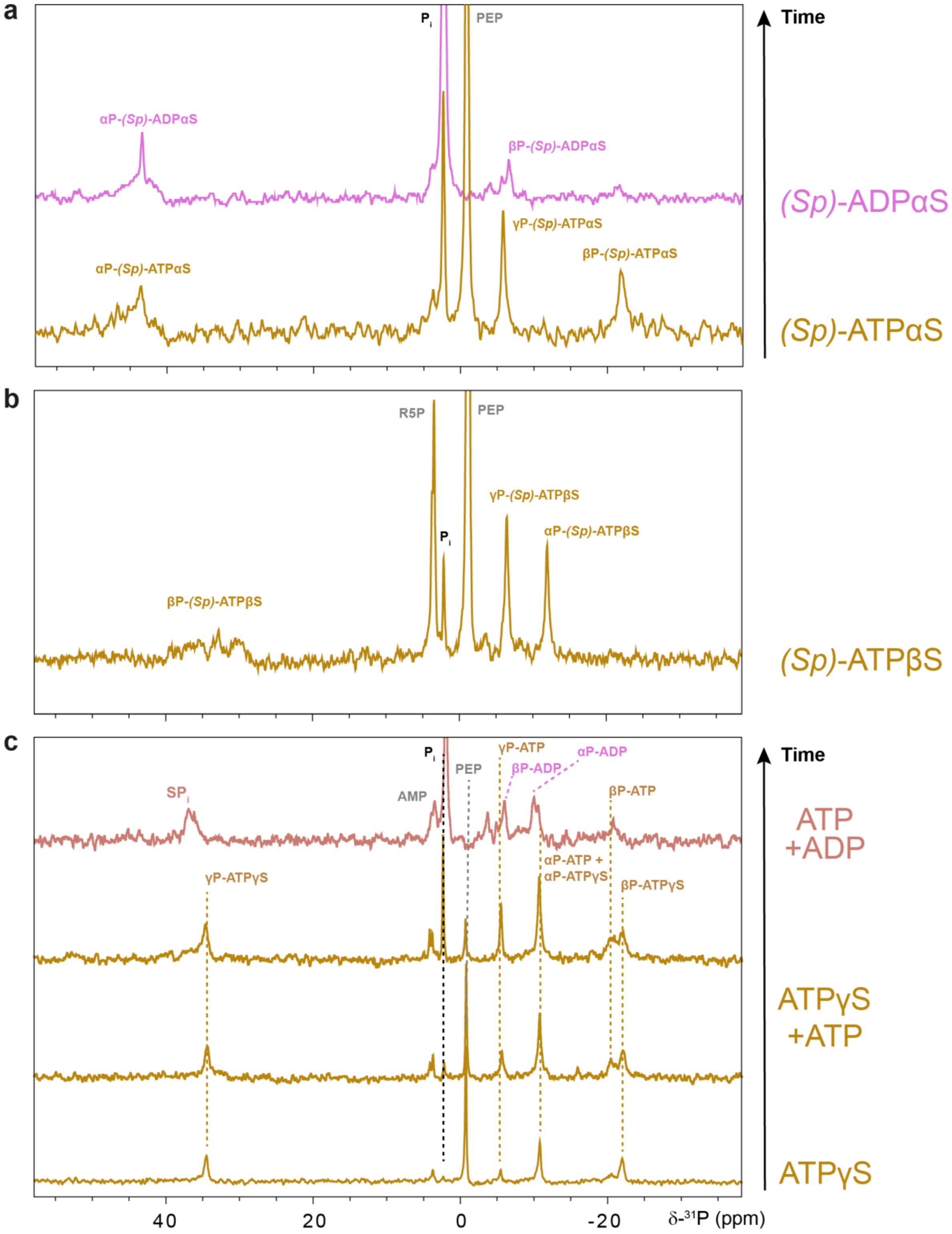


Figure 41. ^{31}P NMR measurements of thiosubstituted ATP analogues in solution. Directly pulsed ^{31}P 1D NMR measurements were used for reaction control during ATPase measurements of ND1L-E305Q with thiosubstituted nucleotide analogues at 7°C. Thiosubstituted phosphate resonances are shifted downfield by ~40 ppm. **a**) In the beginning of the measurement, only *Sp*-ATPαS is present in solution. After 72 h of measurement time, the sample was heated to

25°C for 4 h to speed up the reaction and deplete PEP. After the depletion of PEP, *Sp*-ADPαS starts to emerge, while the βP peak of *Sp*-ATPαS disappears. **b**) No significant PEP consumption was observed during the measurement with ATPβS; spectra remained identical over time. **c**) Measurements of ATPγS with the regeneration system lead to mixed nucleotide species in solution due to regeneration of ATP from the hydrolysis product of ATPγS, ADP. In the beginning, almost exclusively ATPγS is found in solution. Over time, as long as PEP is present, ATP starts to build up (time difference between the second and third spectrum ~70 h). Upon depletion of PEP, both ATP and ATPγS are slowly converted to ADP. After the third spectrum (from the bottom) the sample was heated to 30°C to speed up the hydrolysis reaction. The first spectrum is shown as a reference of a different, identical measurement. All measurements were recorded at 45 kHz MAS. Peaks were assigned following literature values and pre-established rules (see text).

Additionally, chemical shifts of ADPβS differ from the respective shifts of its ATP analogues by maximum 1 ppm. Thus, the vast peak distribution of the βSP resonances must rather be attributed to the presence of several different Mg^{2+} complexes of the aforementioned nucleotides, possibly in an intermediate exchange. In conclusion, the type of nucleotides in solution (aside from *Sp*-ATPβS) cannot be identified with any certainty. Considering that the pyruvate kinase is easily inhibited by ATPβS, it cannot be determined whether the regeneration system functions inside the protein sediment as successfully as in a dilute solution.

ATPγS

When it comes to ATPγS, reaction control becomes more complicated. In the beginning, only ATPγS and PEP are present in solution, but the minor impurity of commercially available ATPγS – ADP – is immediately regenerated to ATP at the start of the reaction (Figure 41c). As with the other nucleotide analogues, peaks are readily assigned: the most downfield peak at 33.8 ppm belongs to the thiosubstituted γSP resonance, the most upfield one to the βP phosphate. By process of elimination, the final peak is assigned to the αP-phosphate with a chemical shift that is very similar to αP-ATP. The remaining minor peaks can be assigned to regenerated ATP. Over time, while ATPγS is slowly hydrolysed, more ATP is built up and the nucleotide ratio in solution changes. PEP is depleted at an ATP:ATPγS ratio of ~50%, after which both molecules are slowly hydrolysed to ADP. To some extent, this can be followed by a build-up of the cleaved inorganic SP_i peak at 36.9 ppm.

Analysis of bound thiosubstituted nucleotide analogues

Since the methyl spectra in solution could not unambiguously establish the state of the protein (pre-hydrolysis or post-hydrolysis/ADP·P_i-like) in the presence of thio-substituted ATP analogues, we analysed whether the nucleotide state could be pinned down from ^{31}P chemical shifts under MAS.

Sp-ATPαS

^{1}H-^{31}P-CP spectra of bound *Sp*-ATPαS in the presence of the regeneration system seem to produce an ATP-like state (Figure 42). Three major groups of peaks are visible in the spectra recorded in the presence of PEP, which can be unambiguously assigned to the respective *Sp*-ATPαS phosphates. However, the centre group around -2 ppm consists of three peaks, with the strongest at -1.46 ppm, the second at -2.43 ppm and the weakest at -3.96 ppm. Over the course of the measurement, a progression can be seen. A mixture of *Sp*-ATPαS and *Sp*-ADPαS is found in the middle CP spectrum, recorded after depletion of PEP and in the presence of mixed nucleotides (*c.f.* Figure 41a). In the last spectrum, only two peaks remain, which can be assigned to *Sp*-ADPαS. Using these chemical shifts as a reference, one can identify a minor contribution of *Sp*-ADPαS in the first spectrum as well, confirming that both ATP and ADP analogues are bound from the start. Considering that both nucleotides were present from the beginning of the measurement – and *Sp*-ADPαS in a much smaller quantity – it is not surprising to detect both an ATP/pre-hydrolysis-like state and a minor ADP-like state. Unfortunately, in this spectrum, no indication of the ADP·P_i-like state can be found, implying that the rate-determining step has changed with this ATP analogue, from phosphate release to ATP hydrolysis.

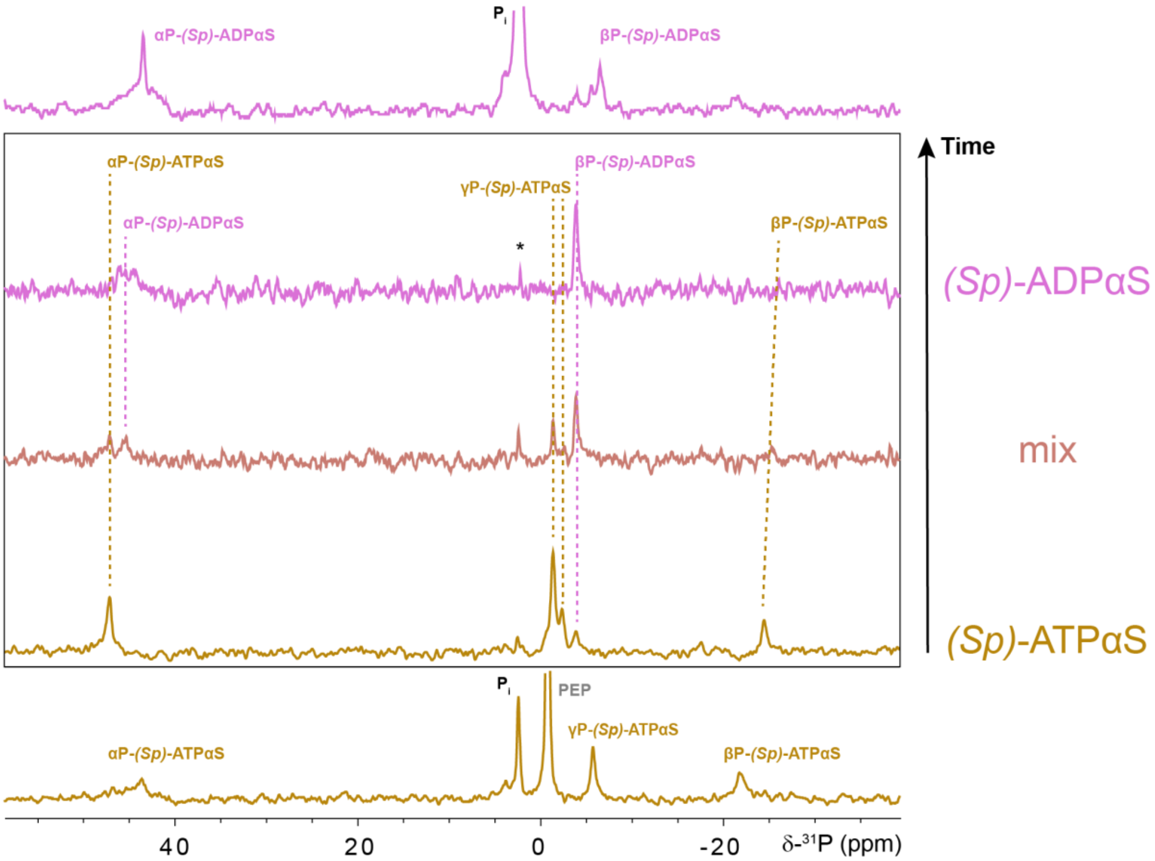


Figure 42. ^{1}H-^{31}P-CP spectra of bound *Sp*-ATPαS. While PEP is present in solution, the main bound species is determined to be *Sp*-ATPαS with a minor presence of *Sp*-ADPαS. After depletion of the regeneration system, ratios of the bound species change towards *Sp*-ADPαS, until only the ADP analogue remains. No indication of an ADP·P_i-like state is visible. The origin of the two peaks between -1-2.5 ppm is unclear, as both resonances could be possibly assigned to the γP phosphate. Reference spectra of analogue nucleotides in solution are shown on top (*Sp*-ADPαS) and bottom (*Sp*-ATPαS). All measurements were recorded at 45 kHz MAS. The spectrum showing pure bound *Sp*-ADPαS is taken from a different, identical measurement. * denotes the unbound P_i.

Curiously, the ^{1}H-^{31}P CP-spectrum displays a greater peak dispersion for the respective *Sp*-ATPαS phosphates than the spectrum of the free nucleotide in solution (Figure 42, bottom), a feature not seen for ATP. Here, the thiosubstituted αSP peak shifts further downfield with respect to the reference, while the βP peak shifts upfield from the free nucleotide reference. Additionally, two peaks could potentially be assigned to the γP resonance, both of which are strongly downfield shifted to -1.46 ppm and -2.43 ppm, respectively, in comparison to the reference. The intensity of the former is also surprising, as it is much more intense than any of the others. A splitting of the γP phosphate has been reported for GMP-PNP bound to the small GTPase Ras [104]. The observation was explained by assuming two separate binding states involving different intermolecular contacts, specifically hydrogen bonds. This argument could also work for p97, especially for a nucleotide analogue displaying non-native binding. For instance, Mg^{2+}-bound and Mg^{2+}-free forms in C2/C3 conformations could produce such an effect. A possible alternative explanation would be to re-interpret the ^{1}H-^{31}P CP spectrum as a mixture of two different nucleotide states: ATP-like and ADP·P_i-like (instead of ADP-like). In this case, the proven presence of *Sp*-ADPαS would be interpreted as a post-hydrolysis state, with the cleaved P_i showing only a single resonance at -1.46 ppm or -2.43 ppm. For such a spectral interpretation to make sense, it would need to be assumed that the ADP·P_i-like state of *Sp*-ADPαS·P_i shares strong similarities with the ADP-like state and that the inorganic phosphate is rigidly bound in a homogeneous environment. Either way, this interpretation would highlight another major difference to the native ADP·P_i state and, thus, the limited usefulness of this ATP analogue. Altogether, this nucleotide analogue appears to be bound by the protein in a non-native way, and the rate-determining step of the ATPase cycle appears to have changed to hydrolysis. The comprehensive interpretation of protein- and nucleotide-detected results is found in the discussion below.

ATPβS

^{1}H-^{31}P-CP spectra of ATPβS in the presence of the regeneration system similarly produce an ATP-like state (Figure 43). Three major peaks are detected, with the one at 33.4 ppm being assigned to the βSP phosphate group. Two other signals correspond to the other two phosphates, but their assignment is ambiguous. Presumably, binding to p97 causes a 3.5-4 ppm downfield shift (see next section), which would imply that the right peak belongs to the αP phosphate and the central peak to the γP phosphate. It is unclear, however, why the terminal phosphate would produce such an intense signal in comparison to the other resonances. A different interpretation of the spectrum could be made by assuming an ADP-like state with a strong additional peak, possibly the inorganic phosphate. Based on the presented data, it is impossible to definitively differentiate between bound ATPβS (with three peaks) and ADPβS with an additional signal. Finally, the binding modes of present *Rp*-ATPβS are completely unknown and could produce different spectra altogether – if the nucleotide is bound at all. Overall, the measured spectrum does not show key attributes of the native ADP·P_i state, suggesting that ATPβS is a poor ATP analogue for p97.

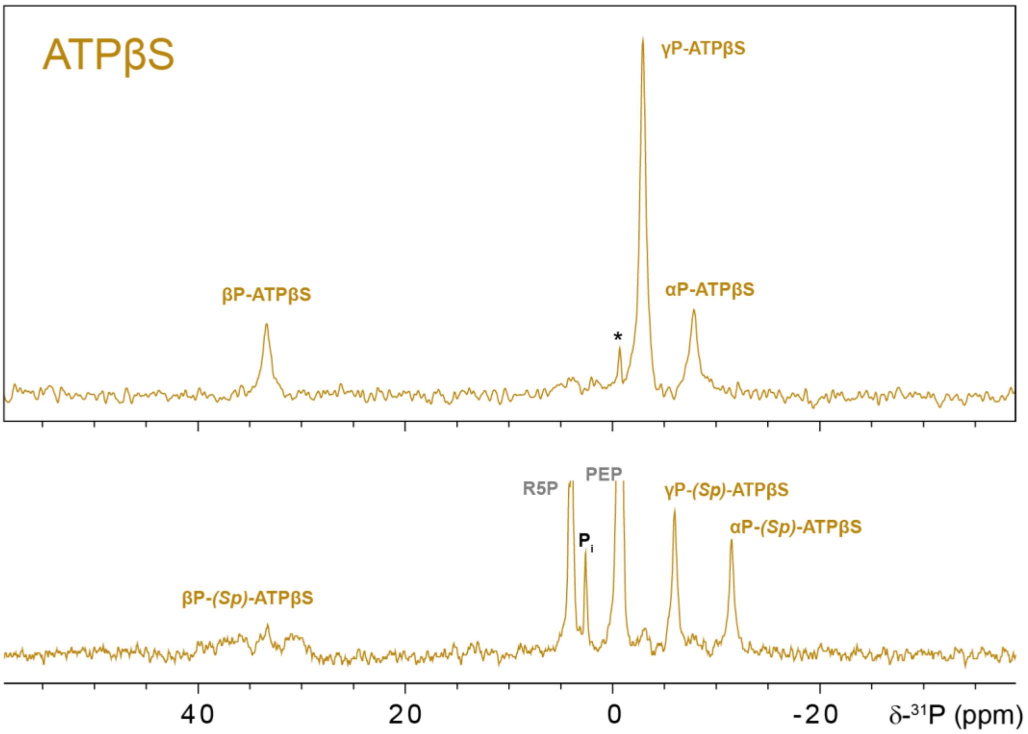


Figure 43. ^{1}H-^{31}P-CP spectra of bound ATPβS. ND1L-E305Q bound ATPβS is measured in the presence of the regeneration system. The main bound nucleotide species is assigned to be ATPβS. While the assignment of the βSP phosphate to the peak at 33.4 ppm is obvious, assignment of the other two peaks to the αP and γP phosphate groups is ambiguous. Further, the reason for the high intensity of central peak is unclear. Reference spectra of *Sp*-ATPβS in solution is shown at the bottom. The measurement was recorded at 45 kHz MAS. * denotes the unbound PEP.

ATPγS

Spectra measured in the presence of ATPγS and a regeneration system indicate mixed nucleotide states, which can be unambiguously assigned (Figure 44). Initially, two bound states are detected: ATPγS bound and ADP·SP$_i$. The ATPγS bound state is assigned following the free nucleotide spectrum reference. The two other major peaks at -4.5 ppm and -16 ppm are very similar to the peaks of the ADP·P$_i$ state (see below) and are thus assigned to an ADP·SP$_i$ state. The heterogeneity of the cleaved P$_i$ observed in the ATP measurement between -2 and -12 ppm is absent, and instead, multiple peaks between 30 and 37 ppm are seen. In reality, a more complicated mixture of states is expected since aside from ATPγS, ATP is also present in the solution (see Figure 41). However, a spectrum with similar features is observed with pure ATPγS, i.e. without a regeneration system (Figure 44, bottom). While the peak ratio of

ATPγS-bound to ADP·SPi-bound protein is shifted towards the pre-hydrolysis/ATPγS state, both states appear without ATP being in solution. The sheer presence of these ADP·P$_i$-like peaks confirms the existence of a third nucleotide state, aside from the pre-hydrolysis ATP(γS) state and the post-hydrolysis ADP state. Moreover, a weak, broad signal is observed between 30 and 37 ppm, clearly downshifted by ~5 ppm from the γSP resonance of the terminal phosphate group, which gives a narrow, obvious signal. Thus, this signal must correspond to the cleaved inorganic thiophosphate.

After the regeneration system runs out, ATP and ATPγS are slowly hydrolysed to ADP, which is eventually the only species bound by p97 (Figure 44, top).

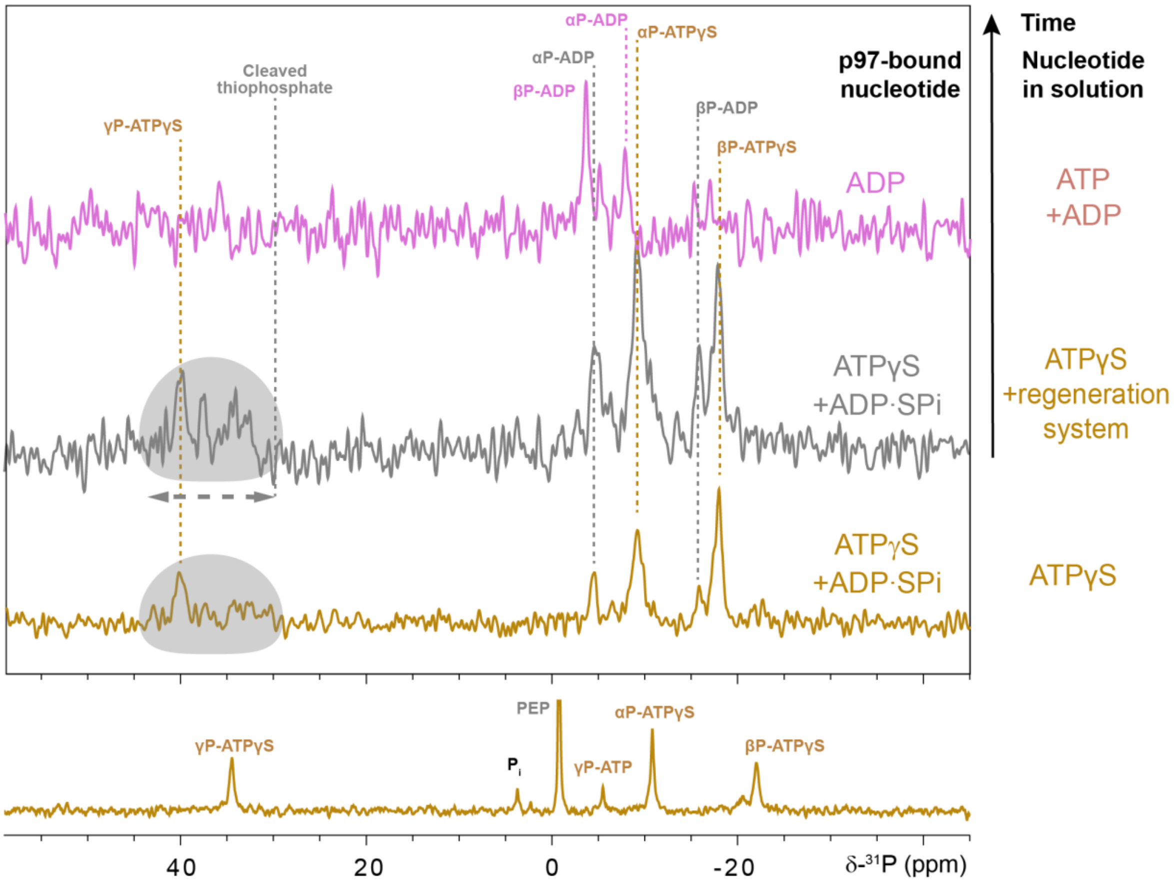


Figure 44. ^{1}H-^{31}P-CP spectra of bound ATPγS. ND1L-E305Q bound ATPγS is measured in the presence of the regeneration system. A mixture of two states is detected: ATPγS and ADP·SP$_i$. Noticeably, the broad peak dispersion of the cleaved P$_i$ observed in Figure 28 is shifted to 30-37 ppm due to the thiosubstitution to SP$_i$. At least three distinct peaks are visible. This behaviour is not caused by the presence of ATP, since the same general pattern emerges in a control measurement without the regeneration system (bottom spectrum). After the depletion of the regeneration system, only bound ADP is detected. A reference spectrum of ATPγS in solution is shown on top. The measurements were recorded at 45 kHz MAS.

4.3.4 Comparison of all ^{1}H-^{31}P-CP spectra

So far, most interpretations of the protein-bound nucleotide states have come from singular measurements, internally comparing the changing spectra over time. In order to prove peak assignment, it is worth comparing different measurements.

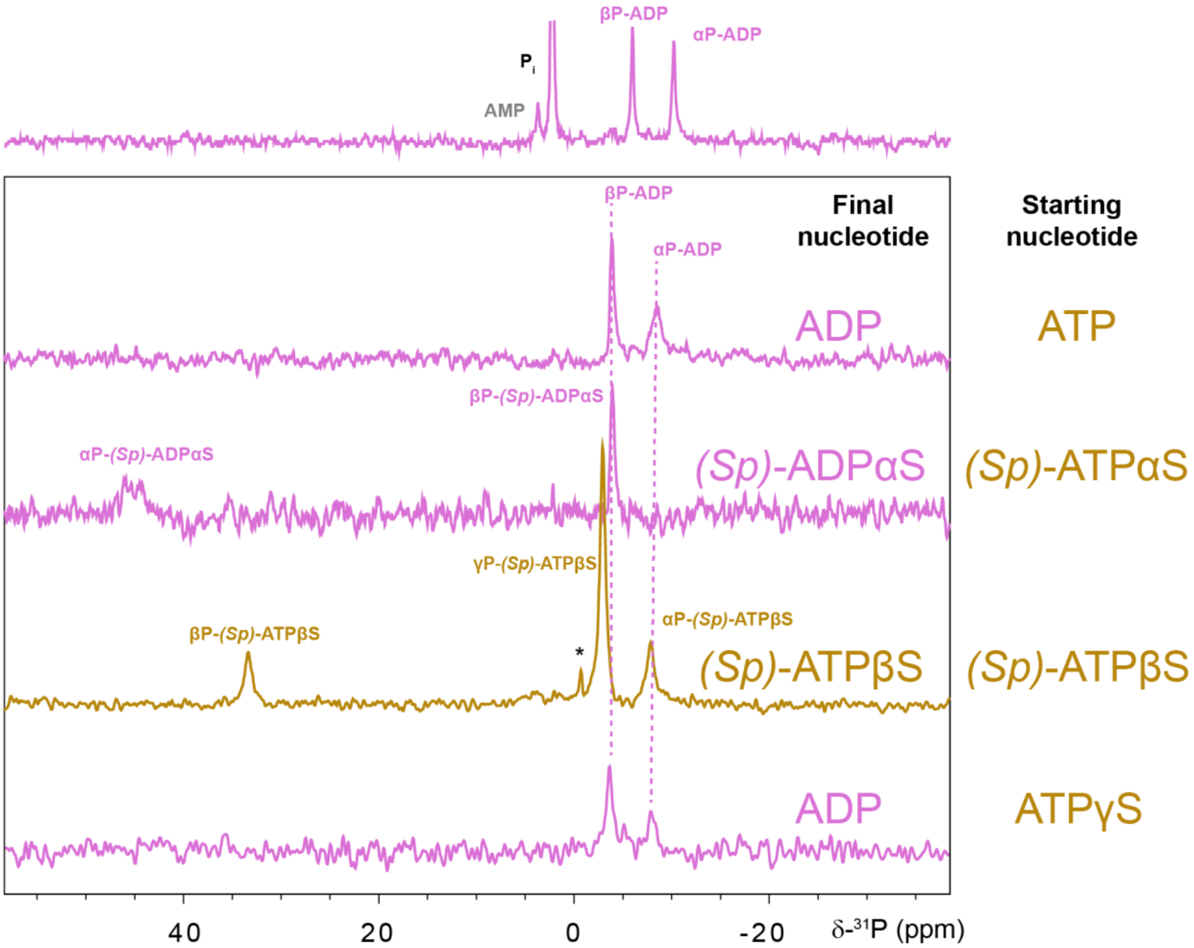


Figure 45. ^{1}H-^{31}P-CP spectra of bound ADP-like states. Comparison of spectra showing ADP-like features. Measurements were conducted starting from ATP, *Sp*-ATPαS, ATPβS or ATPγS. All spectra (aside from ATPβS) show very similar features, mainly of two prominent peaks. Expectedly, ATP and ATPγS spectra are highly similar, as the final nucleotide composition of both measurements is almost identical. In contrast, the αP phosphate of the *Sp*-ATPαS measurement is strongly downfield shifted, as it would be expected for a thiosubstituted phosphate. The same downfield shift is seen for the ATPβS measurement, although the exact type of detected nucleotide is unclear. A reference spectrum of ADP in solution is shown on top. The measurements were recorded at 45 kHz MAS. * denotes the unbound PEP.

First, only ADP-like states are compared in Figure 45: all spectra (aside from the measurement with ATPβS) show two main peaks corresponding to the αP and βP resonances. Spectra recorded starting with ATP and ATPγS show great similarities, both in terms of peak pattern and chemical shifts. This is not surprising, as by the end of both experiments, ADP is the main nucleotide found in the solution, emphasising that these results are generally reproducible. Also, in both cases, the left peak is narrower and more intense than the right one.

The latter was assigned to the αP resonance following the reference shifts of the free nucleotide. This assignment is confirmed by the thiosubstitution-induced downfield shift of the right peak in the spectrum of bound *Sp*-ADPαS. Moreover, the chemical shift of the βP resonance is unaffected, implying a native-like binding of the ADP analogue.

ADPβS can neither be confirmed nor completely ruled out, since the chemical shifts of both the αP and βSP peaks could support the assignment of an ADP analogue. As mentioned previously, the assignment of the nucleotide state of the ATPβS measurement remains ambiguous.

Overall, the similarity in chemical shifts between different ADP-like nucleotides is remarkable but not surprising. Oxygen thiosubstitution can affect modes of metal binding due to the softness of the sulphur in contrast to the hard oxygen [79,114]. However, no metal ions are required for ADP binding [32].

ATPγS

Comparing the spectra showing pre-hydrolysis states, bound ATPγS must show three distinct peaks, one of which (γSP) would be strongly downfield shifted. The assignment of two other peaks is slightly more ambiguous, but a chemical shift of -18 ppm would better fit for the βP resonance in comparison to the free nucleotide reference (Figure 46). Additionally, the stronger intensity and decreased broadness of this peak in comparison to the αP peak is reminiscent of the established line shapes of the αP and βP peaks of bound ADP. Thus, the assignment of the ATPγS spectra is found to be reliable. Similar features (broad left peak; narrow and more intense right peak) are found in the ATP and ADP·P_i bound spectra as well (Figure 46). This finding supports their respective assignments of the αP and βP peaks. The missing γP-phosphate of the ATP-bound spectra thus must be assigned to the weaker peak at -8.7 ppm, while an identical argument can be made for the ADP·P_i state where the cleaved phosphate is detected as multiple, broadly dispersed peaks. These peaks shift to 30-37 ppm in the ATPγS measurement, confirming the identity of the cleaved SP_i – and in retrospect the identity of the cleaved P_i in the ADP·P_i measurement.

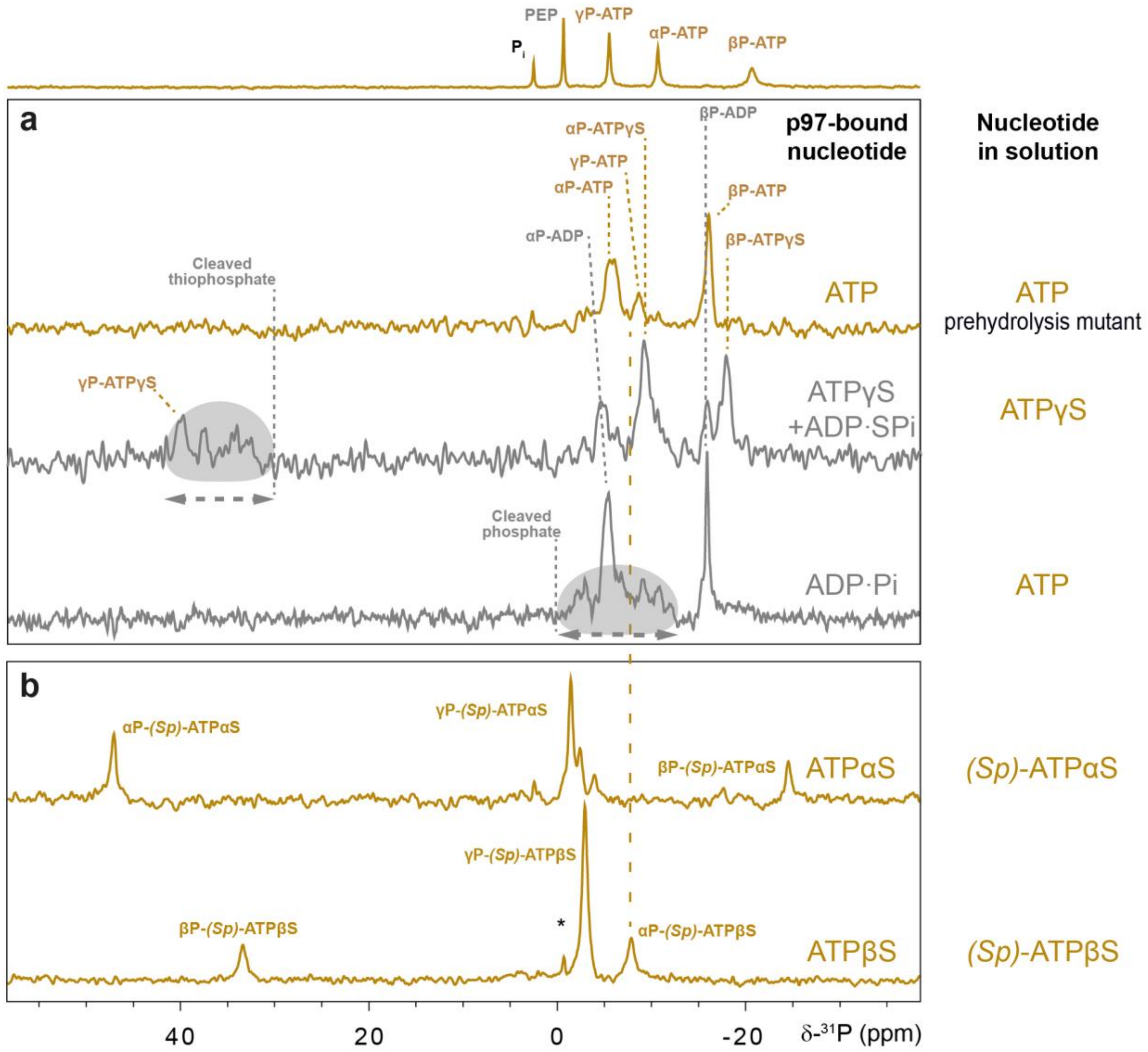


Figure 46. ^{1}H-^{31}P-CP spectra of bound nucleotides in the presence of PEP. a) Comparison of spectra showing pre-hydrolysis and post-hydrolysis features. Measurements were conducted starting from ATP, ATP with a disease mutant or ATPγS. All spectra show very similar features, mainly of two prominent peaks which are assigned to αP and βP resonances. Minor peaks observed in the ATPγS measurement overlap very well with the ADP·P$_i$ reference, proving the observation of an ADP·SP$_i$ state. A reference spectrum of ATP in solution is shown on top. **b**) Comparison of *Sp*-ATPαS and ATPβS spectra. In the presence of both nucleotide analogues the pre-hydrolysis state is observed. The measurements were recorded at 45 kHz MAS. * denotes the unbound PEP.

Sp-ATPαS

The nucleotide state in the presence of *Sp*-ATPαS is most likely the pre-hydrolysis state, as clearly (at least) three distinct peaks are visible. While an *Sp*-ADPαS state can be ruled out due to its unambiguous observation in the same measurement, the *Sp*-ADPαS·P$_i$ state with a strongly shifted, cleaved P$_i$ in a single homogeneous environment is, nonetheless, possible. Due to the dissimilarity of the spectrum in comparison to the ATP, ATPγS and ADP·P$_i$ bound spectra, it cannot be definitively stated what type of nucleotide state is observed. Still, the strong upfield shift of the βP resonance is usually encountered for ATP (and its analogues).

ATPβS

In the ATPβS spectrum, only the assignment of a single peak (βSP) is unambiguous, while the chemical shift of the peak at -7.9 ppm falls close to the resonances of both αP-ADP·P$_i$ and αP-ATPγS, but also γP-ATP. Moreover, as stated above, the chemical shift overlaps with the αP-ADP shift, making a reliable assignment impossible. The intense central peak actually shares some similarities with the γP phosphate of the *Sp*-ATPαS spectrum, thus partially supporting the assignment of the observed state towards the pre-hydrolysis state. However, similar to the *Sp*-ATPαS case, in principle, an ADP-like ADPβS·P$_i$-like state with a single, intense cleaved P$_i$ peak cannot be completely ruled out. Another possible interpretation would be that ATPβS indeed is bound by p97, but Mg^{2+} binding is inhibited, creating an ADP-like environment, which would fit the observed the chemical shift of the αP-resonance.

Table 4. ^{31}P chemical shifts of free and bound nucleotide species. Chemical shift ranges indicate variability between different spectra, temperatures and experiments.

Nucleotide species		**^{31}P chemical shift [ppm]**		
		α-phosphate	**β-phosphate**	**γ-phosphate**
In solution	P$_i$	2.1-2.6	n.a.	n.a.
	SP$_i$	36.5-37.0	n.a.	n.a.
	ADP	-10.0-10.2	-5.8-6.1	n.a.
	ATP	-10.6-10.7	-19.9-20.8[a]	-5.2-5.6
	Sp-ATPαS	43.3-43.8	-21.8	-5.7-5.8
	Sp-ADPαS	43.4	-6.6	n.a.
	ATPβS	-12	30-38[b]	-6.5
	ATPγS	-10.7-10.8	-21.0-22.0[a]	34.5-35.1
Protein-bound	ADP	-7.7-8.5	-3.5-3.8	n.a.
	ATP	-5.6-6.1	-16.1	-8.7
	ADP·P$_i$	-5.0-5.5	-15.9	-2.1 -3 -6.8 -9.1 -10.8 -12[b]
	Sp-ATPαS	47	-24.5	-1.46; -2.43[b]
	Sp-ADPαS	45.4	-4	n.a.
	ATPβS	-7.88	33.4	-2.95
	ATPγS	-9.1-9.2	-17.9-18.0	39.8-40.1
	ADP·SP$_i$	-4.4-4.6	-15.9-16.0	30-38[b]

[a]: broad peaks observed. [b]: multiple peaks observed. n.a.: not applicable

Comparison of all nucleotide analogues

Table 4 summarises chemical shifts of all free and bound nucleotide species as they have been assigned in the previous section. This assignment can be supported by comparing chemical shifts of bound species to the reference ones in solution (Table 5). Since the assignment of the ADP peaks has been proven unambiguously through thiosubstitution, the binding-induced shift of the ADP peaks can be taken as an overall reference. Both phosphate peaks are shifted by 2-2.3 ppm downfield due to protein binding. Interestingly, very similar values (2.0-2.6 ppm) are calculated for *Sp*-ADPαS, suggesting that this trend is generally true for this system. The picture becomes less clear for other nucleotides, though: αP- and βP-ATP resonances shift by 4-5 ppm, and these values are generally encountered for other nucleotide states as well. For instance, ATPγS peaks shift by 1.6-5.2 ppm upon binding, although it is not exactly clear why the terminal phosphate of ATPγS experiences a strong downfield shift while the same group of ATP is shifted upfield. Indeed, some values fall out of place, e.g. the minor upfield shifts of the βP resonances of bound *Sp*-ATPαS and ATPβS.

The advantage of this calculation also lies in the comparison of different states. The chemical shift perturbations observed for the ATP, ADP·P_i and ADP·SP_i nucleotides upon binding are very similar for all three cases, supporting the individual assignments. Additionally, comparisons to the ADP references reveal strong mismatches (4.9-5.6 ppm for αP, but -10.0 ppm for βP), implying that a potential misinterpretation of the ADP state for an ADP·(S)Pi state is highly unlikely. Even more, both absolute and relative chemical shifts (Tables 4 and 5) emphasise the similarities between the ATP, ADP·P_i and ADP·SP_i states.

Table 5. Nucleotide ^{31}P chemical shift perturbations induced by binding to p97. Chemical shifts of bound species are compared to those of the free nucleotides. The chemical shift difference $\Delta\delta$ was calculated following $\Delta\delta = \delta(bound) - \delta(free)$ for the respective phosphate resonances. For this calculation, mean values of chemical shift ranges were used. Positive values imply a binding-induced downfield shift. ADP·$\mathrm{P_i}$(-like) state nucleotide shifts are compared to both ATP and ADP references.

Bound species	**Free nucleotide reference**	**^{31}P chemical shift difference $\Delta\delta$ [ppm]**		
		α-phosphate	**β-phosphate**	**γ-phosphate**
ADP	ADP	2	2.3	n.a.
ATP	ATP	4.8	4.3	-3.3
ADP·$\mathrm{P_i}$	ADP	4.9	-10.0	n.a.
	ATP	5.4	4.5	-1.9[a]
Sp-ATPαS	*Sp*-ATPαS	3.5	-2.7	4.3; 3.3
Sp-ADPαS	*Sp*-ADPαS	2.0	2.6	n.a.
ATPβS	ATPβS	4.1 (9.1)[c]	-0.6	3.6 (-1.4)[c]
	ADPβS (lit.)[b]	4.1	0.4	n.a.
ATPγS	ATPγS	1.6	3.6	5.2
ADP·$\mathrm{SP_i}$	ADP	5.6	-10.0	n.a.
	ATPγS	6.3	5.6	-0.8

[a]: a single mean value of -7.3 ppm was calculated for the cleaved $\mathrm{P_i}$. [b]: literature value is taken from ref. [89] and converted to values from this work. [c]: Values in parenthesis represent identical calculations for the case if the αP and γP-assignments would have been swapped. n.a.: not applicable

Still, the interpretation of the ATPβS measurement remains difficult. Explaining all three major peaks as the actual phosphate groups of the pre-hydrolysis state, it would seem that only the shown αP/γP assignment of Table 4 is plausible: if these chemical shifts were to be swapped, chemical shift differences would be greater and less typical for this system (9.1 and -1.4 ppm instead of 4.1 and 3.6 ppm). However, if one were to ignore the major central peak (assigned to γP) and interpret it as an entirely separate entity, e.g. bound inorganic phosphate, one could only focus on the remaining phosphate groups. These peaks could be interpreted as bound ADPβS, as stated in the previous section, and the resulting chemical shift differences would be almost identical to those from the ATPβS-interpretation. Altogether, the assignment of this nucleotide state remains challenging.

4.4 Discussion

The major characteristic of the ADP·P_i state in the solid-state ^{31}P NMR spectra is the heterogeneous behaviour of the cleaved P_i. Still, an incorrect assignment of the two major peaks (αP and βP) would lead to a wrong interpretation regarding the nature of the intermediate state, especially whether it is more ATP-like or more ADP-like.

The usage of ATPγS as a thiosubstituted ATP analogue has helped to unambiguously prove the peak identity of the cleaved phosphate through the downfield shifting of the corresponding peaks. Moreover, spectra of bound ATPγS provide a solid reference for other nucleotides, especially ATP. The intuitive assignment of the ATPγS-bound peaks helped to support the peak assignment of the ATP-bound state. Due to very similar appearances of the ATP-bound and ADP·P_i states, a straightforward peak assignment transfer is possible. However, as discussed in Chapter 3, spectra of bound nucleotides can unpredictably vary depending on nucleotide type and protein system. In that sense, an ADP-like ADP·P_i state would be entirely possible, and the peak order would be reversed. Thus, peak assignment transfer from ATP to ADP·P_i states remains unproven.

For this reason, two different thiosubstituted nucleotides were chosen to cross-check the assignment. The thiosubstitution-induced downfield shift of either αP (*Sp*-ATPαS) or βP (ATPβS) phosphates could have unambiguously identified the respective peaks. Unfortunately, these nucleotide analogues turned out to be poor ATP mimics, as they have already shown in preparatory experiments prior to the solid-state NMR measurement: both the failing ATPase assay measurements and the poor chemical stability prevented rigorous testing of the nucleotides. The lack of commercially available ATPβS was solved by an enzymatic regeneration of ADPβS with pyruvate kinase. The key issue of these nucleotides is the introduction of additional *P*-chiral centres, which can affect both the ATPase (p97) as well as the kinase (pyruvate kinase). This complicated the setup of the ATP regeneration system and even introduced product mixtures: ADPβS was partially regenerated to *Rp*-ATPβS alongside the main diastereomer (*Sp*-ATPβS).

Most importantly, thiosubstitution affects metal ion binding. The hard cation Mg^{2+} preferentially forms chelating complexes with the hard oxygen but not with the soft sulphur. Inside the binding pocket, however, stereochemistry is mainly dictated by the positioning of the amino acid side chains. A repositioning of the metal ion to favour a specific chalcogen binding may not be possible. For example, it has been shown that the maximum activity of the yeast hexokinase drops by a factor of 12 when *Sp*-ATPαS is exchanged for *Rp*-ATPαS, but a decrease by a factor of 590 is observed upon *Rp*-ATPβS exchange for *Sp*-ATPβS [119]. A similar stereospecificity has been reported for the acetate kinase [121]. It is thus not surprising that p97 would be affected by the newly introduced chirality and preferentially process one diastereomer over the other. Even more, the introduced thiosubstitution could affect the mode of Mg^{2+} binding and change the location of the nucleotide or the metal ion. The contradicting interpretations of the nature of the bound nucleotides seen in solution-state NMR spectra (e.g. I206 more similar to an ATP-analogue, while the L268 peak detecting a nucleotide akin to ADP) are the consequence of the non-native nucleotide and/or metal ion positioning. In contrast, the methyl TROSY spectrum in the presence of *Sp*-ADPαS is highly similar to the one with natural ADP (with minor CSPs detected). A supporting observation of identical ^{31}P chemical shifts for βP-ADP and βP-*Sp*-ADPαS is made in solid-state NMR. This affirms the previous hypothesis: since Mg^{2+} is not required for ADP binding, *Sp*-ADPαS can adopt the native-like conformation inside the binding pocket more easily.

Sp-ATPαS

The non-native binding of *Sp*-ATPαS during ATP turnover is also reflected in the ^{31}P NMR spectra. Based on the chemical shifts, the bound nucleotide has been identified as the pre-hydrolysis *Sp*-ATPαS, although the peak splitting of the presumed γ-phosphate hints at the possible presence of a secondary binding mode. Although the αP-phosphate is not directly involved in the hydrolysis step, its chemical modification can still affect the nucleotide-protein interaction inside the binding pocket.

Both ^{1}H NMR and ^{31}P NMR experiments confirm that *Sp*-ADPαS can be successfully converted to *Sp*-ATPαS by the regeneration system (Figures 37a and 41a). However, in ^{31}P NMR experiments, a minor population of *Sp*-ADPαS was still detected in solution and bound by the protein, even though PEP was still available, suggesting a slow turnover by the pyruvate kinase and limited compatibility of the analogue with the regeneration system. Possibly, the ATPase assay failed due to a much faster *Sp*-ATPαS-hydrolysis rate of wild-type p97-ND1L in comparison to the regeneration rate by the pyruvate kinase. This problem was circumvented in the other experiments due to the usage of the slower ND1L-E305Q mutant.

Based on the fact that nucleotide turnover is confirmed via PEP consumption (Figure 41a), one must assume that *Sp*-ATPαS can be processed by p97. However, ^{31}P NMR spectra of p97-bound nucleotides imply the binding of the pre-hydrolysis *Sp*-ATPαS instead of a post-hydrolysis $ADP\cdot P_i$-like nucleotide (Figure 42), suggesting that the rate-determining step of the ATPase cycle changed from P_i release to *Sp*-ATPαS hydrolysis. While this is not obviously reflected in the solution-state NMR spectra, the many differences in regards to the native ADP·P_i state, incl. the mixed NTD positioning, suggest that the nucleotide analogue is improperly identified and processed by p97, leading to an unclear protein state. As is discussed in Chapter 5, the state of the nucleotide (i.e. the nucleotide identity) is communicated from the binding pocket towards the outer regions of the protein, including the N-terminal domain. It is, thus, conceivable that the pre-hydrolysis state of *Sp*-ATPαS could emulate an ADP·P_i-like state, similar to ND1L-wt without Mg^{2+} (see Section 5.2.4).

ATPβS

A similar assessment can be made regarding the protein state in the presence of ATPβS and the regeneration system. ^1H NMR and ^{31}P NMR experiments confirm the conversion of ADPβS to ATPβS by the regeneration system (Figures 37b and 41b), even though some compatibility issues were still encountered (see pyruvate kinase inhibition by ATPβS). Still, the lack of PEP consumption in these experiments suggests that the nucleotide mimic cannot be processed by p97.

Under solution-state NMR conditions, a mixture of apo- and ADP·P_i-like spectra was observed. The mixed NTD-positioning could stem from an incorrect nucleotide type assessment by p97. In ^{31}P NMR measurements, no unambiguous assignment of the nucleotide type is possible, but a bound ATPβS seems to be more likely. One of the arguments for this assessment is the lack of PEP consumption during the measurement (Figure 41b). Additionally, the chemical shift of the presumed αP-resonance of bound ATPβS is highly similar to that of bound ADP, suggesting that ATPβS is bound identically to ADP – i.e. without Mg^{2+}. This interpretation could also explain the absence of ATPase activity, since the metal ion is required for hydrolysis.

Following this argument, methyl TROSY spectra of p97 in the presence of ATPβS and the regeneration system can be interpreted in a similar way as the ones in the presence of *Sp*-ATPαS: the chemical modification leads to a partial misinterpretation of the nucleotide type by the protein. For that reason, it remains unknown whether the rate-determining step has changed to hydrolysis, as seen for *Sp*-ATPαS.

ATPγS

Contrasting, another chemically modified ATP analogue, the non-*P*-chiral ATPγS, is correctly interpreted by p97. The reference spectra of both solution and solid-state NMR depict a pre-hydrolysis state (Figures 17 and 44), while the ones recorded in the presence of the regeneration system (at least partially) find the post-hydrolysis ADP·SP$_i$ state.

In conclusion, ATPγS is a functional ATP analogue which is correctly identified both in the absence and presence of the regeneration system. In contrast, both *Sp*-ATPαS and ATPβS are poor ATP analogues: *Sp*-ATPαS changes the rate-determining step of the hydrolysis reaction and is non-natively bound inside the binding pocket. ATPβS is poorly processed by p97, being a potential inhibitor of the ATPase. However, both nucleotide analogues are (slow) substrates of the pyruvate kinase.

Chapter 5: Defining features of the ADP·P_i state

5.1 Introduction

The observation of a native ADP·P_i state leads to a new fundamental question: what are the defining features of this state? How does it differ from the other protein states in the hydrolysis cycle? Which interactions inside the binding pocket support this conformational state, and which ones destabilise it? Why does the signal of the cleaved inorganic phosphate detected by ^{31}P NMR appear in the form of multiple weak peaks? Is the protein state dynamic or is it static? These questions lead to a collaboration of multiple research groups and methods (NMR, cryo-EM, MD) with the goal of studying the intricacies of the ADP·P_i state [52]. In the following chapter, results obtained from solution-state NMR and biophysical tests will be discussed. Those results will be related to observations made by the other two methods of structural biology, and conclusions will be drawn.

As discussed in Chapter 2, methyl labels spread throughout the protein provide effective sensors for probing conformational changes induced by nucleotide binding and processing. To deepen the functional understanding, it is necessary to study the contributions and interactions of individual amino acids known for their key roles in structural motifs, such as the Walker A, Walker B, or arginine finger (see Chapter 2). Unfortunately, many of the key residues making up these motifs are charged or polar and thus do not appear in methyl TROSY spectra. As a consequence, a direct study of these residues using the ILVM labelling technique is impossible, and in the best-case scenario, the effects of specific amino acids can still be analysed by observing the peaks of their neighbours.

A more direct approach involves mutational studies, in which the amino acid of interest is point mutated (often to an apolar amino acid, e.g. alanine [122]), and the effect of the mutation is analysed. In this work, we performed mutational studies on key residues and analysed their effects by means of solution-state NMR and biophysical characterisation, especially ATPase assay. Identical studies performed on the wild-type protein were used as a reference.

5.2 Roles of amino acids in the ADP·P_i state as studied by site-directed mutagenesis

5.2.1 Motivation

This thesis focuses on deepening the understanding of the ADP·P_i state observed for p97 [55]. In simpler terms, the questions become: what are the key characteristics of the state, and what contributes to its observation and stability? One approach to these questions is to focus on specific residues located in the binding pocket which could interact with the nucleotide. Crucially, many amino acids are described as being part of protein sequence motifs such as Walker A, Walker B, arginine finger or sensor. While the general function these motifs play is widely known, in this work, the target is to investigate the role these motifs and residues play in the longevity of the ADP·P_i state.

These protein sequence motifs are highly conserved (Figure 47), so the investigation of all residues would have been of great interest but also would go beyond the scope of this work. Instead, we decided to focus on key residues from each motif, many of which have been studied previously, with few exceptions (see below). Our approach involves site-directed mutagenesis in order to create single (or double/triple) mutants primarily of the ND1L construct in order to study the effects of these amino acid point-mutants. The mutations have been chosen for their prevalence in literature or to maximise the stability of the newly generated mutant.

The effects of these mutations on the ND1L construct have been evaluated using several criteria: oligomerisation, nucleotide binding, ATPase activity, and conformational behaviour, as observed by solution-state NMR. Here, four states have been measured: apo state (nucleotide-free), pre-hydrolysis state (with ATPγS and in some cases with AMP-PNP) and the post-hydrolysis ADP·P_i (under ATP turnover) and ADP states. For the NMR measurement, three general questions have been asked: do any of these measured mutant spectra differ (substantially) from the reference wild-type spectra? Is ATP turnover observed? Is the ADP·P_i state formed?

In order to determine the functional defect of any given mutant, experimental results were passed through a pipeline. In case the mutant was found unable to fulfil the current criteria, the mutant was dropped from the pipeline as subsequent analysis was deemed irrelevant. Potential deficiencies were assessed in the following order: 1) oligomerisation state, evaluated by size exclusion chromatography (SEC); 2) nucleotide binding (ATPγS and ADP), evaluated by isothermal titration calorimetry (ITC); 3) NTD coupling to nucleotide type, evaluated for apo, ATPγS/AMP-PNP and ADP states by solution-state NMR; 4) ATP hydrolysis, evaluated by solution-state NMR; 5) formation of the ADP·P_i state, evaluated by solution-state NMR; 6) efficient P_i release, evaluated by ATPase assay. An overview of all mutants and performed experiments is presented in Table 6.

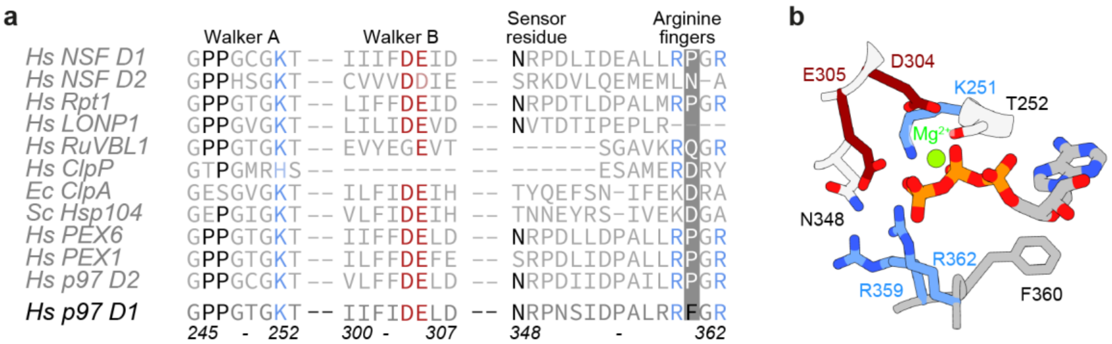


Figure 47. Sequence alignment of p97 D1 residues with other AAA+ family proteins. a) Sequence alignment of Walker A, Walker B, sensor residue and arginine finger motifs show that many of these residues are conserved throughout the AAA+ protein family. Residues studied in this work are highlighted: neutral residues are highlighted in black, negatively charged residues in red and positively charged residues in blue. F360 is emphasised using a dark background. Amino acids numbers of p97 are given below. Hs: *H.sapiens*, Ec: *E.coli*, Sc: *S.cerevisiae*. Accession IDs: P35998, P46459, Q16740, P36776, Q13608, Q9Y265, P0ABH9, P31539 and O43933. **b**) Zoom in on the nucleotide binding pocket in the ATP bound state (PDB ID: 7lmy). Key residues are shown as sticks. Figure adapted from ref. [52].

Table 6. Overview of mutants and experimental techniques used for mutagenesis studies. A given "✓" symbolises that the mutant was evaluated using the respective technique. [a]: ITC measurements of R362A were performed with poor repeatability; results are not shown. [b]: NMR measurements of R362A and K251A were evaluated despite the inability of the mutants to bind nucleotides.

ND1l-construct	Protein sequence motif	Oligomerisation (SEC)	Nucleotide binding (ITC)	NMR	ATPase assay
Wild type	-	✓	✓	✓	✓
K251A	Walker A	✓	✓	✓[b]	✓
P246T		✓	✓	✓	✓
P247A		✓	✓	✓	✓
P247G		✓	✓		✓
P247K		✓	✓	✓	✓
D304N	Walker B	✓	✓	✓	✓
E305Q		✓	✓	✓	✓
N348Q	Sensor residue	✓	✓	✓	✓
R359A	Arginine fingers	✓	✓	✓	✓
R359K		✓	✓	✓	✓
F360A		✓	✓	✓	✓
F360P		✓	✓	✓	✓
R362A		✓	(✓)[a]	✓[b]	✓
R362K		✓			✓
R359A-R362A		✓	✓		✓

5.2.2 Results of biophysical characterisation

Oligomerisation state evaluation

The functional minimal unit of p97 is a homohexamer. Since the nucleotide binding pocket is located between two neighbouring protomers, any deficiency in hexamerisation will subsequently compromise all other functions as well. For this reason, as the first step of the pipeline, the oligomerisation status of mutants was assessed. The technique of choice for this question was size exclusion chromatography, which, while already the last step of the protein purification process, can be used to estimate the size of the construct from the elution volume.

Two protein states were evaluated: apo state and ADP bound state. It is known that nucleotide binding increases the thermal stability of p97 by supporting hexamerisation [65].

Elution profiles of apo state proteins are shown in Figure 48. Most constructs elute as hexamers, similarly to ND1L-wt. A few mutations, however, lead to deficiencies which emerge as two separate groups. In one group (E305Q and P247K), proteins elute as tetramers while showing a strong tailing, hinting at the presence of smaller oligomers. In the other group, which consists strictly of arginine finger mutants, the constructs elute as even smaller oligomers, rather corresponding to trimers.

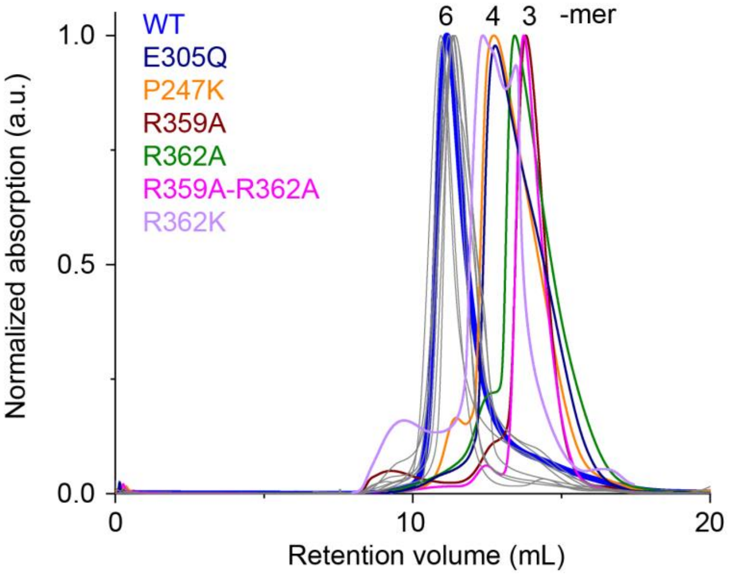


Figure 48. Evaluation of elution profiles of p97 mutants in apo state to assess the oligomerisation status. Apyrase digested ND1L mutants were run over a Superdex 200 Increase to assess the oligomerisation status of the respective apo states. The oligomeric state was evaluated based on an internal calibration (Supplementary Figure 2). Most mutants eluted from the column in as hexamers, similarly to the wild type protein (blue). However, a few mutations lead to a compromised oligomerisation ability, resulting in tetra- or trimers. Figure adapted from ref. [52].

In the case of E305Q, P247K and R359A, the compromised oligomerisation ability recovers when bound to ADP (Table 7). On the other hand, mutations involving R362 are so severe that the oligomerisation status remains faulty even in the presence of ADP.

Table 7. Comparison of oligomeric states of ND1L-mutants in apo and ADP bound states. Oligomeric states have been determined from SEC elution volumes. Most mutants form stable hexamers both in the presence and absence of nucleotides. E305Q, P247K and R359A do not form hexamers in the apo state but recover when bound to ADP. Mutations involving R362 severely affect the oligomerisation so that the hexameric form is not (fully) formed even in the presence of ADP. The oligomeric state of K251A in the nucleotide-bound state cannot be assessed since the mutant is nucleotide-binding deficient. Hexamerisation notation: + (yes), ~ (partial), + (no).

Oligomeric state in	Wild type	P246T	P247A	P247G	P247K	K251A	D304N	E305Q	N384Q	R359A	R359K	F360A	F360P	R362A	R362K	R359A-R362A
apo state	+	+	+	+	-	+	+	~	+	-	+	+	+	-	~	-
ADP state	+	+	+	+	+		+	+	+	+	+	+	+	-	~	-

Evaluation of nucleotide binding abilities

Assessing the contribution of a certain residue to the stabilisation of the ADP·P$_i$ state becomes more ambiguous in cases when the corresponding mutation affects nucleotide binding. Any observed effect regarding a potential deficiency could result from the general lack of nucleotide binding instead of being a characteristic of the ADP·P$_i$ state.

Nucleotide-binding affinities of ND1L mutant constructs towards ATPγS and ADP were measured by ITC. ND1L-wt was used as a reference. The nucleotide-binding deficient K251A mutant [35,47,123] was used as a negative control for ADP binding. Since metal ions are required for ATP binding, ND1L-wt in the absence of Mg^{2+} was used as a negative control for ATPγS binding.

Measurements indicate that all mutants are generally capable of binding nucleotides (Table 8). Binding affinities appear to vary between mutants, typically by a factor of 2 to 3. While the negative control of K251A resulted in a complete lack of ADP binding, ND1L-wt was still able to bind ATPγS in the absence of metal ions, albeit by a five-to-sixfold lower affinity. This emphasises that despite the lack of Mg^{2+}, p97 is still able to bind an ATP analogue, although with a weaker affinity, resulting in an efficient setup for a negative control.

Overall, the affinities obtained for wild-type ND1L are comparable to published values determined by surface plasmon resonance (SPR) for ADP binding but are approximately five times larger than those reported for ATPγS [47]. In contrast, the values measured in this work are about an order of magnitude smaller than published values determined by ITC [54].

Table 8. Nucleotide binding affinities of ND1L mutants measured by ITC. All mutants are capable of binding nucleotides. Values are given as averages ± standard deviation of several experiments. ND1L-K251A serves as a negative control for ADP binding, ND1L-wt without Mg^{2+} as a negative control for ATPγS binding. n: number of technical replicates. n.d.: not detected. [a]: Measurements of R362A resulted in non-reproducible data. [b]: ITC measurements were performed despite the mutants' inability to form stable hexamers in the ADP-bound state. ITC measurements were performed by K. D. Leitl.

p97-ND1L	**ADP K_D (μM)**	**ATPγS K_D (μM)**
wt	0.08 ± 0.01 (n=3)	0.19±0.07 (n=3)
P246T	0.02 ± 0.0007 (n=2)	0.09 ± 0.05 (n=3)
P247A	0.08 ± 0.02 (n=3)	0.18 ± 0.16 (n=2)
P247K	0.05 ± 0.03 (n=3)	0.01 ± 0.004 (n=2)
D304N	0.05 ± 0.002 (n=2)	0.30 ± 0.08 (n=3)
E305Q	0.32 ± 0.001 (n=2)	0.08 ± 0.001 (n=3)
N348Q	0.18 ± 0.04 (n=3)	0.68 ± 0.06 (n=2)
R359A	1.35 ± 0.78 (n=3)	0.04 ± 0.02 (n=3)
R359K	0.11 ± 0.03 (n=3)	0.21 ± 0.16 (n=3)
F360A	0.06 ± 0.03 (n=3)	0.23 ± 0.01 (n=2)
F360P	0.17 ± 0.02 (n=3)	0.14 ± 0.02 (n=2)
R362A[a,b]		0.07 ± 0.05 (n=2)
R359A-R362A[b]	0.05 ± 0.03 (n=3)	0.11 ± 0.02 (n=2)
K251A	n.d. (n=2)	
wt without Mg^{2+}		1.08 ± 0.007 (n=2)

Evaluation by solution-state NMR

Selected mutants were expressed with an ILVM labelling suitable for methyl TROSY experiments and measured by solution-state NMR. For each mutant, a set of four experiments was measured, corresponding to the four nucleotide states of the ATPase hydrolysis cycle: apo, ATP(γS), ADP·P_i and ADP. For three mutants whose behaviour in the ATPγS state differed from the wild-type reference (D304N, F360P, R359A), additionally, spectra in the presence of AMP-PNP were recorded.

Mainly, mutant spectra were compared to wild-type references of the same nucleotide type with a focus on 1) the overall protein conformation, 2) apparent nucleotide identity as determined by methyl probes located close to the binding pocket and 3) NTD positioning (up or down).

A summary of the mutant behaviour as observed by NMR is shown in Table 9. Most mutant spectra showed a close resemblance to the respective wild-type spectra in all four states. Especially in the apo state, all mutants behaved similarly to wild-type ND1L. However, differences arose in the other three states. In the presence of ATPγS, most mutants adopted conformations similar to pre-hydrolysis states, but four constructs showed distinct behaviour which falls into two categories: 1) P247K and R359A both initially showed ATPγS-like spectra, but over the course of the measurement started to partially form the ADP·P_i state. In contrast, R359A only showed pre-hydrolysis spectra when bound to AMP-PNP. 2) Mutants of the other category (D304N and F360P) do not adopt pre-hydrolysis conformations in the first place and show ADP·P_i-like spectra instead, both in the presence of ATPγS and AMP-PNP.

Similarly, under ATPase conditions, most constructs, with the exception of three mutants, form the ADP·P_i state. N348Q and R362A produce spectra which show strong similarities to the pre-hydrolysis spectra and an NTD located in mixed positions. R359A shares similarities with both these two mutants and all other constructs, as it produces spectra with features of ATPγS- and ADP·P_i states.

While almost all mutants were able to adopt ADP-like states in the presence of ADP, R359A, R359K, and R362A additionally displayed mixed conformations of the NTD positioning with the NTD being partially located above the D1 plane (“up”).

Finally, K251A acts completely differently from all other mutants, as all its spectra in the presence of nucleotides are identical, irrespective of the nucleotide type. Instead, these spectra display both features of the apo and ADP states with the NTD located in mixed conformations. The individual results are discussed in detail below.

Table 9. Summary of mutant behaviour as observed by solution-state NMR. Mutant behaviour was assessed in comparison to the wild-type protein. In cases where the mutant and wild-type spectra show strong overlap, mutant behaviour is denoted "as wt". Behaviour in the presence of AMP-PNP is written in italics. P247K and R359A initially produced a pre-hydrolysis state in the presence of ATPγS, but over time, transitioned to the post-hydrolysis ADP·P$_i$ state. In the presence of ADP, R359A, R359K, and R362A produce ADP-like spectra, but the NTD adopts a mixed up-down conformation. The approximate up-to-down ratios are given as percentages.

ND1L-construct	**Apo state**	**Pre-hydrolysis state (ATPγS/ *AMP-PNP*)**	**ATP turnover (ADP·P$_i$)**	**Post-hydrolysis (ADP)**
Wild type	NTD up	NTD up / *NTD up*	NTD down	NTD down
K251A	as wt	similar to ADP and apo states, NTD up and down mixed		
P246T	as wt			
P247A	as wt			
P247K	as wt	as wt; slow hydrolysis to ADP·SP$_i$	as wt	
D304N	as wt	ADP·P$_i$ like / *ADP·P$_i$ like*	as wt	
E305Q	as wt			
N348Q	as wt		high similarity to ATPγS state, NTD up:down split, ~70:30	as wt
R359A	as wt	as wt; slow hydrolysis to ADP·SP$_i$ / *as wt*	similar to ADP·P$_i$ and ATPγS states, NTD up:down split, ~40:60	NTD up:down split, ~10:90
R359K	as wt			NTD up:down split, ~20:80
F360A	as wt			
F360P	as wt	ADP·P$_i$ like / *ADP·P$_i$ like*	as wt	
R362A	as wt	similar to ATPγS and apo states, NTD up	high similarity to ATPγS state, NTD up:down split, ~90:10	similar to ADP and apo states, NTD up:down split, ~60:40

Evaluation of the ATPase activity

ATP processing is carefully regulated by the design of the nucleotide binding pocket. Although the rate of ATP hydrolysis does not correlate with the rate of substrate processing [45,124], it is an important indicator of the severity of a mutation and can hint at the function of the mutated residue.

We measured the ATPase rates at 37°C and 50°C. While 37°C is the physiologically relevant temperature, measurements at 50°C are useful for two reasons. Firstly, to estimate the rate of PEP consumption during the NMR measurement (which is usually performed at 50°C) and, secondly, to probe the mutant for a response to temperature increase. Previously, p97 has been reported to play a key role in responding to heat-induced stress [26,64], and it has been shown that the D1 domain is sensitive to heat stimulation [56]. ATPase measurements at two different temperatures can help to identify residues involved in heat shock responses and/or to assess their general functions.

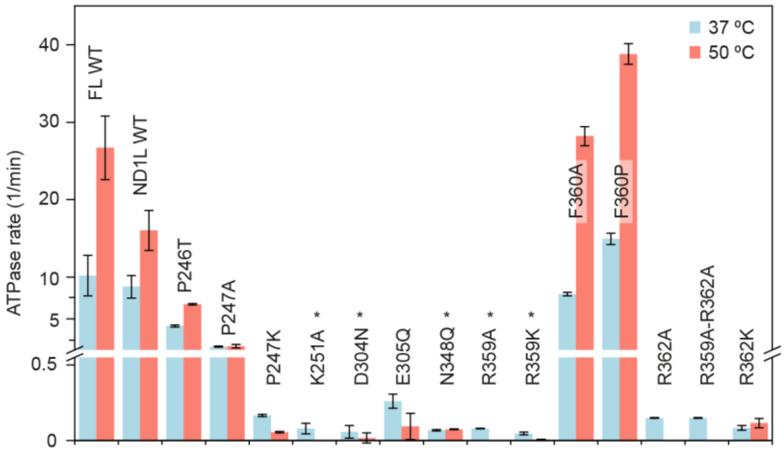


Figure 49. ATPase rates of ND1L mutants. ATPase rates were measured at 37°C (blue) and 50°C (red). ATPase rates are shown as mean values. Error bars represent standard deviation for n = 3-4 technical independent replicates. The ATPase rate of K251A at 50 °C could not be measured due to the low thermal stability. * denotes ATPase inactive mutants.

The results of the ATPase measurements are shown in Figure 49. In most cases, the introduction of a mutation in the binding pocket leads to a severe reduction of the ATPase activity. P246T, R349A and P247A retain some of the original activities, but many mutations lead to strong losses of hydrolysis capabilities, even to the point of complete inactivity. One exception is found with F360X mutants, which at 37°C show ATPase rates comparable to the wild type but strongly gain in activity upon heating. Wild-type p97 itself shows a 1.5 to threefold increase in ATPase activity upon heating to 50°C. This heat-induced increase in ATPase rates was observed for all ATPase-capable mutants as long as their ATPase rate was substantially higher than zero. For instance, P247K and E305Q are slowly active mutants but show a decrease in the ATP hydrolysis rate upon heating. Additionally, all arginine finger mutants (R359X, R362X) except R362K lost their ATPase activity completely at 50°C while it was still above the detection limit at 37°C.

5.2.3 Mutational analysis: results

Walker A motif

The Walker A motif, also known as the P-loop, has a classical sequence of (G/A)XXXXGK(T/S), with X being any residue [16,48]. In the D1 domain of p97, the motif takes the form of 245-GPPGTGKT-252 (Figures 14 and 50). While only the lysine and serine/threonine residues have been explicitly shown to be crucial for Mg^{2+} and ATP binding and thus for the functionality of this motif [16,35,41], the whole sequence is heavily conserved among the members of the AAA+ family (Figure 47). Thus, although the roles of K251 and T252 have been discussed in multiple studies (e.g. [54,59,63,123,125]), very little has been published on the non-canonical Walker A residues, such as P246 and P247 [45].

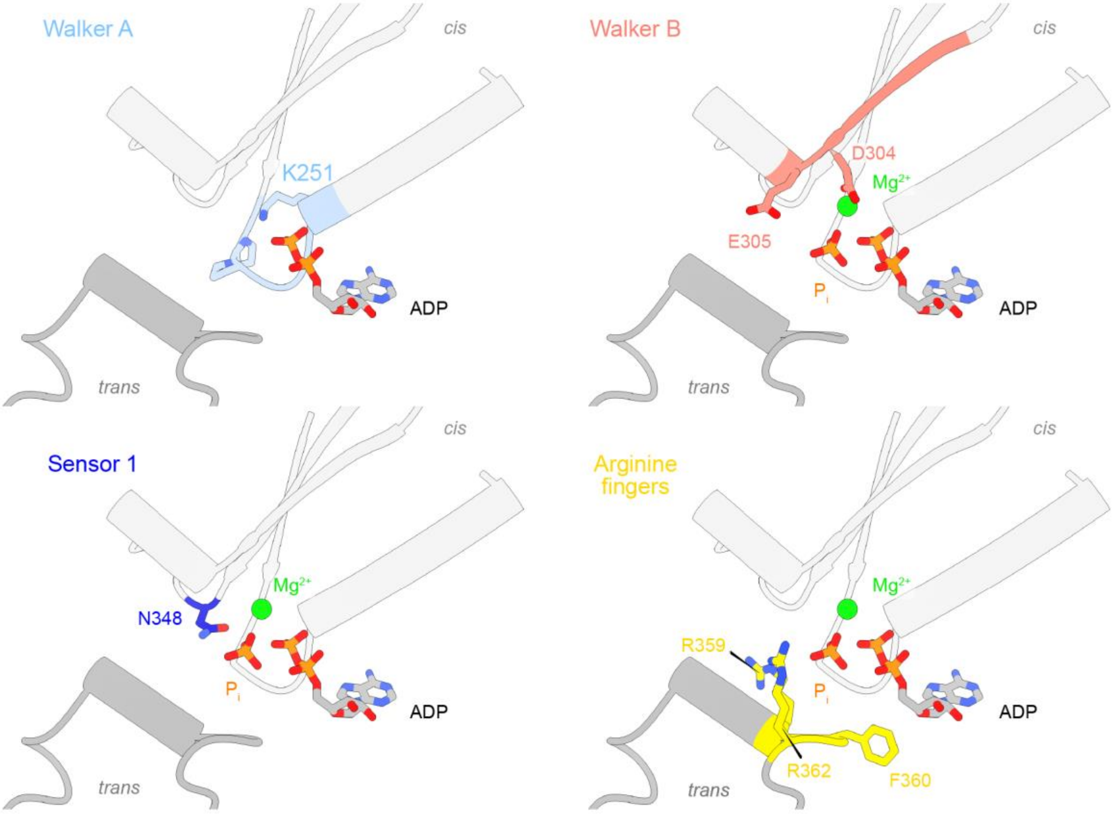


Figure 50. Conserved motifs of p97-D1 in the context of the ADP·P_i state. The canonical protein sequence motifs are shown in colour. Mutated motif residues are additionally shown as sticks. PDB ID: 8ooi.

K251A

ND1L-K251A was selected to act as a negative control for other mutants. The mutant was not found to be hexamerisation impaired, in contrast to the results of some previous studies working with full-length p97 [40,126]. The deficiency in nucleotide binding, however, as measured by ITC (Table 8), is well established in the literature [35,41]: in our hands, ND1L-K251A is the only construct which is purified directly in its apo state, without pre-bound ADP.

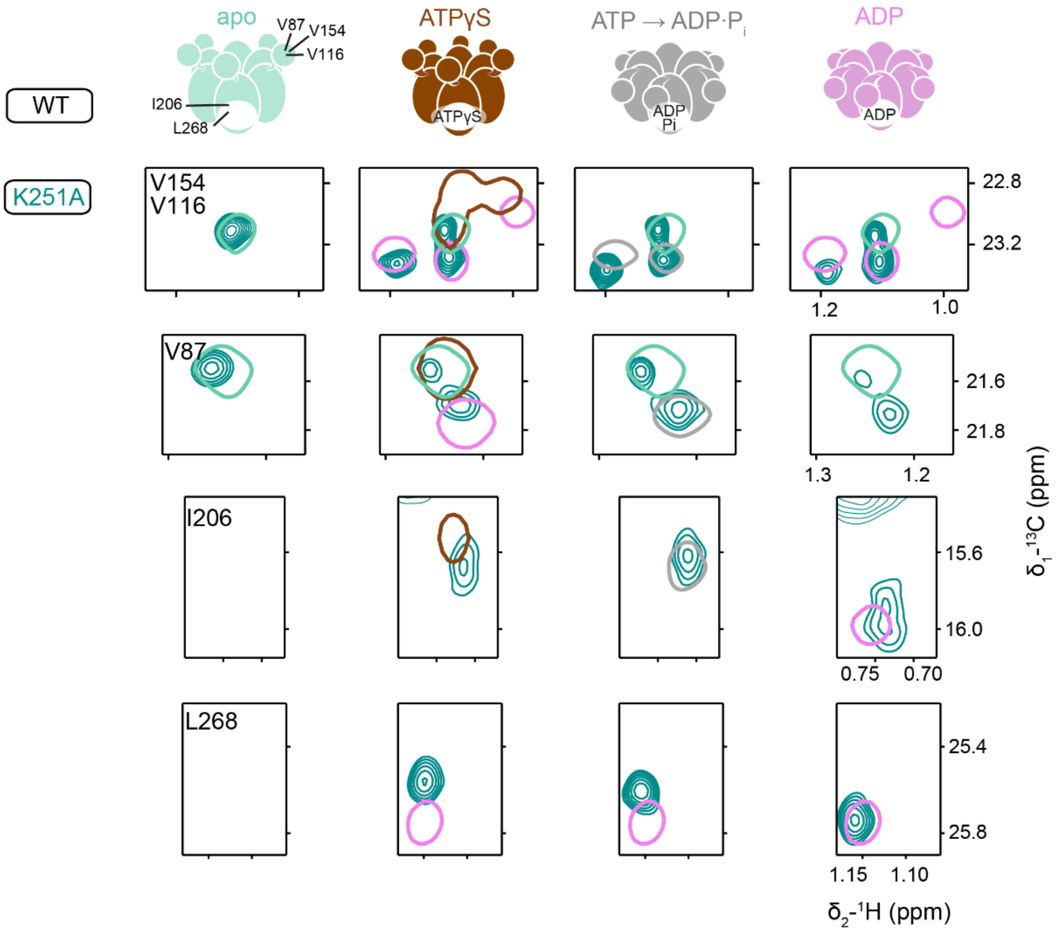


Figure 51. Methyl TROSY spectra of ND1L-K251A in different nucleotide states. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. Residues V154, V116 and V87 report on the NTD position, while residues I206 and L268 inform on the type of nucleotide bound in the binding pocket. In single contours, corresponding wild type spectra are shown. In cases where the mutant behaviour differs from wild type, additionally, ADP spectra are shown. In the apo state, ND1L-K251A adopts a conformation similar to the wild type. In the presence of nucleotides, the mutant produces highly similar spectra, irrespective of the nucleotide type. These spectra share qualities of both the apo and ADP states with the NTD being located in mixed positions. The mutant cannot properly process the nucleotides, as can be seen from methyl probes located close to the binding pocket: I206 correctly identifies the nucleotides, while L268 always detects an ADP-like nucleotide. This figure was previously published in ref. [52].

The lack of nucleotide binding capabilities is reflected in the ATPase assay, where the mutant is classified as “inactive”. While the ADP binding affinity was below the detection limit of ITC (Table 8), a response to the presence of nucleotides is observed by NMR (Table 9, Figure 51). Irrespective of the nucleotide type, ND1L-K251A produces spectra which share strong similarities with both apo and ADP bound states. Consequently, the N-terminal domain is found in mixed up/down positions. The observation of mixed protein states fits the known extremely low binding affinity, thus leading to a slow exchange with the apo state. Nonetheless, the mutant seems to possess nucleotide binding capabilities to some extent but is unable to correctly bind these: for instance, I206 correctly identifies the bound nucleotide state, while L268 always detects ADP-like nucleotides.

Prolines 246-247

The only obvious effect the mutation P246T leads to is the twofold reduction of ATPase activity. Aside from that, the mutant is capable of forming stable hexamers, binding nucleotides, and responding to their presence similarly to that of the wild type. Thus, the exact reason for the decrease in ATPase rate is not obvious but could result from the partial loss of stiffness in the P-loop, previously imposed by the proline.

The reduction in ATPase activity is more pronounced upon mutating P247. While mutating P247 to alanine leads to a tenfold reduction in ATPase activity, both P247G and P247K are considered (almost) inactive. Since both P247A and P247G constructs form stable hexamers and are able to bind nucleotides, it was decided that only P247A by NMR should be analysed. Here, no substantial differences in comparison to wild-type spectra were observed, which is reminiscent of the case of P246T. Again, it appears that the rigidity provided by P-loop prolines is required for fast ATP processing, but both mutants are capable of forming stable ADP·P_i states. Thus, the rate-determining step remains the product release, which must be impaired (based on the ATPase assay). In comparison, the mutation to glycine instead of alanine could introduce undesired flexibility, compromising the overall structure of the loop and explaining the complete loss of activity. Since no NMR measurements of this mutant have been conducted, it remains unknown whether the mutant is, in fact, capable of ATP hydrolysis.

In contrast, P247K shows deficiencies in hexamer formation, but only in the apo state, potentially due to the presence of an additional positive charge inside the binding pocket –

unbalanced without the negative charges of the nucleotide. This misplacement of the new lysine side chain apparently affects product release: under ATPase conditions, P247K forms the ADP·P_i state, although its ATPase rate is extremely reduced. It is conceivable that the additional positive charge leads to a more effective binding of the cleaved, negatively charged phosphate ion, slowing down P_i release. This assumption is supported by the observation that during the NMR measurement in the presence of ATPγS, the protein seems to slowly adopt the ADP·(S)Pi state. ATPγS (auto)hydrolysis at 50°C is unavoidable, but the saturation of the solution with ATPγS should lead to rapid product exchange with fresh ATPγS. The slow accumulation of the ADP·P_i-like state in the ATPγS state measurement is evidence that the cleaved thiophosphate remains trapped inside the binding pocket.

Walker B motif

In literature, there has been disagreement on the exact sequence of the Walker B motif in p97. Canonically, the Walker B motif of the form hhhhD(D/E), where h denotes a hydrophobic amino acid, has been described as being crucial for ATP hydrolysis by activating a water molecule for the nucleophilic attack on the tertiary phosphate [50,51]. In an alternative definition, the focus lies on the DExx box, a series of only four amino acids [54,59]. This disagreement is reflected in the diminished conservation of the protein sequence among the members of the AAA+ family (Figure 47). Still, both approaches put emphasis on the two negatively charged residues, aspartic and glutamic acid (Figure 50). D304 has been shown to interact with the bound nucleotide either via a hydrogen-bonded water molecule or via direct bonding to the Mg^{2+} ion and has been reported to stabilise the Walker A residue T252 [52,54,59]. The more prominent E305 is related to hydrolysis, as its mutation to glutamine has been associated with a deficiency in ATP hydrolysis, which is the reason why E305Q mutants have been readily employed in structural studies [40,44,45,47,55,56].

E305Q

This ubiquity of studies using E305Q constructs to slow down hydrolysis without affecting the rest of the protein is justified, considering that in this work, the mutant displays differences from the wild-type protein only in two categories. First, in the apo state, ND1L-E305Q is only able to partially form hexamers (Table 7). The likely reason for this defect is that in the wild-type form, E305 forms a salt bridge with the *trans*-acting R362, thus stabilising the oligomeric form [59]. This interaction significantly weakens in the absence of the negative charge of the glutamate.

The second distinguished behaviour of the mutant concerns its low ATPase activity. In this case, our observation fits the canonical interpretation of this Walker B mutation, explaining that E305 is crucial in ATP hydrolysis. However, since ND1L-E305Q still forms the ADP·P_i state under NMR conditions, it must be assumed that the rate-determining step remains the product release [55]. Thus, it must be assumed that both the hydrolysis and P_i release are slowed down due to the mutation.

D304N

The introduction of the D304N led to unexpected results. Unlike ND1L-E305Q, ND1L-D304N has no difficulties with hexamer formation. Additionally, the D304N mutant was found to be completely ATPase inactive, even though the nucleotide-binding affinities of both Walker B mutants were determined to be similar.

However, NMR measurements paint a more detailed picture: the mutant shows wild-type-like behaviour in the context of ADP, but in the presence of ATPγS and AMP-PNP, the spectra resemble the ADP·P_i state instead (Figure 52a, Table 9). Overall, no clear distinction can be made between different protein states since all spectra of ND1L-D304N (aside from the apo state) show very strong overlap. Moreover, the mutant is incapable of binding the nucleotide natively, as is evidenced both by the major CSP detected for the L268 methyl probe and by the fact that this peak appears in all four nucleotide-bound states. For comparison, in wt-spectra, this peak only appears in ADP-bound states.

The misinterpretation of nucleotide identity by the mutant as a result of non-native ATP binding explains the observed protein state in the presence of pre-hydrolysis ATP analogues: communication pathways usually relaying ATP binding outwards transduce the presence of an ADP analogue, leading to the NTD down position.

At this point, the origin of this non-native binding should be discussed. Since D304 is associated with Mg^{2+} binding, it can be surmised that the mutant either cannot bind metal ions altogether or the binding leads to an unusual Mg^{2+} coordination. To investigate the hypothesis that the mutant is incapable of Mg^{2+} binding, we measured wild-type ND1L in the presence of ATPγS and EDTA, thus effectively chelating metal ions. Based on ITC measurements, p97 ND1L is able to bind ATPγS even without the metal ion, although with a five to sixfold reduced binding affinity (Table 8). The observed spectrum (Figure 52b) shows a striking resemblance to the ADP bound state. This proves that to assume the pre-hydrolysis conformation, the binding of ATP alone is insufficient; additionally, correct Mg^{2+} binding is required.

Since D304N partially correctly reports on nucleotide identity (see Figure 52a, I206), unlike ND1L-wt with ATPγS and EDTA, which only detects ADP, it should be assumed that the mutant allows metal binding in general, but Mg^{2+} becomes incorrectly positioned due to the lack of D304 coordination. This interpretation fits the classical understanding of the residue function, namely the interaction with Mg^{2+} via hydrogen bonds with a water molecule.

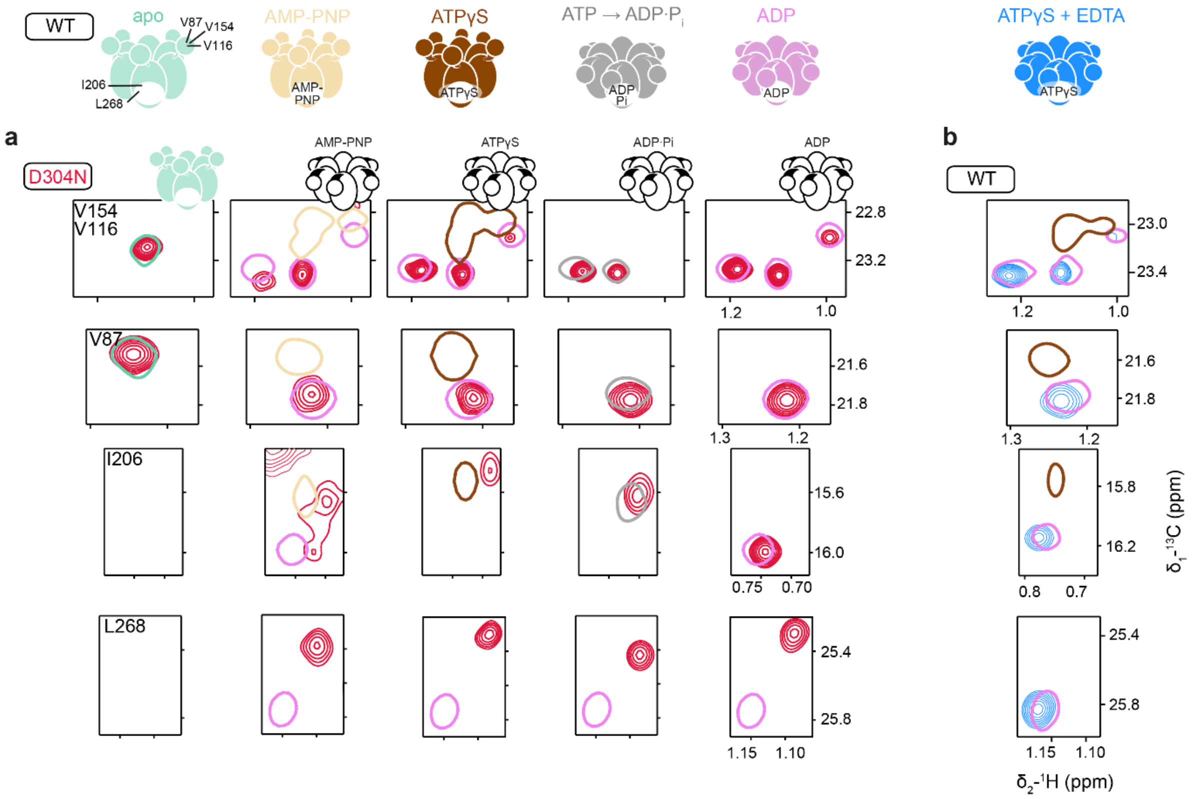


Figure 52. Methyl TROSY spectra of ND1L-D304N in different nucleotide states. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. Residues V154, V116 and V87 report on the NTD position, while residues I206 and L268 inform on the type of nucleotide bound in the binding pocket. In single contours, corresponding wild type spectra are shown. In cases where the mutant behaviour differs from wild type, additionally, ADP spectra are shown. **a**) D304N seems to be unable to properly identify the type of bound nucleotide: In the presence of ATPγS and AMP-PNP the mutant adopts an ADP·P_i like state instead of the pre-hydrolysis conformation. Additionally, the strong CSP observed for L268 indicates non-native nucleotide binding. The appearance of the same peak in all four nucleotide bound spectra supports the notion that the mutant cannot tell nucleotides apart from each other. The identity of the shifted peak was confirmed using a ND1L-D304N-L268M mutant. **b**) In the presence of ATPγS and EDTA, i.e. without Mg^{2+}, wild type p97 produces ADP-like spectra. Figure adapted from ref. [52].

Unfortunately, it remains unclear whether the mutant is capable of ATP hydrolysis at all since it is inactive in the ATPase assay. The similarity of protein spectra in the presence of different nucleotides does not allow us to distinguish the pre-hydrolysis (ATP) state from the post-hydrolysis ADP·P$_i$ state. It is conceivable that both the incorrect metal binding prevents ATP hydrolysis and that the release of the cleaved P$_i$ cannot be initiated for exactly the same reason.

Sensor-1 residue/N348Q

The naming of the protein sequence motif "sensor 1" (canonically Thr or Asn) results from the supposed role of the sensor-1 residue (N348 in p97) to *sense* the nucleotide [58,59], but it has also been reported to play a catalytic part in hydrolysis via hydrogen bonding with the attacking water [41,60].

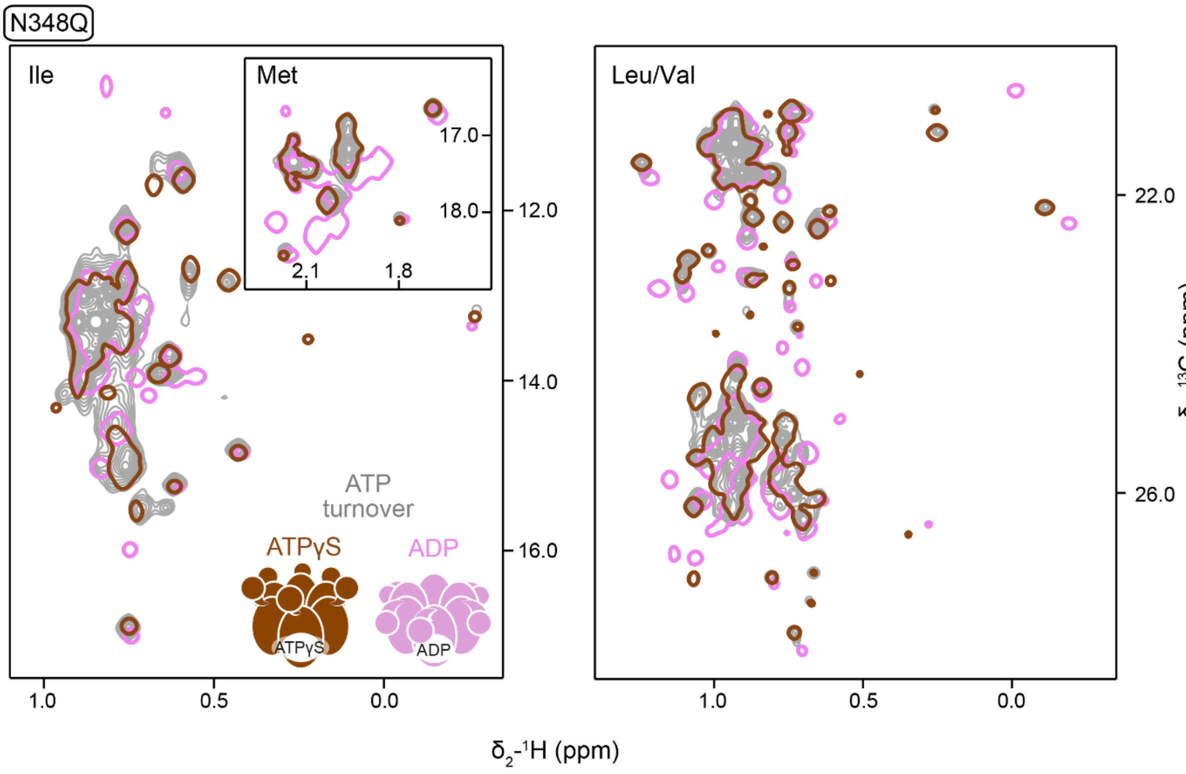


Figure 53. ND1L-N348Q is hydrolysis deficient. Spectral regions of ^{1}H-^{13}C-HMQC spectra of ND1L-N348Q are shown. N348Q can clearly adopt native pre- and post-hydrolysis states in the presence of ATPγS (brown) and ADP (pink), as evidenced by the single contour spectra. However, under ATP turnover the resulting spectrum overlaps greatly with the ATPγS reference, indicating that the protein conformation is that of the pre-hydrolysis state. Only a minor population indicates the presence of the ADP·P$_i$ state conformation. This figure was previously published in ref. [52].

Among all studied constructs, ND1L-N348Q stands out as the only one which shows no deficiencies in either hexamer formation or nucleotide binding but does not form the ADP·P_i state in NMR under ATP turnover. Instead, the mutant displays a pre-hydrolysis-like protein conformation, similar to the state in the presence of ATPγS (Figure 53).

In light of the fact that the protein is able to adopt both native pre- and post-hydrolysis conformations in the presence of the respective nucleotides, the observed ATPγS-like state under ATP turnover leads to the conclusion that the mutant is hydrolysis deficient. This hypothesis is corroborated by the lack of ATPase activity determined in the ATPase assay. MD simulations of the ADP·P_i state give an explanation for this deficiency [52]: in the wild-type protein, N348 hydrogen bonds with the terminal phosphate, positioning it for the nucleophilic attack by the water molecule. The mutation to glutamate extends the side chain length, leading to steric clashes with the catalytically active E305 and R359. Consequently, the catalytically active conformation inside the binding pocket is less frequently observed.

In the full-length context, FL-N348Q and FL-N348A have both been reported to show similarly decreased ATPase rates (25 vs 28% of the wild-type reference) [41,127]. Though the implication of these reports is that the residual ATP hydrolysis activity arises from the D2 domain, it is difficult to envision the ND1L-N348A mutant suffering from steric clashes similarly to N348Q, based on the side chain length alone. This would support the idea that N348 is a catalytically active residue, and its hydrogen bonds with phosphate are crucial for the subsequent attack by the water molecule.

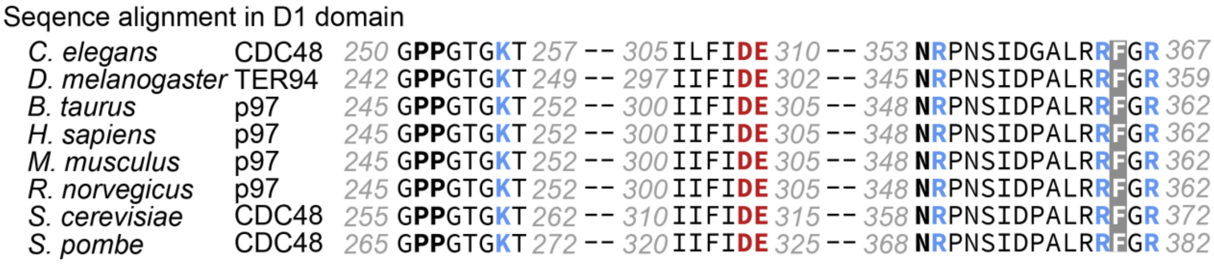
Finally, a minor contribution of the ADP·P_i-like state in the spectrum, leading to the partial NTD down positioning (~40% down), could imply that occasionally, hydrolysis-active conformations can be achieved despite the mutation, an interpretation supported by MD simulations [52].

Second region of homology

The second region of homology (SRH) is a conserved region of AAA proteins that is formed by the β4 strand, a loop, the α4 helix and the loop between the α4 helix and β5 strand and is responsible for communication with the neighbouring subunit (Figure 14) [19,39]. This region includes two specific motifs, starting with the sensor-1 residue and finishing with the arginine fingers.

A *trans*-acting arginine, pointing towards the nucleotide of the neighbouring protomer, is referred to as an arginine finger (R-finger) [19]. The role of such an Arg-finger is to support hydrolysis by interacting with the γ-phosphate of ATP and stabilising accumulated negative charges [50]. In combination with the ISS, the arginine finger motif has been associated with sensing the nucleotide of the adjacent subunit [62]. Members of the AAA subfamily lack another motif, the sensor-2 arginine, but possess instead a second arginine finger residue, typically separated from the first one by two residues (Pro-Gly) [61]. In this thesis, both p97-D1 residues R359 and R362 are referred to as arginine fingers, despite conflicting interpretations over the inclusion of R362 (even though the latter is actually conserved in all AAA+ proteins) [35,36,61,75,116,127].

The ATPase p97 differs from the rest of the AAA+ family members by replacing the aforementioned proline residue with phenylalanine – but only in the D1 domain. This unique substitution is, notably, conserved among all mammalian p97 homologues and in yeast (Figure 54). During the ATPase cycle, F360 has been reported to flip its side chain rotamer according to the nucleotide state in D1, thus acting as a replacement sensor residue [41,63]. Moreover, its rotamer conformation is coupled to the motion of the arginine finger R359 [52].


Seqence alignment in D1 domain

C. elegans	CDC48	250	GPPGTGKT	257	--	305	ILFIDE	310	--	353	NRPNSIDGALRRFGR	367
D. melanogaster	TER94	242	GPPGTGKT	249	--	297	IIFIDE	302	--	345	NRPNSIDPALRRFGR	359
B. taurus	p97	245	GPPGTGKT	252	--	300	IIFIDE	305	--	348	NRPNSIDPALRRFGR	362
H. sapiens	p97	245	GPPGTGKT	252	--	300	IIFIDE	305	--	348	NRPNSIDPALRRFGR	362
M. musculus	p97	245	GPPGTGKT	252	--	300	IIFIDE	305	--	348	NRPNSIDPALRRFGR	362
R. norvegicus	p97	245	GPPGTGKT	252	--	300	IIFIDE	305	--	348	NRPNSIDPALRRFGR	362
S. cerevisiae	CDC48	255	GPPGTGKT	262	--	310	IIFIDE	315	--	358	NRPNSIDPALRRFGR	372
S. pombe	CDC48	265	GPPGTGKT	272	--	320	IIFIDE	325	--	368	NRPNSIDPALRRFGR	382

Figure 54. Sequence alignment of p97 homologues. The p97-D1 phenylalanine between the arginine fingers (F360) is conserved among all organisms. Mutated residues are highlighted in colour. Accession IDs: Q7KN62, Q9P3A7, P25694, P46462, Q3ZBT1, P54812, Q01853 and P55072. The sequence alignment was performed by S. R. Rout using Clustal Omega [128]. This figure was previously published in ref. [52].

Arginine fingers

In this work, multiple arginine finger mutants were created and subjected to the aforementioned screening. While all tested ND1L mutants (R359A, R359K, R362A, R362K and R359A-R362A) are able to bind nucleotides, most of them display severe hexamerisation deficiencies (Tables 7 and 8). R362 appears to be crucial for oligomer stability, as its mutation leads to more severe deficiencies, both in the apo and ADP-bound states. This observation agrees with previous reports [59,127], which highlight the importance of the salt bridge R362 forms with E305. This significance is reflected in the partially compromised hexamerisation of ND1L-E305Q in the apo state, although the extent of the deficiency is reduced in comparison to ND1L-

R362A. Potentially, hydrogen bonds formed between Q305 and R362 can partially compensate for the lost electrostatic interaction. This hypothesis is corroborated by the observation that the R362K mutant is only moderately oligomerisation deficient, implying that the positively charged lysine can support the native salt bridge to some extent. Alternatively, this hypothesis could also be tested with the E305A mutation, but the ND1L-E305A construct has not been characterised in the literature. In contrast, a charge flip (R362E) leads to a hexamerisation deficiency even in the full-length context, even though full-length constructs are considerably more stable [127].

R359 is known to contact the nucleotide, but no reports of salt bridges have been published. In this work, ND1L-R359A only displayed deficiencies in hexamer formation in the apo state, contrary to previous studies [127]. It is not entirely clear why the lack of the Arg positive charge would only affect the stability of the empty binding pocket and why this residue specifically would contribute to hexamerisation. It is possible that the mutation leads to a strong hexamerisation deficiency, which is partially compensated by nucleotide binding in the ADP-bound state, leading to a normal elution volume during SEC. It has been shown that apo state p97 is less stable than its nucleotide-bound form [65]. Again, the oligomerisation defect was compensated in the R359K mutation, emphasising that the positive charge of the amino acid 359 is important for hexamer formation. Unsurprisingly, the double mutant ND1L-R359A-R362A showed the strongest oligomerisation defects out of all constructs.

Since the goal of this work is to study the contributions of individual amino acids to the stability of the ADP·P_i state, we decided to investigate mutants of both arginine fingers by NMR. However, considering that mutations of R362 lead to strong hexamerisation defects which could not be fully compensated (R362K), we decided to focus only on one single mutant R362A, discarding R362K and R359A-R362A.

R359X

Under NMR conditions, ND1L-R359K protein states share strong similarities with the wild-type protein, irrespective of the nucleotide state. Indeed, the formed ADP·P_i state is highly stable, as seen by the extremely low ATPase rates. MD simulations of the R359K mutant suggest that the protein is trapped in the post-hydrolysis state, with K251 and K359 symmetrically stabilising the nucleotide. This residue positioning is highly similar to the

ADP·P$_i$ state A, implying that the mutation prevents a successful transition to state B [52]. This ADP·P$_i$ trapping is further evidenced by the low mobility of the mutant sensor loop (see Section 5.6.2) [52]. In contrast, NMR spectra of R359A show several dissimilarities in comparison to the wt protein. In the presence of ATPγS, ND1L-R359A starts with the pre-hydrolysis-spectra but slowly forms an ADP·SP$_i$-like state (Figure 55). However, under ATP turnover, the ADP·P$_i$ state is only partially formed, as the protein conformation additionally shares similarities with the pre-hydrolysis state, resulting in a split up-down position of the NTD. These observations fit the classical interpretation of an Arg-finger being involved in hydrolysis [50]. More specifically, the positive charge appears to be necessary to counterbalance the phosphate charge since R359K is able to fully hydrolyse ATP and form the ADP·P$_i$ state, while R359A displays a mixture of pre- and post-hydrolysis states. Additionally, R359 seems to be involved in product release, which is evidenced both by the trapped ADP·P$_i$ state of R359K and by the slow formation of the ADP·SP$_i$ state of R359A. No ADP·P$_i$-like state is trapped by R359A in the presence of AMP-PNP, although the nucleotide is subjected to thermal hydrolysis similarly to ATPγS (see Section 4.2.1). This could imply that the trapping of the post-hydrolysis state of AMP-PNP (AMP-NH$_2$·P$_i$) has a different set of requirements in which R359 plays a less crucial role.

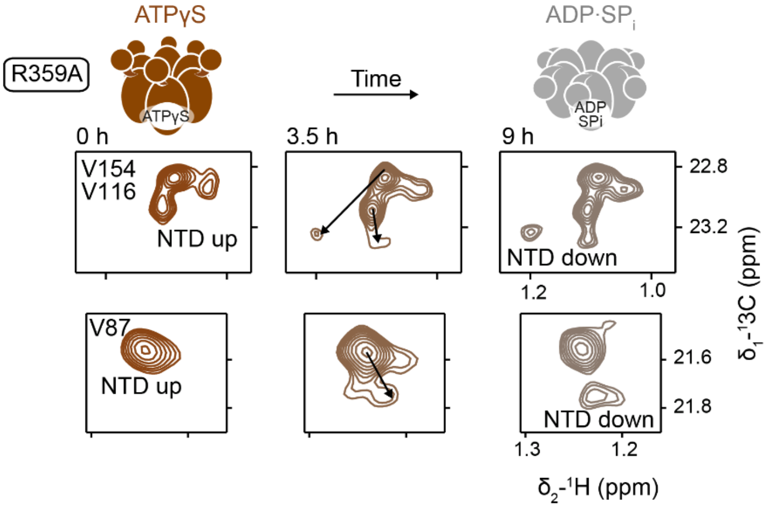


Figure 55. ND1L-R359A transitions from the ATPγS to ADP·SP$_i$ state. Spectral regions of ^{1}H-^{13}C-HMQC spectra of ND1L-R359A are shown. Residues V154, V116 and V87 report on the NTD position. R359A can clearly adopt the native pre-hydrolysis state in the presence of ATPγS (brown), but over the course of few hours starts to form the ADP·SP$_i$ state (grey).

Even though both ND1L-R359A and ND1L-P247K show similar features in slowly producing ADP·SP$_i$ states in the presence of ATPγS, whether the mechanistic origins are identical is unknown. In the case of P247K, it is hypothesised that the ADP·P$_i$ state experiences an enhanced stabilisation by the new positive charge. In contrast, R359A could be lacking a part of the native product release mechanism.

R362A

As discussed above, ND1L-R362A displays a significant deficiency in oligomerisation. This is reflected in the NMR spectra in which the static nucleotide states show partial features of the apo state. Generally, the spectra could be described as showing mixtures of both nucleotide-bound and free states and, consequently, the NTD in mixed conformations. Unsurprisingly, under ATP turnover, the mutant does not form the ADP·P_i state, but instead, the spectrum shares similarities with the ATPγS-bound conformation. Thus, the deficiency in hexamerisation leads to an ATP-locked, pre-hydrolysis state of the protein, which is reflected in the reported drastic decrease of cooperativity [126]. This major defect prevents further analysis of the mutant down the introduced pipeline. Consequently, it cannot be ruled out that R362 has an additional function in ATP hydrolysis and/or product release, following the classical interpretation of an arginine finger. For instance, in MD simulations, R362 has been shown to coordinate the cleaved P_i [52], it is known to play a role in interprotomer communication via electrostatic interactions with D333 (ISS) [129], and it affects the ability of p97 to bind polyubiquitin chains [127].

F360X

Both mutations of F360 share a distinction from all other investigated mutants: the mutation led to a gain in ATPase activity in comparison to the wild type, which was even more pronounced at 50°C (Figure 49). Since both mutants were able to form the ADP·P_i state under ATP turnover conditions, it must be assumed that the mutations lead to an increased rate of product release. Overall, no deficiencies in hexamerisation or nucleotide binding were observed, and ND1L-F360A displayed wt-like spectra under NMR conditions. Based on these observations alone, it was unclear what caused the increase in ATPase activity. The mutation of F360 to proline, inspired by the conserved residue among the AAA+ family members (incl. p97-D2), however, led to an unexpected result. In the presence of both ATPγS and AMP-PNP, F360P adopts an ADP·P_i-like conformation (Figure 56). Unlike ND1L-D304N, however, which suffered from incorrect Mg^{2+} binding and subsequent ATPase inactivity, ND1L-F360P is highly active, and the lack of NTD response to ATPγS binding is of a different origin. The similarity of both spectra in the presence of slowly hydrolysable nucleotides makes it unlikely that F360P is simply very efficient at hydrolysing ATPγS but incapable of ADP·SP_i release.

Instead, it appears more likely that the mutation prevents proper communication between the nucleotide binding pocket and the NTD, leading to a lack of NTD response. In the ADP·P_i state, F360 rotamers are coupled to the side chain orientation of R359 and the positioning of P_i and Mg^{2+} [52]. One of the rotamers contacts the α7 helix ($\alpha_{407–423}$), which leads from the binding pocket to the N-terminal domain and could be (partially) responsible for NTD motion control [52]. In the MD simulation of ND1L-F360P, the sensor loop (Section 5.6.2) displays increased mobility of all subunits, regardless of the nucleotide state [52]. This could signify the loss of an inter-protomer co-regulation of hydrolysis, explaining the increased ATPase rate and compromised NTD motion. Since the NTD response defect is not observed for F360A, it must be assumed that the imposed rigidity of the arginine finger loop due to the introduction of a proline residue is responsible for it [129]. In the context of p97-D1, the otherwise conserved proline needed to be replaced by the phenylalanine to tightly regulate the ATPase cycle.

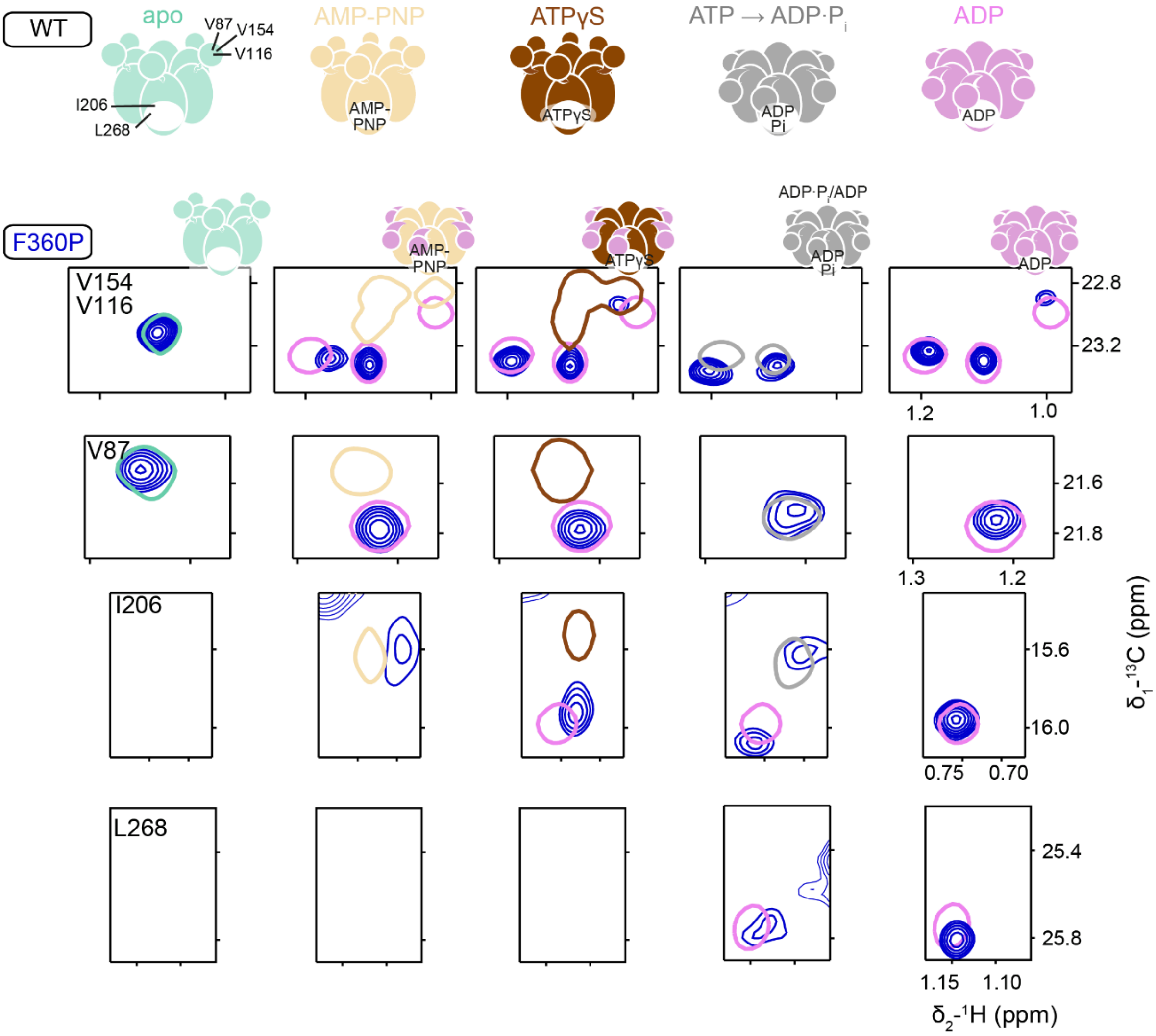


Figure 56. Methyl TROSY spectra of ND1L-F360P in different nucleotide states. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. Residues V154, V116 and V87 report on the NTD position, while residues I206 and L268 inform on the type of nucleotide bound in the binding pocket. In single contours, corresponding wild type spectra are shown. In cases where the mutant behaviour differs from wild type, additionally, ADP spectra are shown. F360P is unable to form a pre-hydrolysis conformation and adopts ADP·P$_i$-like states in the presence of AMP-PNP and ATPγS. The mutation leads to an increased product release rate, leading to a mixture of ADP·P$_i$ and ADP states during ATP turnover. Figure adapted from ref. [52].

5.2.4 Mutational analysis: discussion

Using the described pipeline, we investigated the functions of several conserved residues, which are part of important protein sequence motifs. The results are summarised in Figure 57.

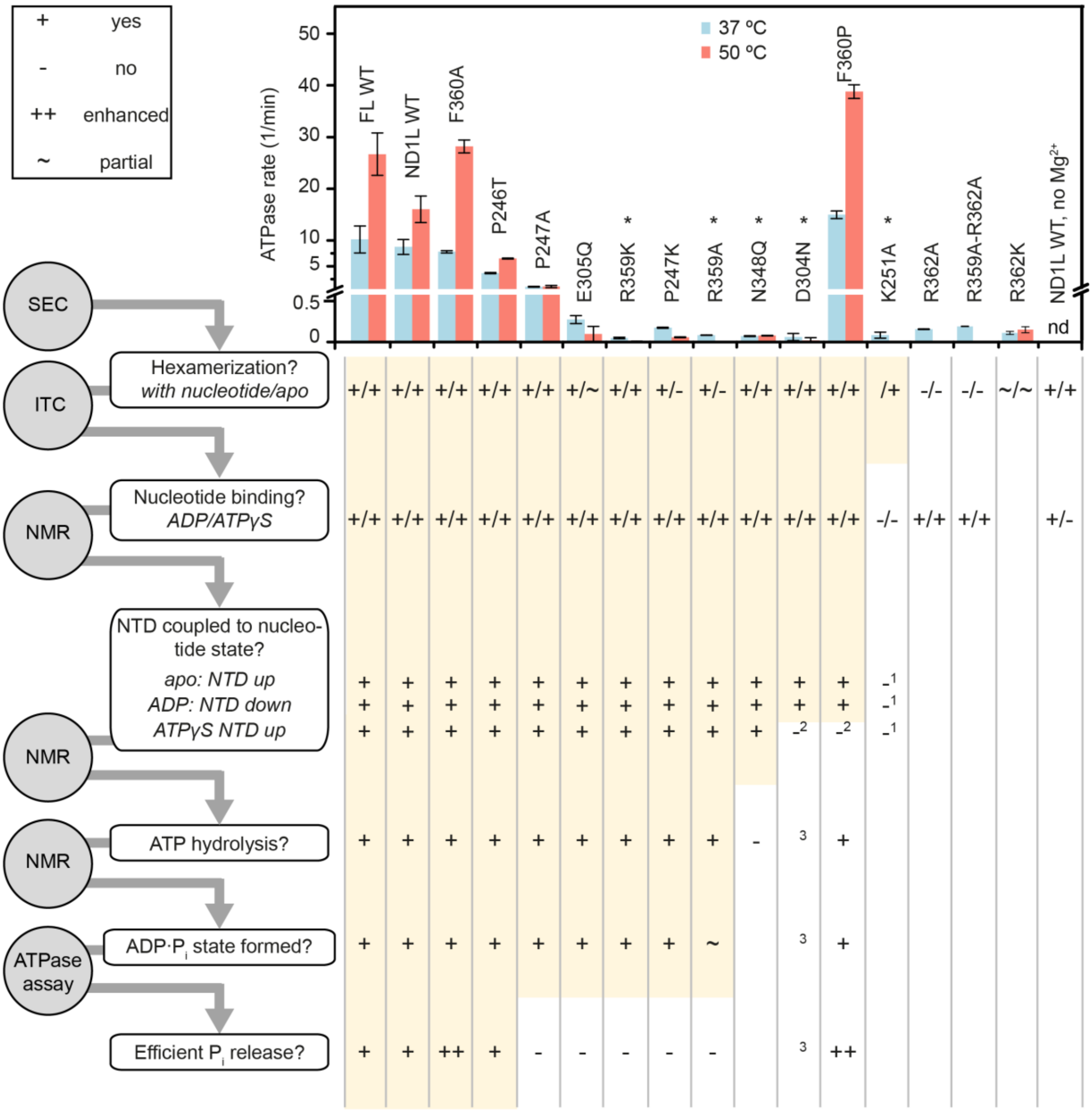
We confirmed that the arginine fingers R359 and R362 are involved in hexamer formation, and the mutation of the latter produces more severe deficiencies. Electrostatic interactions and salt bridges seem to play a role in oligomerisation, and the lysine substitution of the missing positive charge from arginine (R359K, R362K) can partially alleviate the defects.
Consequently, deficiencies in hexamerisation lead to further defects down the line: arginine finger mutants produce spectra which imply impaired nucleotide binding (R362A in the ADP state) and hydrolysis (R362A and R359A under ATP turnover). Conversely, these measurements also confirm that the lack of hexamerisation in the apo state does not prevent nucleotide binding.

In the ATPase assay, R-finger mutants were inactive or barely active. Partially, this can be tied to poor hexamerisation. Since the nucleotide is bound between two neighbouring protomers, defects in hexamer formation would affect the functionality of the nucleotide binding pocket. Additionally, these defects can play another role in the heat-activation measurements: all mutants displaying any kind of oligomerisation deficiencies either show reduced (E305Q, P247K), non-detectable (R359A, R362A, R359A-R362A) or identical (R362K) ATPase rates at 50°C in comparison to the 37°C values. Considering the decreased stability of ND1L in the apo state, it is conceivable that the ATPase measurements at 50°C do not reflect the actual mutant activities, but the lower values are the result of protein melting and/or aggregation [65]. The only exception to this rule is the R359K mutant, which shows a decrease in ATPase activity at 50°C but does not display any oligomerisation deficiencies.

Classically, mutations of the Walker A residue K251 have been associated with nucleotide binding deficiencies [63,123,125]. Although we were also unable to detect ADP binding of ND1L-K251A using ITC, NMR experiments show that the mutant, in fact, reacts to the presence of nucleotides, although with a reduced binding affinity, leading to apo state-like spectra. Furthermore, the mutant seems to be incapable of differentiating nucleotides, as all spectra (aside from the apo state) look identical, regardless of the present nucleotide. It is unclear, however, whether this deficiency in nucleotide identification is caused by an incorrect nucleotide binding inside the binding pocket or due to a fast exchange between the apo and nucleotide-bound states. In contrast, mutations of the Walker A motif prolines (P246T, P247A) did not compromise nucleotide binding but only caused a strong reduction in product release rates.

Most mutants show a coupling of the NTD position to static nucleotide states. Two exceptions need to be highlighted: ND1L-D304N and ND1L-F360P fail to adopt pre-hydrolysis states in the presence of ATPγS, but due to different reasons. The mutation of the Walker B residues (D304N) compromises Mg^{2+} binding, without which p97 is unable to differentiate between ATP(γS) and ADP. The severity of this mutation prevents a full categorisation of the mutant since it is unclear whether the mutant is trapped in the pre- or post-hydrolysis state under ATP turnover. In contrast, the hyperactive mutant F360P is fully capable of hydrolysing ATP due to an increased product release rate. The sidechain of F360 is coupled to the movement of the arginine finger R359 and acts as a sensor of the nucleotide state [52,63]. By contacting the α7 helix ($\alpha_{407\text{–}423}$) by one of the rotamers (χ_1~180°), F360 can transfer the information regarding the nucleotide state towards the N-terminal domain. The loop stiffness imposed by replacing phenylalanine with the classically conserved proline prevents information transfer via this network.

Overall, the ADP·P_i state appears to be very stable since most mutations allow the formation of this post-hydrolysis state under ATP turnover, provided the mutant is able to hexamerise and bind nucleotides. The only exception is found in the ND1L-N348Q mutant, which fails to produce an ADP·P_i state due to its deficiency in nucleotide hydrolysis. In summary, one could postulate that the NTD down conformation alone is not sufficient for ATP hydrolysis. This is seen in the behaviour of N348Q, R362A and D304N mutants, which are all ATPase inactive but have their NTD partially located in the down positions in the presence of the ATP regeneration system.



	FL WT	ND1L WT	F360A	P246T	P247A	E305Q	R359K	P247K	R359A	N348Q	D304N	F360P	K251A	R362A	R359A-R362A	R362K	ND1L WT, no Mg^{2+}
Hexamerization? *with nucleotide/apo*	+/+	+/+	+/+	+/+	+/+	+/~	+/+	+/-	+/-	+/+	+/+	+/+	/+	-/-	-/-	~/~	+/+
Nucleotide binding? *ADP/ATPγS*	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	-/-	+/+	+/+		+/-
NTD coupled to nucleotide state? *apo: NTD up* *ADP: NTD down* *ATPγS NTD up*	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +	+ + -^{2}	+ + -^{2}	-^{1} -^{1} -^{1}				
ATP hydrolysis?	+	+	+	+	+	+	+	+	+	-	^{3}	+					
ADP·P$_i$ state formed?	+	+	+	+	+	+	+	+	~		^{3}	+					
Efficient P$_i$ release?	+	+	++	+	-	-	-	-	-		^{3}	++					

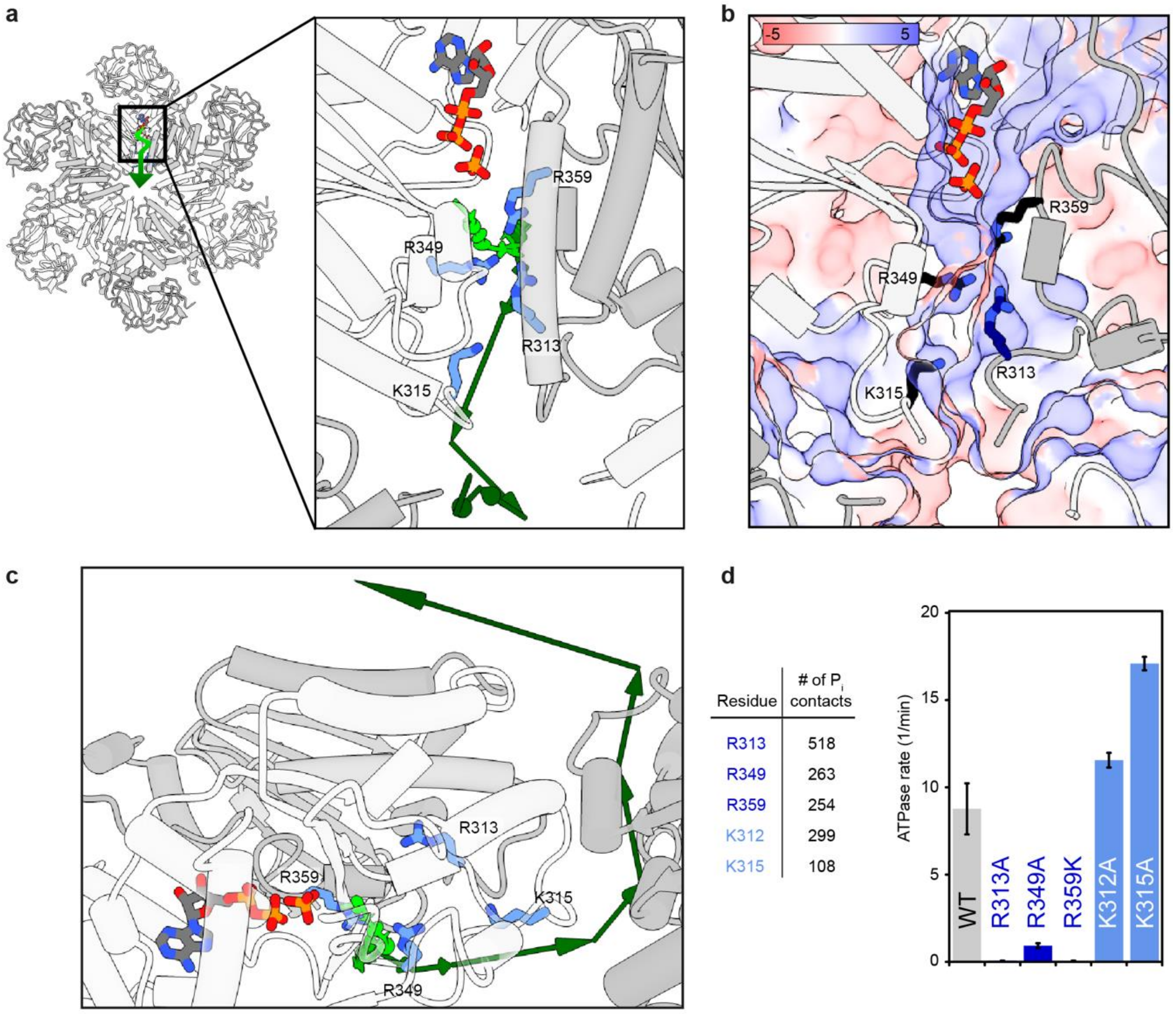
Figure 57. Summary of structural and functional defects of ND1L mutants. Left: Mutant behaviour in respect to their structural and functional defects was assessed in a pipeline going from top to bottom. The biophysical methods are shown in grey circles, the properties in white boxes. Right top: ATPase rates of the mutants. Right bottom: Assessment of mutant behaviour was continued following the pipeline as long no defects were encountered. Categories: behaviour in agreement with wt (+, yellow background), partially in agreement (~) or different from the wt (-). Abbr. SEC: size-exclusion chromatography, ITC: isothermal titration calorimetry, NMR conformational analysis, ATPase rate measurement; Annotations: *: Mutant deemed ATPase inactive; nd: no data; ^{1}: spectral change detected in the presence of any nucleotide, but indifferent to type, NTD position mixed; ^{2}: NTD in 'down' position in presence of ATPγS and AMP-PNP; ^{3}: categorisation is impossible, since ADP·P$_i$ state cannot be distinguished from the ATP state by NMR. The corresponding NMR spectra are shown in Figures 16, 51-53 and 56. The nucleotide dissociation constants are listed in Table 8. Data are presented as mean values. Error bars represent s.d. for n=3-4 technical replicates. ATPase rate of K251A at 50 °C could not be determined due to low thermal stability.

5.3 Pathway of phosphate dissociation

5.3.1 Phosphate dissociation channel leading to the central pore

A much-neglected part of the ATPase cycle is the study of the cleaved P_i release from the nucleotide binding pocket. This process must be the connecting step between the ADP·P_i and ADP states. Its duration impacts the lifetime of both states and could even define the boundaries between them.

The ADP·P_i state has a lifetime of multiple seconds [55] , but MD simulations can only cover much shorter time scales. In the MD simulation of the ADP·P_i state, lasting 2 µs, no direct evidence of phosphate release was seen [52]. While the small size of the inorganic phosphate leads to the belief that the ion could diffuse out from any pore opening, there is no direct evidence for that assumption. In order to facilitate the P_i dissociation process, the stabilising Mg^{2+} ion was removed, leading to occasional phosphate transfer out of the binding pocket. Hereby, the phosphate was transported from the binding pocket through an internal channel towards the central pore, from where the ion separated and diffused away from the protein (Figure 58a,c). Indeed, electrostatic potential (APBS) calculations of p97 show that the internal channel is highly positively charged, facilitating the transport of the negatively charged ion (Figure 58b).



Residue	# of P_i contacts
R313	518
R349	263
R359	254
K312	299
K315	108

Figure 58. Pathway of the cleaved P_i through the internal channel towards the central pore. a) Zoom-in on the nucleotide binding pocket. MD simulations observe a P_i release trajectory from the nucleotide binding pocket through an internal channel towards the central pore, from which the phosphate is ultimately released. The trajectory is shown in green, while positively charged residues coming in contact with the P_i in the simulations are shown in blue. **b**) Electrostatic potential calculations (APBS) of the p97 confirm the existence of a positively charged internal channel, leading from the nucleotide binding pocket towards the central pore. The view corresponds to the one shown in **a**). **c**) shows the side view of the zoom-in from panel **a**). **d**) Residues with the highest numbers of contacts with the P_i in the MD simulation were subjected to mutational analysis. Contacts are defined as distance below 3 Å in a single frame. Only the mutation of Arg (but not Lys) residues lead to a decrease in ATPase rates. Figure adapted from ref. [52].

5.3.2 Mutational analysis of residues involved in the P_i release

In the MD simulation of the P_i release, several residues come into regular contact with the transported phosphate, implying an important contribution to the release mechanism. Amino acids with the highest amount of number of contacts in the simulation are presented in Figure 58d. As shown in Figure 58a, these amino acids line the internal channel, creating an obvious transport pathway for Pi, which could be "handed over" from one residue to another. These residues – arginines and lysines – were subjected to mutational analysis, removing the positive charge. Afterwards, the mutants – all capable of hexamerisation and nucleotide binding – were investigated using the ATPase assay to determine whether their positive charges were indeed important in the ATP hydrolysis cycle. The results show that only the mutation of arginines, but not lysines, leads to a drastic decrease of the ATPase activity (Figure 58d). These results particularly support the hypothesis of the channel transport since (other than R349 and R359) R313 is clearly located outside of the binding pocket and should not influence the hydrolysis process itself.

Following these results, we wanted to determine whether the exact positioning of the arginines was of relevance and whether the ATPase rates could be enhanced by replacing lysines with arginines. To achieve this, several additional mutants were created. Seeing how the ND1L-R313A mutant was ATPase inactive, we investigated the possibility of substituting the missing charge by creating R313A-double mutants with the neighbouring residues (K312 and E314) being mutated to arginines. As is seen in Figure 59, both double mutants remained ATPase inactive, emphasising that the exact positioning of R313 cannot be changed.

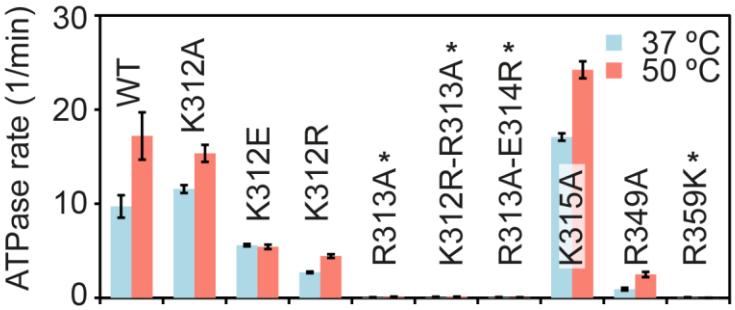


Figure 59. ATPase rates of P_i release associated mutants. Residues with the highest contact numbers to the phosphate during P_i release were subjected to mutational analysis. The mutation of R313, R349 and R359 leads to a drastic reduction of ATPase rates. The arginine positive charge at R313 cannot be replaced with another neighbouring arginine in positions 312 or 314. P_i release rates cannot be increased by mutating involved lysines to arginines (K312R). ATPase rates were measured at 37°C and 50°C. * denotes ATPase inactive mutants.

At the same time, ATPase rates cannot be sped up by replacing lysines with additional arginines (K312R) or flipping the charge (K312E) since both mutations lead to a decrease in ATPase activity. It is unclear how exactly these K312X mutations affect the protein in order to reduce the ATPase rates. Possible explanations could be impeding the phosphate transfer through charge reversal (K312E) but also trapping the released phosphate with the new arginine (K312R). It is also worth noting that K312 is located near the central pore at the interface between two neighbouring P_i release channels and thus could fulfil additional roles mediating between two binding pockets.

The involvement of K315 in the P_i release is uncertain since the residue is known to be trimethylated *in vivo* [130], thus eliminating the positive charge. Moreover, other studies have reported decreased ATPase rates for K315A/L mutants in the full-length context, highlighting potential other functions of this residue which were neglected here [126,130].

Mutations of all aforementioned arginines lead to drastic reductions in ATPase assays. Out of the three amino acids, mutation of R349 appears to have the least effect since the ATPase activity is not completely abolished, and the NMR spectra of the corresponding mutant do not significantly differ from those of the wild type (Section 5.7.3). In comparison, mutations of R359 have a profound effect on the ATPase ability of the protein, as thoroughly discussed in Section 5.2.3. Although in the MD simulation, both R349 and R359 show similar contact numbers, R349 could play a more significant role in P_i transport, considering that R359 is directly involved in the stabilisation of the ADP·P_i state, thus potentially leading to overcounting. Considering that an arginine in position 359 is a key component in transitioning from ADP·P_i state A to state B, it is highly likely that this transition effectively initiates the P_i release [52]. R313, however, is located far away from the nucleotide binding pocket, but its mutation renders the mutant completely inactive without being able to counteract it. Unsurprisingly, this drastic defect is reflected in the high contact number in the MD simulation, highlighting the importance of P_i-R313 interaction and the overall tight coordination of the phosphate release process.

5.4 ADP·Pi mimicking

The defining feature of the ADP·P_i state is, undoubtedly, the presence of the phosphate ion. Studying this intermediate state inadvertently involves the capture of the phosphate ion inside the nucleotide binding pocket or finding substitutes to mimic this state. In previous studies, trapping of the post-hydrolysis state has been achieved by supplying proteins with external phosphate [70], arsenate [69] or sulphate [41,70,72] or through long incubations in the presence of ATP [69,71,82,131]. However, these studies may not represent the relevant short-lived intermediates after ATP hydrolysis, and none of the structures involves members of the AAA+ ATPases.

Even so, these approaches do not seem to work with p97: incubating p97 with ATP leads to rapid ATP hydrolysis and to observation of the ADP bound state (the effect is directly visible in the NMR measurements in the presence of the regeneration system starting from the moment PEP runs out). The other method involves the incubation of ADP-bound p97 with external phosphate. Solution-state NMR confirmed that the addition of up to 200 mM inorganic phosphate to the ADP-bound p97 did not change the protein state (Table 10) [55]. We also tested the hypothesis that the ADP·P_i state can be enforced with the help of additional ions which are present in the ATP turnover measurement setup (Mg^{2+}, K^+). Still, the ADP state remained unchanged (Table 10). The same observation was made when the ions were added to ADP-bound ND1L-R359K and ND1L-P247K. Both these mutants effectively trap the ADP·P_i state in the presence of the ATP regeneration system but are incapable of product release. This further proves that the ADP·P_i state cannot be artificially induced in p97.

Cryo-EM structures of ADP·P_i and ADP states are highly similar, but the secondary structure of the sensor loop in the ADP·P_i state is more akin to its conformation in the pre-hydrolysis state, confirming that the subsequent phosphate release is linked to final structural rearrangements [52]. These structural changes could potentially prevent phosphate from entering the nucleotide binding pocket. Investigations of F-actin have shown that local side chain movements are sufficient to affect the conformation of the protein interior, thus altering pore sizes through which phosphate release is controlled [132].

Table 10. Induction of the ADP·P_i state in p97 through the addition of external phosphate. Summary of solution-state NMR experiments performed on ND1L-wt to test for its ability to produce an ADP·P_i bound state upon incubation with external ions. Three different initial protein states were chosen for the experiments: ADP bound, apo and ATPγS bound. On the left, tested conditions are shown. The ADP·P_i state cannot be artificially produced from the ADP state. Additional experiments were performed with R359K (R) and P247K (P) mutants, leading to the same observations. [a]: ND1L-wt experiments starting from the ADP bound state were performed with 50 mM P_i.

Ions to be added	Initial protein state		
	ADP bound	**apo**	**ATPγS bound**
+100 mM P_i	ADP state [a,P,R]	ADP·P_i state	
+100 mM P_i +1 mM EDTA		ADP·P_i state	
+4 mM $MgCl_2$ +100 mM P_i (or $HAsO_4^-$/SO_4^{2-})		ADP·P_i state	ADP·P_i state
+100 mM P_i +50 mM KCl	ADP state [a,R]		
+100 mM P_i +50 mM KCl +4 mM $MgCl_2$	ADP state [a,P,R]		
+100 mM KCl			ATPγS state

In contrast, an ADP·P_i-like state could be induced starting from the apo state. Phosphate ion titrations to wild-type ND1L showed that 75 mM P_i was sufficient to induce a complete NTD down motion (Figure 60a). Moreover, the resulting spectra share strong similarities with the native ADP·P_i state. Interestingly, the addition of phosphate ions alone was sufficient to observe the ADP·P_i-like state; Mg^{2+} was not required (Table 10). MD simulations of ND1L-wt in which the nucleotide was replaced with two inorganic phosphates confirm the stability of the ADP·P_i-mimic state (Figure 60b) [52]. Specifically, in the simulations, the ions are bound in the same positions as the βP-ADP and cleaved phosphate, but the presence of Mg^{2+} is not required. Instead, a potassium ion from the buffer solution provides the bridging positive charge.

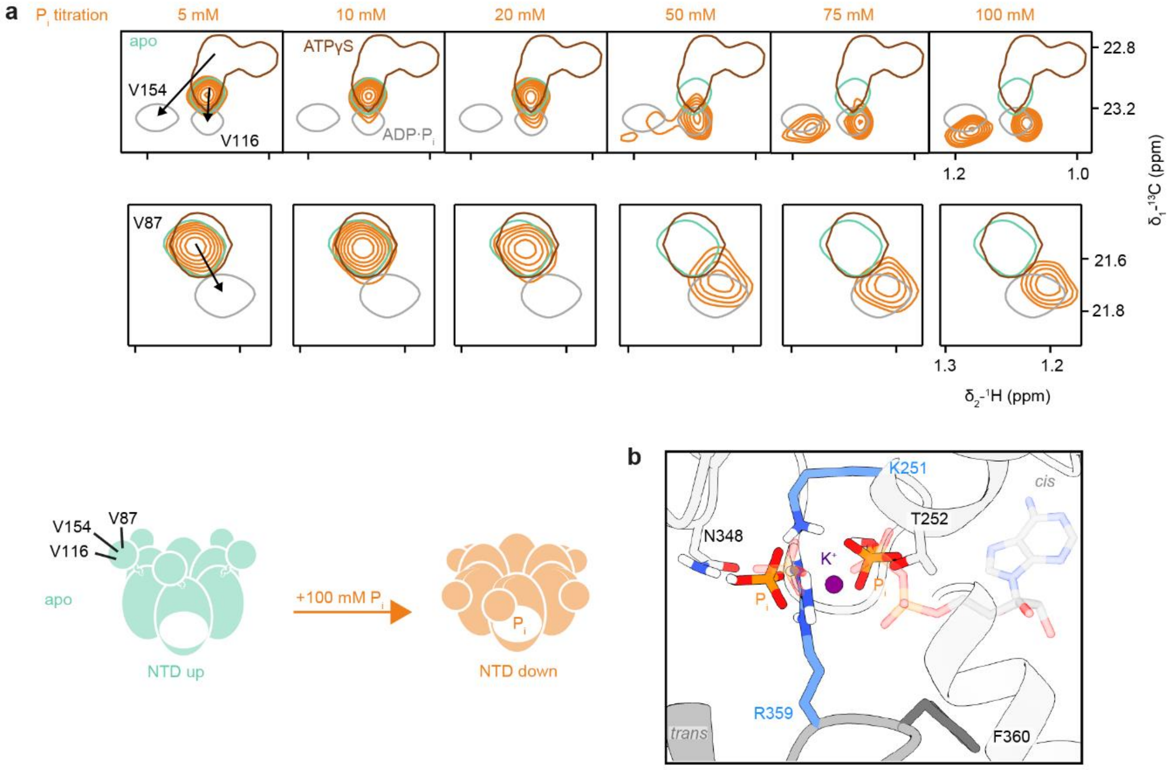


Figure 60. Exogenous addition of inorganic phosphate induces an NTD down motion in ND1L-p97. a) Inorganic phosphate was titrated to apo ND1L-wt, which causes an NTD down motion. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. Residues V154, V116 and V87 report on the NTD position. In single contours, corresponding wild type spectra are shown: ATPγS (brown), ADP·P_i (grey) and apo (green). The arrows symbolise the CSPs upon NTD down movement. **b**) In MD simulations of the apo state, two phosphate ion are mimicking the ADP·P_i state by being positioned in place of βP-ADP and the cleaved phosphate. Nucleotide positioning of the ADP·P_i state is shown in transparent. Figure adapted from ref. [52].

Surprisingly, the ADP·P_i state can also be induced from an ATPγS bound state, suggesting that a surplus of the ions competes with the nucleotide for the binding inside the pocket. As a control, we added 100 mM KCl to the ATPγS bound protein and confirmed that the increase in ionic strength is not the leading cause of the NTD down motion (Table 10). Still, an alternative explanation of this effect could be that the additional phosphate ions in solution lead to chelation of Mg^{2+}, practically removing the metal ion from the binding pocket. As discussed in Section 5.2.3, in the absence of Mg^{2+} ND1L-wt mistakenly identifies ATPγS for ADP, leading to an ADP-like state. It should be noted, however, that the spectra of p97 in the presence of ATPγS, Mg^{2+} and P_i display an ADP·P_i-like state instead of an ADP-like state, rendering this explanation less likely.

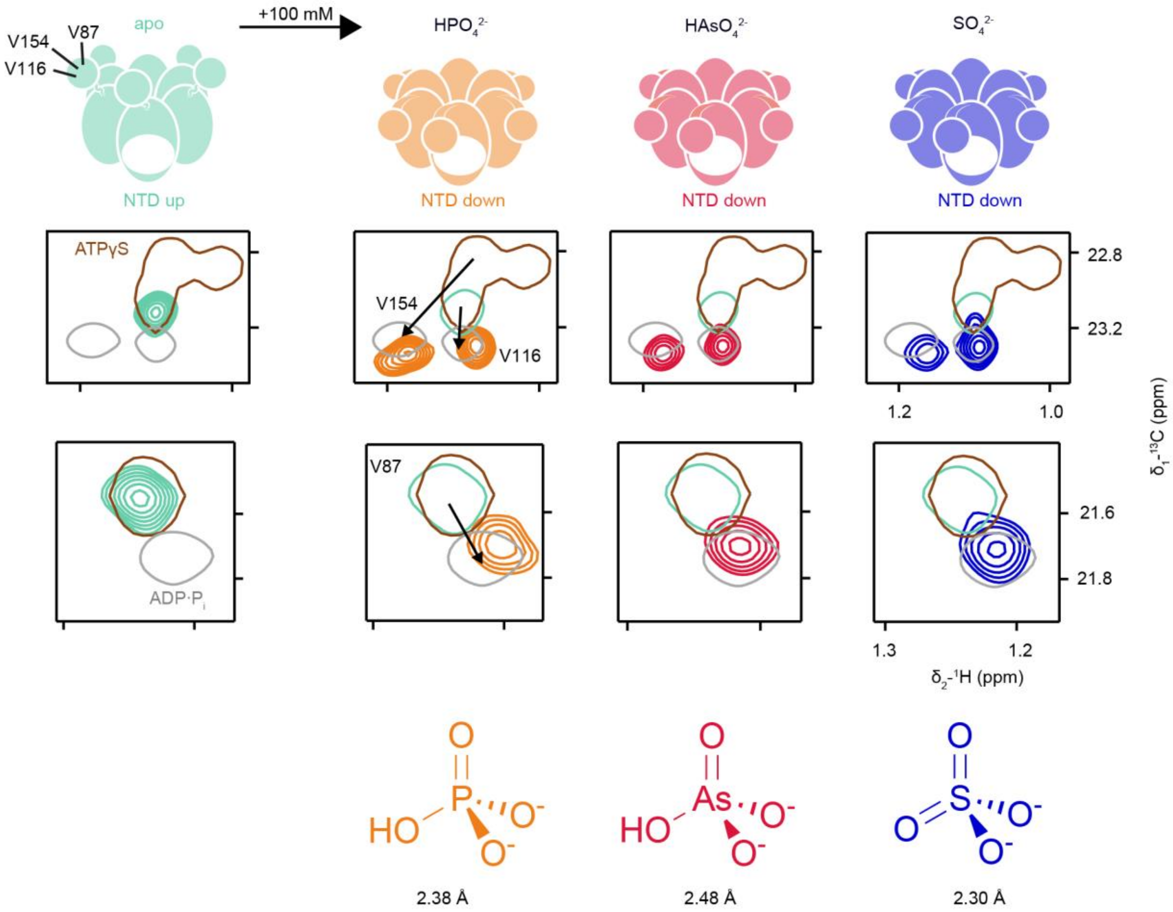


Figure 61. Phosphate mimics arsenate, and sulphate can induce the NTD down movement of apo state p97. The addition of 100 mM of inorganic phosphate leads to an NTD down movement of apo state ND1L. The same movement can be achieved by substituting phosphate with its known mimics, arsenate and sulphate. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. Residues V154, V116 and V87 report on the NTD position. In single contours, corresponding wild-type spectra are shown: ATPγS (brown), ADP·P$_i$ (grey) and apo (green). The arrows symbolise the CSPs upon NTD down movement. The thermochemical radii of the three ions are given below [133]. Figure adapted from ref. [52].

Similar to phosphate, the same NTD down movement can be induced by adding established phosphate mimics arsenate and sulphate (Table 10, Figure 61). The thermochemical radii of the three ions are similar in size (2.38 Å vs 2.48 Å vs 2.30 Å at pH 7.5 [133]), rendering a phosphate substitution possible. Both arsenate [69,74] and sulphate [72] have been previously used to study phosphate binding positions, and identical binding sites between phosphate and the mimicking ions have been found. Thus, it is not surprising that 100 mM of arsenate or sulphate ions were sufficient to fully induce the ADP·P$_i$ state.

Phosphate titration experiments have also been performed with full-length p97 starting from the apo state (Figure 62). Contrary to the behaviour of the ND1L, FL-p97 does not respond to the presence of phosphate ions by transitioning to the ADP·P_i-like state. Even 300 mM of inorganic phosphate is not enough to induce a complete NTD down motion, and the overall protein conformation remains in the apo state. Although it is known that the position of the N-terminal domain is regulated by nucleotide binding in D1 [32], these results imply that the D2 domain also contributes to the control of NTD.

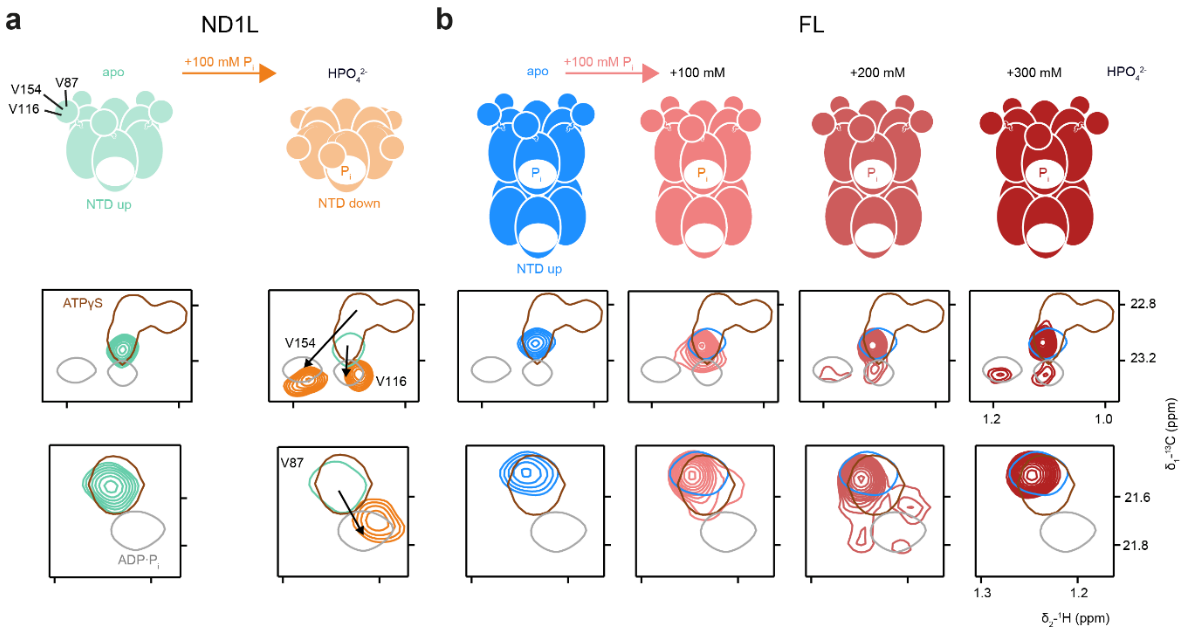


Figure 62. Addition of phosphate to full-length p97 apo state does not induce a complete NTD down motion. a) Addition of 100 mM of inorganic phosphate leads to an NTD down movement of apo state ND1L-wt. **b**) The same movement cannot be achieved by with FL-wt. The NTD position responds weakly to the presence of phosphate and even 300 mM P_i are insufficient to induce the ADP·P_i-like state. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. Residues V154, V116 and V87 report on the NTD position. In single contours, corresponding wild type spectra are shown: ATPγS (brown), ADP·P_i (grey) and apo (green or blue). The arrows symbolise the CSPs upon NTD down movement.

In order to confirm that the NTD down movement is correlated to phosphate binding inside the nucleotide binding pocket and not, e.g. to the NTD itself, the effect of phosphate titration was tested on all ND1L mutants. The results of the experiments are summarised in Table 11. Mutants which show deficiencies in hexamer formation in the apo state also do not show complete NTD down movements in the presence of P_i. This correlation confirms that a proper hexamer (and thus nucleotide binding pocket) formation is a prerequisite to the binding of phosphate, linking phosphate binding directly to the conformational state of the binding pocket.

Two exceptions are noticeable: P246T and K251A mutants do not react or only respond partially to the titration of Pi, although both mutants form normal hexamers in the apo state. The response of K251A can be explained by its general inability to stably bind nucleotides and the involvement of K251 in phosphate binding, as observed in the MD simulation (Figure 60b). The behaviour of P246T is surprising since this mutant does not show any other obvious defects. It is unclear whether this phosphate binding defect is connected to the mutant's inability to bind phosphate ions in the first place or to the lack of signal communication towards the NTD. Either way, it appears that the secondary structure of the P-loop is directly involved in nucleotide binding and signal transduction. This also confirms that the NTD down movement is indeed induced by phosphate binding inside the binding pocket.

Table 11. Mutants without oligomerisation deficiencies can react to phosphate titration. ND1L-mutants were tested for their ability to react to phosphate titration (100 mM inorganic phosphate, 4 mM $MgCl_2$). Mutants which do not show any deficiencies in hexamer formation in the apo state form ADP·P_i-like states upon phosphate addition. Notation: + (yes), ~ (partial), + (no).

	Wild type	P246T	P247A	P247K	K251A	D304N	E305Q	N384Q	R349A	P355A	R359A	R359K	F360A	F360P	R362A
Hexameric in apo state?	+	+	+	-	+	+	~	+	+	+	-	+	+	+	-
NTD down movement upon P_i addition?	+	-	+	~	~	+	~	+	+	+	~	+	+	+	-

5.5 Crosslinking study

5.5.1 Motivation

During ATP turnover, F360 has been shown to act as a switch and sensor [52,63]. In ADP·P_i state B, the F360 rotamer comes into contact with the α7 helix ($\alpha_{407\text{–}423}$), transducing nucleotide state information towards the NTD (Figure 63a). By affixing the rotamer to the helix, potentially the protein could be locked in the ADP·P_i state, preventing further ATP processing. This approach can be achieved by covalently bonding the helix and the *trans*-acting loop via disulphide bonds. In previous studies, this approach has been used on p97 to alter NTD positioning: for this, the N-terminal domain with its disease-associated R155C mutation was disulphide bonded to N387C of the D1 domain. This effectively locked the NTD in the down conformation [66,134] and reduced the ATPase rate significantly [44].

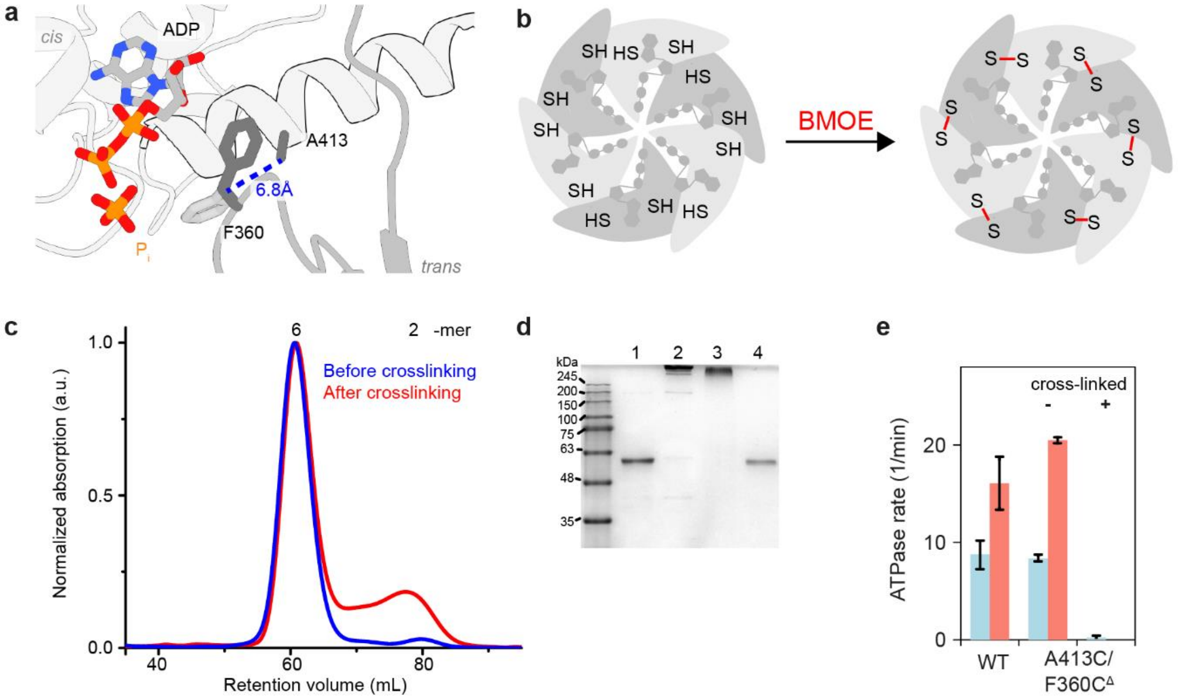


Figure 63. Crosslinking of ND1L-ΔCys-F360C-A413C. a) During ATP turnover, one of F360 rotamers comes into contact with α7 helix. The helix ($\alpha_{407–423}$) is highlighted in bold. In the ADP·P_i state, the distance between the Cβ atoms of F360 and A413 is 6.8 Å. **b-e**) Biophysical characterisation of BMOE-crosslinked ND1L-ΔCys-F360C-A413C. **b**) Schematic depiction of neighbouring protomers being cysteine-crosslinked using BMOE. **c**) Free (blue) and crosslinked (red) protein retain hexameric SEC elution profiles. **d**) Non-reducing SDS-PAGE confirms crosslinkage. Samples: 1 – SEC peak corresponding to hexamers before the reaction with BMOE; 2 – BMOE/p97 reaction mixture; 3 – SEC peak corresponding to hexamers after the reaction with BMOE; 4: ND1L-wt reference. Similar results were achieved in other biologically and/or technically independent experiments. **e**) ATPase assay of free and crosslinked samples. BMOE crosslinking renders the protein ATPase inactive. $^{\Delta}$ designates the cysteine-free p97 variant. Figure adapted from ref. [52].

Using the same idea, we envisioned cloning double cysteine mutants, with F360C on one side and a corresponding cysteine on the side of the helix. For this, three candidates (D410, A412 and A413) were selected due to their orientation towards the F360 rotamer. One also needs to consider the average distances between the residues, shown as distances between the corresponding Cα atoms in different nucleotide states (Table 12). On average, the distances between the Cα atoms in all three nucleotide-bound states are 8-10 Å, which is significantly longer the average distance of 4-5 Å and a maximum distance of 7 Å between Cα atoms in native disulphide bonds [135]. In contrast, the distance between the Cα atoms of R155 and N387 in the NTD down conformation is only 5.3 Å (PDB ID: 5ftk), making disulphide bond linkage more likely. Still, the arginine finger loop containing F360 is clearly dynamic, and the distance between the loop and helix might even further change over the course of the ATP hydrolysis cycle, rendering the cysteine-cysteine bonding possible.

Table 12. Cα-Cα distances between F360 and potential crosslinking candidates. Cα-Cα distances of F360 to the potential crosslinking candidates D410, A412 and A413 in different nucleotide states are shown. Corresponding PDB IDs are given in parentheses.

Distance from F360 to	D410	A412	A413
ATPγS (5ftn)	8.7 Å	9.8 Å	8.9 Å
ADP·P_i (8ooi)	8.3 Å	9.9 Å	8.6 Å
ADP (5ftk)	8.1 Å	9.9 Å	8.6 Å

5.5.2 Results

Initial trials were conducted with double mutants (ND1L-F360C-[D/A]41[0/2/3]C) in ADP bound and apo states via oxidation on air, following the protocols from ref. [66]. Disulphide bond formation was tested via reducing and non-reducing SDS-PAGE, and hexamer formation was controlled via SEC. Although partial inter-protomer linkage was visible on gels, most of the protomers remained unlinked (data not shown). We surmised that the cysteine C415 located in close proximity to residues A412 and A413 could interfere with disulphide formation (average Cα-Cα distance between C415 and A412/3 is approx. 5.1-5.6 Å). Thus, we cloned triple mutants ND1L-F360C-[D/A]41[0/2/3]C-C415A, which, however, suffered from the same issues: the ratio of crosslinked-to-monomeric protein remained unfavourable. The oxidation reagent dichloro-Cu-phenanthroline was employed to force disulphide bridge formation. Unfortunately, this method has proven to be too aggressive for these mutants since sample oxidation led mostly to aggregation. It is likely that the long distance between the cysteines led to a conformational strain upon disulphide formation, potentially pulling two neighbouring protomers together and increasing the gap between the other monomers. This would explain why the oxidation on air only ever led to incomplete oxidation. In contrast, forced disulphide bond formation using dichloro-Cu-phenanthroline possibly led to further stress and unfolding or, when formed between hexamers, to aggregation.

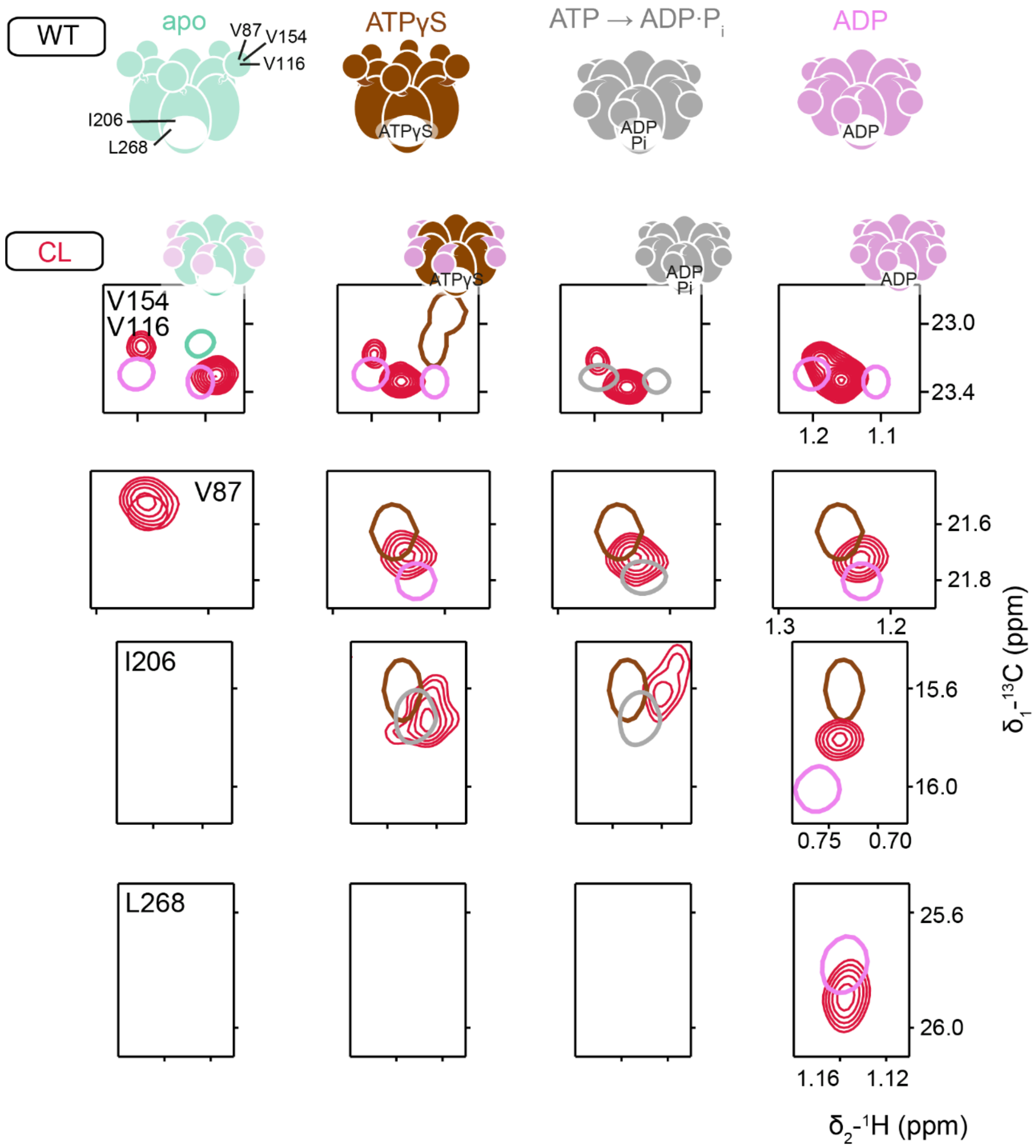


Figure 64. Cysteine-free crosslinked (CL) ND1L mutant fails to couple NTD positioning to the nucleotide state. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. Residues V154, V116 and V87 report on the NTD position, while residues I206 and L268 inform on the type of nucleotide bound in the binding pocket. In single contours, reference wild type spectra are shown. Irrespective of the nucleotide type, the NTD is located in the down position. Mutation of cysteines alone renders the mutant incapable of coupling the NTD position to the nucleotide state.

To combat these issues, we decided to employ a linker agent, *bis*-maleimidoethane (BMOE), with a spacer arm length of 8 Å to cover the long distance between the cysteines. The main disadvantage of this approach is that the linkage is irreversible, rendering control measurements more complicated. Moreover, due to the increased linker length, undesirable crosslinking reactions with other cysteines (intra- and interprotomer) are more likely to occur. Thus, a cysteine-free ND1L mutant, ND1L-ΔCys, was used as a background for the crosslinking mutations (a complete list of mutations is given in Methods). Out of the three potential candidates, ND1L-ΔCys-F360C-A413C showed the best performance. Successful crosslinking was confirmed by reducing SDS-PAGE, and both the free and crosslinked protein eluted as hexamers on SEC (Figure 63c,d). The biological activity of the free mutant was confirmed in an ATPase assay, where the protein demonstrated ATP hydrolysis rates comparable to the wild-type ND1L reference (Figure 63e). In contrast, upon crosslinking, the sample became ATPase inactive. These promising results indicated that the crosslinking indeed could have led to the trapping of the ADP·P_i state. In order to confirm this, we planned to perform solution-state NMR measurements on the mutant. However, it turned out that the cysteine-free mutant ND1L-ΔCys already had a major conformational defect, as it failed to couple the NTD position to the nucleotide state (data not shown). This deficiency was also observed in the crosslinked mutant (Figure 64), rendering further analysis impossible.

In conclusion, we successfully crosslinked neighbouring protomers by covalently binding the arginine finger loop to the α7 helix, which come into contact naturally during ATP turnover. The linkage prevented proper ATP processing, although it is unclear which step is actually affected. In NMR experiments, the crosslinked mutant was still able to react to nucleotide binding (*c.f.* apo state spectra vs. nucleotide bound spectra, Figure 64), showing that the crosslinking reagent did not sterically prevent nucleotide binding. Still, it remains unclear whether the ATP hydrolysis or product release is affected by crosslinking.

5.6 Intersubunit communication

5.6.1 Introduction

Undoubtedly, p97 does not work in its monomeric form, automatically requiring cooperation between the involved protomers. The location of the nucleotide binding pocket at the interface of two neighbouring protomers implies coordination and signal transduction. Several key residues of this process have been discussed in detail in Section 5.2 (arginine fingers, F360, ISS). Aside from these known features, no working hypothesis of ATP hydrolysis has been proven for p97, although counter-clockwise hydrolysis around the ring is assumed in reference to the working model of other AAA+ ATPases [22]. Several observations support this hypothesis: In solution-state NMR measurements under ATP turnover, low concentrations of ADP in solution (amounting to a single bound ADP nucleotide per hexamer) were sufficient to stall the ATPase activity of the full ND1L ring [55]. In MD simulations of the ND1L-wt construct, the occurrence of hydrolysis active conformations was correlated to the distance from the ADP·P_i bound binding pocket in a counter-clockwise manner [52]. These findings suggest that the nucleotide identity in one binding pocket is communicated around the whole ring, requiring established communication pathways. In the following section, several communication features and pathways are investigated and discussed.

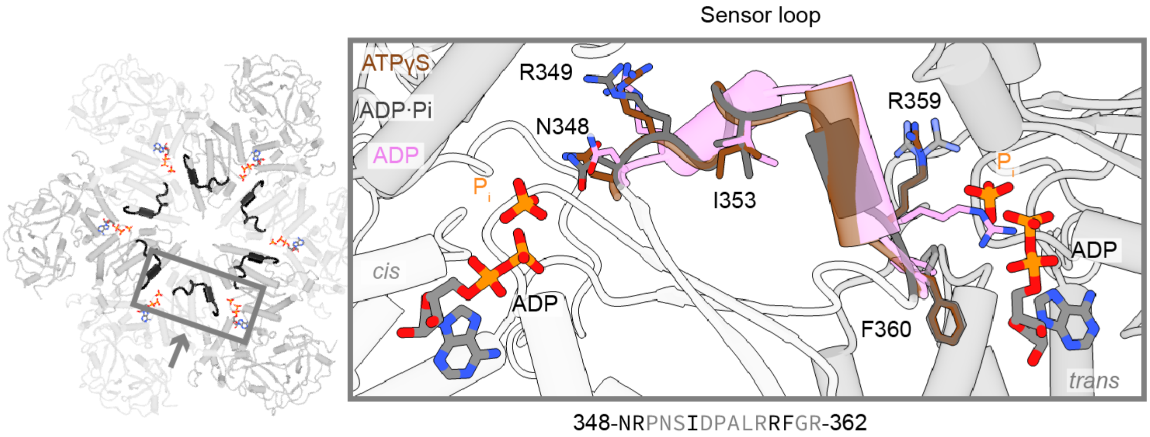


Figure 65. The sensor loop directly connects one nucleotide pocket to another. Two neighbouring nucleotide binding pockets are directly connected via the sensor loop. It starts with the sensor-1 motif (N348) in one binding pocket and ends with the arginine fingers (R359, R362) in the neighbouring one. One the left, in the top view, the loop is seen connecting two nucleotides. On the right, the secondary structures of the sensor loop in three different nucleotide states (ATPγS (5ftn): brown; ADP·P_i (8ooi): grey; ADP (5ftk): pink) are shown. The ATPγS and ADP·P_i structures are overlapping, while the loop adopts a different conformation in the ADP state. This figure was previously published in ref. [52].

5.6.2 Sensor loop dynamics

After ATP hydrolysis, the ADP·P$_i$ state is formed. The cryo-EM structure of wild-type full-length p97 under ATP turnover, with the ADP·P$_i$ state in D1, is highly similar to the ADP structure of the same construct (PDB ID: 5ftk), with the RMSD being 2.2 Å, while the RMSD between the ADP·P$_i$ state and the ATPγS state (PDB ID: 5ftn) is 12.1 Å due to the large difference in NTD positioning. However, even disregarding the flexible N-terminal domain, the RMSD between the two structures is still larger (3.0 Å) than between the ADP and ADP·P$_i$ states.

One of the few key structural differences is found in the sequence of the Secondary region of homology. The secondary structure element of the ADP·P$_i$ state starting with the sensor-1 motif (N348) in one nucleotide binding pocket and directly leading to the neighbouring one (up to arginine finger R362) shares more similarities with the ATPγS bound conformation instead (Figure 65). This direct connectivity led us to term this secondary structure element as “sensor loop”. Upon visual inspection, it can be seen that the loop shares its secondary structure between the ATPγS and ADP·P$_i$ states but adopts a different one in the ADP state. The visual assessment is supported by quantified analysis performed by the STRIDE algorithm and by Ramachandran plot analysis (Figure 66). According to STRIDE [136], two structural elements are identified in the loop – a turn motif and a 3_{10}-helix. The turn, however, is only present in the ATPγS and ADP·P$_i$ bound states but changes to a second 3_{10}-helix in the ADP state (Figure 66b). A similar comparison is drawn from the Ramachandran plot analysis: backbone dihedral angles of residues which are part of the first half of the loop display minor changes between the ATPγS and ADP·P$_i$ states, but change more dramatically from the ADP·P$_i$ to the ADP states (Figure 66c). We surmise that the sensor loop is one of last secondary elements to restructure from its ATPγS-like to the ADP-like conformation. By virtue of being associated with the transition between the ADP·P$_i$ and ADP states, its conformational change must be coupled to the release of phosphate.

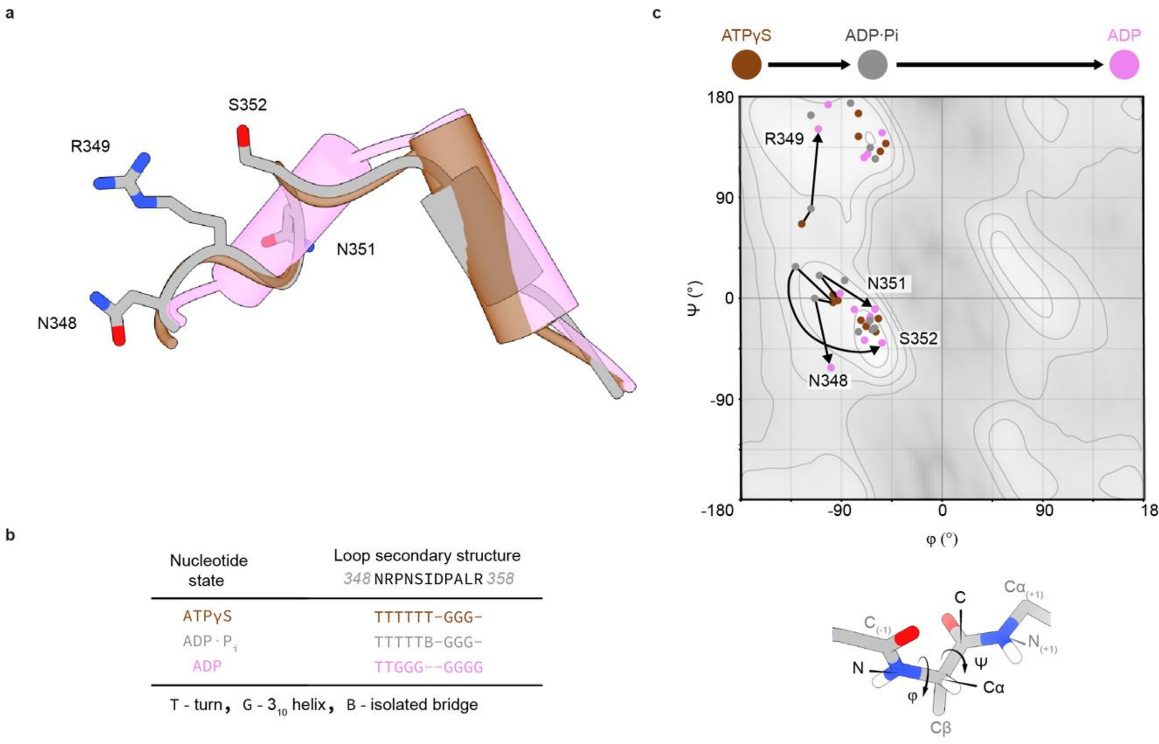


Figure 66. Secondary structure analysis of the sensor loop. a) Overlay of the sensor loop in the ATPγS (brown, PDB: 5ftn), ADP·P_i (grey, PDB: 8ooi) or ADP (pink; PDB: 5ftn) bound states. Residues analysed in **c**) are shown as sticks. **b**) STRIDE algorithm [136] identifies similarities in secondary structure between the three states. In the ADP state, the first half of the loop adopts a 3_{10}-helical structure instead of a turn. **c**) Backbone conformation of the sensor loop as identified by Ramachandran plot analysis. The conformational changes are more pronounced between the ADP·P_i and ADP states than between the ATPγS and ADP·P_i states. The coloured dots represent the individual amino acids in given states. The arrows form connections between subsequent states of hydrolysis of the same residue. Colours are identical to the ones in **a**). The defining dihedral angles are explained on the stick model at the bottom. The Ramachandran plot analysis was performed by S. R. Rout. This figure was previously published in ref. [52].

To test this hypothesis, we cloned several ND1L mutants with mutations specifically located in the first half of the sensor loop. ATPase activity measurements showed that the mutations of the sensor loop only partially reduce the ATPase rate as long as the residues are not directly involved in the ATP processing (Figure 67). N348Q is deficient in ATP hydrolysis and thus inactive, while all other mutants show diminished ATPase activity. Moreover, R349A, P350A and N351A all display similar ATP hydrolysis rates, while P355A is the only mutant whose ATPase rate is comparable to the wild-type rate. Knowing of the involvement of R349 in P_i dissociation (Section 5.3), we decided to use solution-state NMR to investigate both the R349A mutant (as a representative of the slow mutants in this group) and the P355A mutant (as the only fast mutant). However, neither mutant displayed any substantial deficiencies in the NMR measurements, and the spectra overall resembled the wild-type references (data not shown).

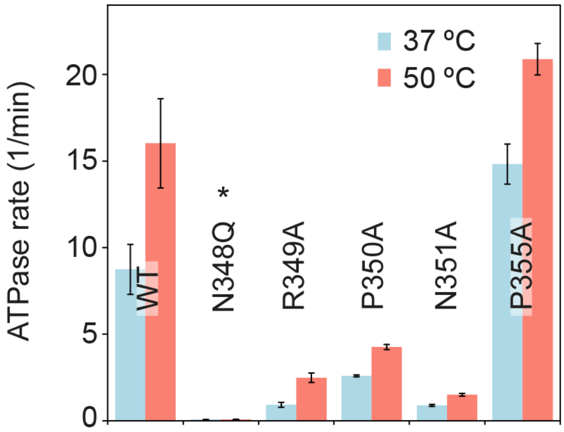


Figure 67. ATPase rates of sensor loop mutants. Residues which are part of the first half of the sensor loop have been mutated. N348Q is inactive due to its inability to hydrolyse ATP. * denotes ATPase inactive mutants.

Thus, we instead decided to focus on loop dynamics by solution-state NMR. Two methyl groups are part of the loop – I353 and L357. The peak position of L357 confirms the similarity between the ATPγS and ADP·P$_i$ states [55]; however, the overall CSP induced by ADP binding is rather small, rendering this methyl probe less useful. In fact, this is not surprising, considering that L357 is located in the second half of the sensor loop and is part of a 3_{10}-helical motif in all three nucleotide-bound states (Figure 66b). In contrast, I353 is part of the first half of the loop and reacts strongly to nucleotide changes [55]. In the ^{13}C dimension alone, the peak shifts by 1.1 ppm between the ATPγS and ADP bound states, a feature possible for isoleucine side chains upon transition from *trans*- to *gauche*-states [137]. Thus, this residue can be used as an effective probe of loop conformational changes.

Loop mobility is clearly visible in the wild-type measurement under ATP turnover. The I353 peak is strongly broadened due to increased dynamics and is positioned between the two extremes. As a follow-up, we compared I353 peak position to the ATPase activity for multiple ND1L mutants (Figure 68a). The results show a clear correlation between the ATPase rate and the completeness of loop rearrangement from the ATPγS state to ADP state during ATP turnover: for inactive or slow mutants, I353 displays a peak position very similar to the ATPγS state, while hyperactive F360[A/P] mutants show complete loop conversion to the ADP state. These findings prove that the sensor loop secondary structure is coupled to the rate of P$_i$ release.

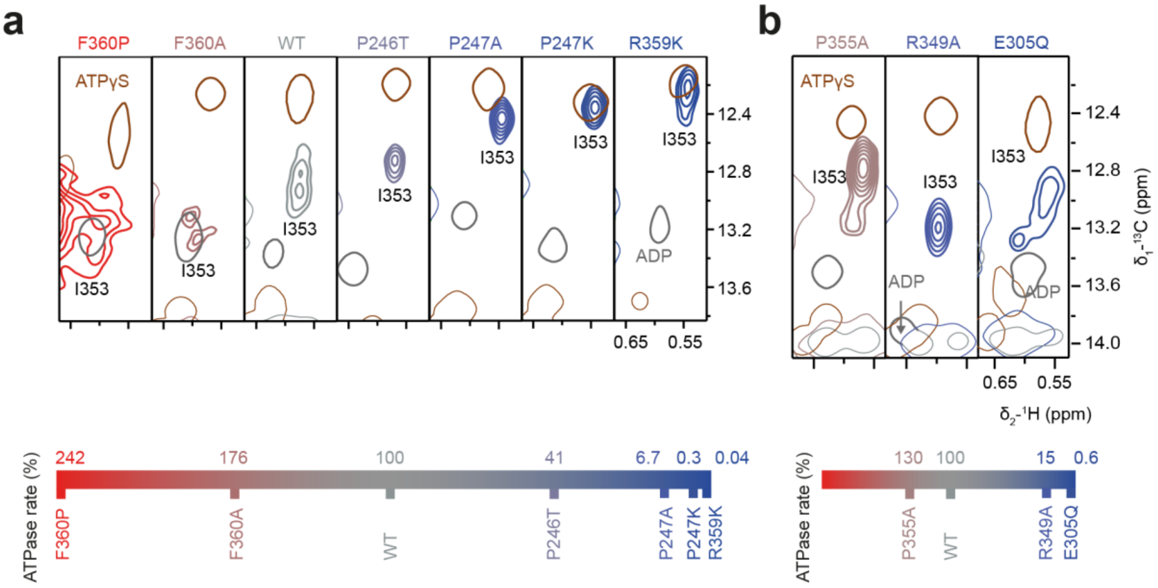


Figure 68. Sensor loop dynamics are related to ATPase activity. $C\delta_1$-methyl group of I353 is used as a probe for structural conformation of the sensor loop. In the ADP·P_i state of wild type p97, the peak is positioned in the middle between the two reference peaks (ATPγS: brown, ADP: grey) and strongly broadened due to increased mobility. **a**) Mutants capable of ATP hydrolysis display a correlation between ATPase activity and peak position, relating loop secondary structure to the rate of P_i release. Bottom: comparison of mutant relative ATPase rates. **b**) I353 peak positions of non-conforming mutants are shown. R349 and P355 are part of the sensor loop. All three mutants form ADP·P_i states under ATP turnover. Panel **a**) was previously published in ref. [52].

As a caveat, it should be mentioned that all displayed mutants fully form the ADP·P_i state under ATP turnover. I353 peak positioning of mutants deficient in ATP hydrolysis (e.g. N348Q, R359A, R362A) can differ from the established correlation. Three other exemptions are shown in Figure 68b: I353 peaks of E305Q, R349A and P355A do not adhere to the shown rules, although all three mutants form ADP·P_i states during ATP turnover. Since R349 and P355 are part of the sensor loop, mutations of these residues may affect its secondary structure. For instance, the reference peak of I353 in the ADP state of ND1L-R349A is strongly shifted downfield. In contrast, the strong peak broadening and the wt-like positioning exhibited by I353 of the E305Q mutant is surprising, considering the low ATPase rate of this Walker B mutant. This observation potentially hints at the more distinct role of E305, which is discussed in Section 5.6.3.

Increased sensor loop dynamics have also been observed in MD simulations. The sensor loop adjacent to the nucleotide binding pocket with the ADP·P$_i$ state showed elevated mobility, while ATP-adjacent loops were less mobile [52]. This suggests subsequent information transfer from the ADP·P$_i$ nucleotide binding pocket through the sensor loop towards neighbouring ATP binding sites.

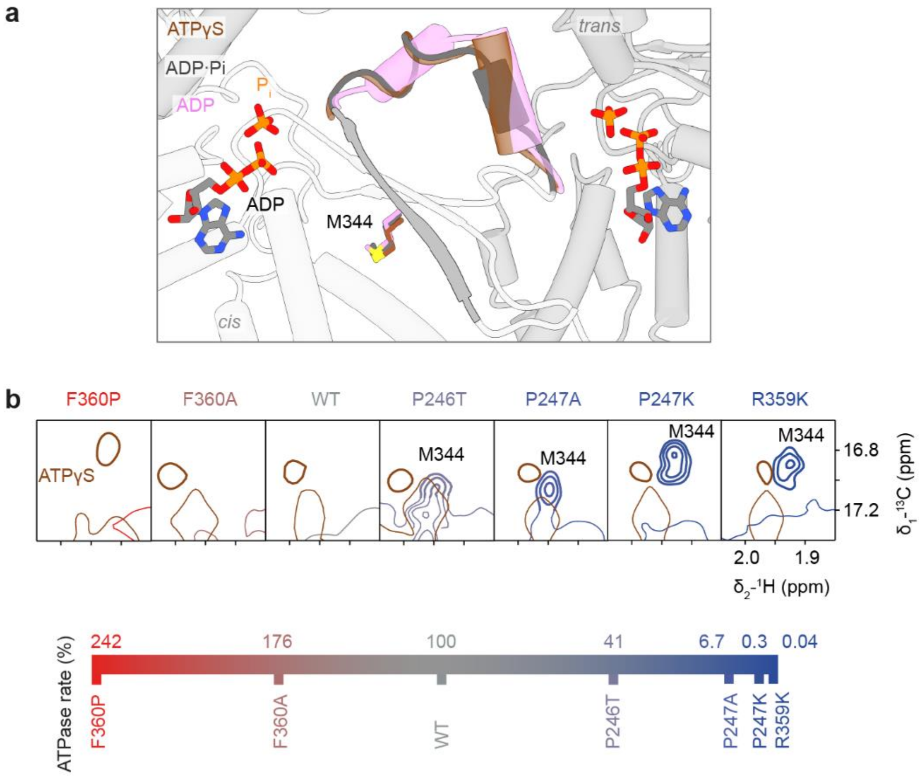


Figure 69. M344 acts as a sensor for the ATPase activity of ND1L mutants. a) Zoom in on the sensor loop from the same view as Figure 65. M344 is part of the β4-sheet ($\beta_{341\text{-}347}$) which leads up directly to the sensor loop. **b**) Cε_1-methyl group of M344 has two distinct peak positions in the static nucleotide states. Under ATP turnover, the peak disappears. Its peak position and intensity is coupled to the ATPase rate. For faster mutants under ATP turnover, the M344 moves away from its ATPγS reference and broadens beyond recognition.

Interestingly, a similar correlation between the ATPase rate and peak position has been noticed for M344. In static nucleotide states, peak positions of the Cε_1-methyl group of M344 are well defined, but the peak disappears in ND1L-wt spectra under ATP turnover (Figure 69). However, the M344 peak re-emerges in the ADP·P$_i$ states of slow ND1L mutants, correlating the side chain conformation of M344 with the phosphate release. It should be noted that the methionine itself is neither part of the sensor loop nor is its side chain in contact with the loop at any given protein state. However, it is part of the β4 sheet, which directly ends in residue N348 and starts the sensor loop (Figure 69a).

5.6.3 The functional relevance of E305

The glutamic acid residue 305 of p97 deserves a separate discussion. Being a conserved residue of the Walker B motif (Figure 47), its role and the effects of its mutation have been well established: Walker B residue E305 is considered to be involved in hydrolysis by activating a water molecule for a nucleophilic attack on the γ-phosphate of ATP [50,51]. Conversely, its mutation is associated with deficiencies in hydrolysis, but specifically not ATP binding [50]. Based on ATPase activity measurements of different Walker A and Walker B mutants and an ND1 construct, Song *et al.* [56] determined that D2 is the main ATPase domain of p97. This notion was corrected when the following studies showed that the D1 domain was indeed able to hydrolyse ATP as long as the D1-D2 linker was attached [47]. The current understanding is that the D1 domain significantly (~30%) contributes to the total ATPase rate [39].

Partially, the misassessment of the ATPase contribution of the D1 domain can be attributed to the setup of some ATPase measurements in which the produced ADP is not immediately regenerated, leading to the accumulation of ADP and thus stalling the protein [39,55,56]. Evidently, ATPase rates of the same mutants may differ between studies. Small variations are to be expected, but the minor reduction of ATPase activity (~75% of wt) in the FL-K251T mutant observed by Song *et al.* strongly contradicts the results of other studies (21% [41], 27% [47], <1% [126], 15% [63]). The main argument for the weak contribution of the D1 domain to the total ATPase rate lies in the comparably stronger effect on the ATPase rate upon mutating Walker A or Walker B residues in the D2 domain [41,47,56].

ATPase measurements show that in the context of full-length p97, E305Q mutations lead to a ~30% increased ATPase rate, but the corresponding D2 mutation E578Q reduces it by ~70% (Figure 70). These results lead to the conclusion that only the D2 domain plays an active role in ATP hydrolysis, while the hydrolysis deficient mutation E305Q is largely irrelevant for the protein. However, in the context of the ND1L construct, the E305Q mutation indeed strongly reduces the ATPase rate. This reduction is significant since the ATPase activity of ND1L-wt is actually comparable to the one of full-length wild-type p97. Additionally, the double Walker B mutant FL-E305Q-E578Q is largely ATPase inactive. If the D1 domain truly did not contribute to ATP hydrolysis, then the introduction of the D1-E305Q mutation into the FL-E578Q mutant should not affect its ATPase activity. Instead, solid-state ^{31}P NMR measurements of the double mutant find a bound ADP·P$_i$ state in the D1, confirming ATP hydrolysis in that domain.

These conflicting findings immediately question the role of the E305 residue and its involvement in the D1-D2 communication. Another clue is found in comparisons of the ATPase activities of p97 homologues (Table 13). Walker B mutations of the D1 domains in the homologue full-length proteins lead to strong reductions in the overall ATPase rates, but this is not observed for the human p97. Considering the analogous roles these proteins play in their respective organisms, it seems likely that the homologues would process substrates in a similar fashion, and the D1 domains would fulfil identical functions. For that reason, the differences observed between the ATPase activities of the D1-Walker B mutants must be more subtle.

Table 13. Compilation of ATPase rates of p97 homologues with a Walker B mutation in the D1 domain. In p97, the Walker B mutation E305Q (in the D1 domain) does not significantly reduce the ATPase activity or even increases it. In contrast, in other homologues, the total ATPase rates are significantly reduced in the mutated constructs. The relative ATPase rates are given in comparison to full-length wild-type proteins. n.d.: not detected

Organism	Protein	Mutation	Rel. ATPase activity [% of FL-wt]	Reference
H. Sapiens	VCP/p97	E305Q	125%	This work
H. Sapiens	VCP/p97	E305Q	133%	Ref. [41]
H. Sapiens	VCP/p97	E305Q	95%	Ref. [47]
H. Sapiens	VCP/p97	E305Q	~90-95%	Ref. [45]
H. Sapiens	VCP/p97	E305Q	~80%	Ref. [56]
H. Sapiens	VCP/p97	E305Q	~225%	Ref. [124]
H. Sapiens	VCP/p97	E305Q	~150%	Ref. [63]
T. acidophilum	VAT	E291A	~15%	Ref. [60]
S. cerevisiae	Cdc48p	E315Q	10%	Ref. [125]
S. cerevisiae	Cdc48p	E315D	23%	Ref. [125]
S. cerevisiae	Cdc48p	E315A	n.d.	Ref. [138]
S. cerevisiae	Cdc48p	E315Q	~25%	Ref. [138]
C. elegans	CDC-48.1	E311Q	30%	Ref. [125]
C. elegans	CDC-48.1	E311N	41%	Ref. [125]
C. elegans	CDC-48.1	E311A	30%	Ref. [125]
C. elegans	CDC-48.1	E311D	33%	Ref. [125]
A. thaliana	Atcdc48	E308Q	70%	Ref. [139]

Still, the conventional interpretation of the E305Q mutation effect needs to be questioned. NMR measurements of the ND1L-wt and E305Q constructs under ATP turnover detect a very similar ADP·P_i state (Table 9) [55]. This implies that while the ATP hydrolysis rate might be slowed down due to the mutation, the actual rate-determining step of the hydrolysis process remains the phosphate release. Moreover, the spectrum indicates a uniform protein conformation, implying that all six subunits are trapped in the ADP·P_i state. Thus, the effect of the mutation does not lead to unregulated, independent ATP processing.

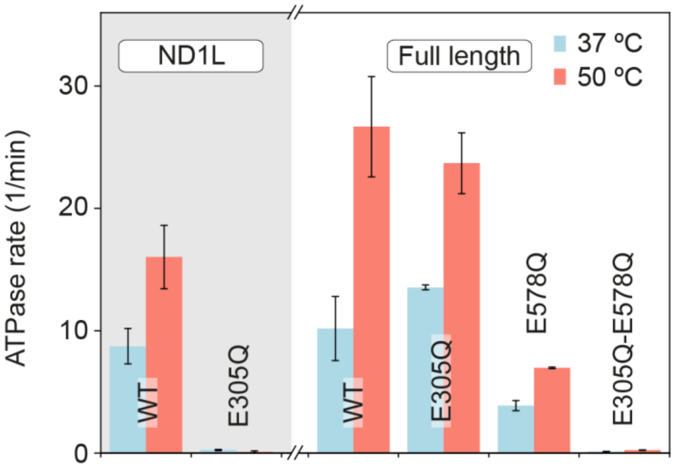


Figure 70. ATPase measurements of Walker B mutants. Mutation of the D1-Walker B residue E305 in the full length protein leads to an increased ATPase rate, while the mutation of the corresponding residue in D2 (E578Q) strongly reduces the activity. In the context of the ND1L construct, E305Q strongly reduces the ATPase rate.

Strangely, in MD simulations, the residue E305 has not been found to play any active role in the stabilisation ADP·P_i state, but only in the hydrolysis and P_i release processes [52]. We hypothesise that the glutamate fulfils a secondary function, possibly in interprotomer coordination. This is supported by the reported competing interaction of R362 with E305 and the ISS residue D333 (Figure 71) [41]. In the Walker B mutant, D333 and R362 would only interact with each other, failing their *trans*-sensing functionality and disrupting this communication network. Other disruptions could be possible, e.g. Q305 might interact with the glutamate switch residue N270 [57], a feature not observed for the wild-type protein in any nucleotide states despite their close proximity.

These considerations can further explain why phosphate release remains the rate-determining step of the ATPase reaction of the ND1L-E305Q mutant. With its negative charges, E305 contributes to the destabilisation of the ADP·P_i state through electrostatic repulsion to the cleaved phosphate [52]. Contrasting, R362 is involved in the coordination of the P_i ion [52]. Thus, it is likely that the formation of the E305-R362 salt bridge prevents ADP·P_i state stabilisation. In the ND1L-E305Q mutant, however, these destabilizing interactions are absent and the ATPase activity is significantly reduced.

Another observation speaks for disrupted interprotomer coordination. In solution-state NMR measurements of ND1L-E305Q under ATP turnover show a similar sensor loop conversion state as the one in ND1L-wt (Figure 68b). As shown previously, for most other ADP·P_i state forming mutants, the conformational state of the sensor loop is coupled to the ATPase rate. In this case, however, the slow E305Q mutant indicates a loop secondary structure similar to that of the fast wild-type protein. The origin of this behaviour is unknown and could possibly result from a new interaction of Q305 or from the enhanced interaction of the sensor loop residue R362 with D333.

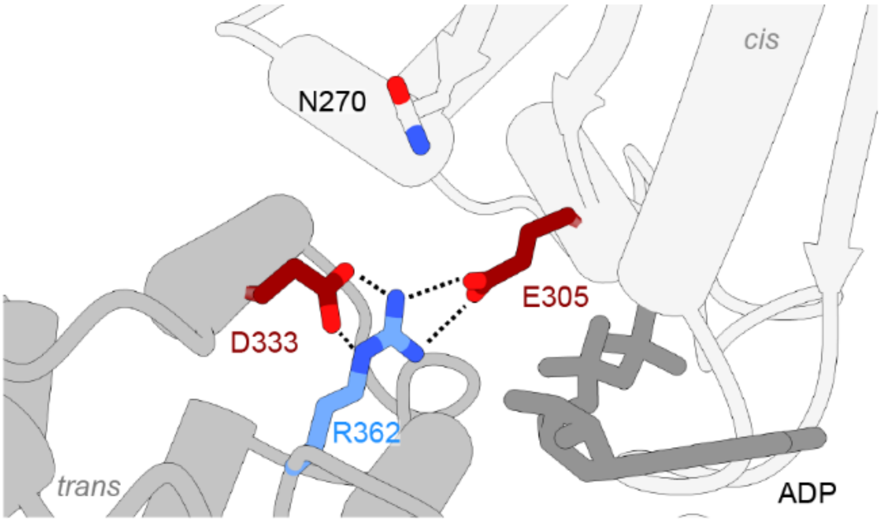


Figure 71. Interaction network of R362. Arginine finger R362 is known to form salt bridges with the Walker B residue E305 and the ISS residue D333 [41]. In the Walker B mutant E305Q this network is disrupted, leading to a strengthened R362-D333 bond. In this mutant, Q305 might form new interactions such as with the glutamate switch residue N270.

In the cryo-EM structure of full-length p97 under ATP turnover, the ADP·P_i state was only found in the D1 domain while ATP was detected in the D2 domain, suggesting a D1→D2 hydrolysis order [52]. The lack of ATP hydrolysis in D2 also leads to the conclusion that prior to D2-ATP hydrolysis, phosphate must be released from the D1 binding pocket. Considering that the loop conversion supposedly signals that release, a mimicking signal through enhanced loop dynamics from the ADP·P_i trapped E305Q mutant might stimulate hydrolysis in D2, explaining the fast ATPase rates of FL-E305Q (Figure 73).

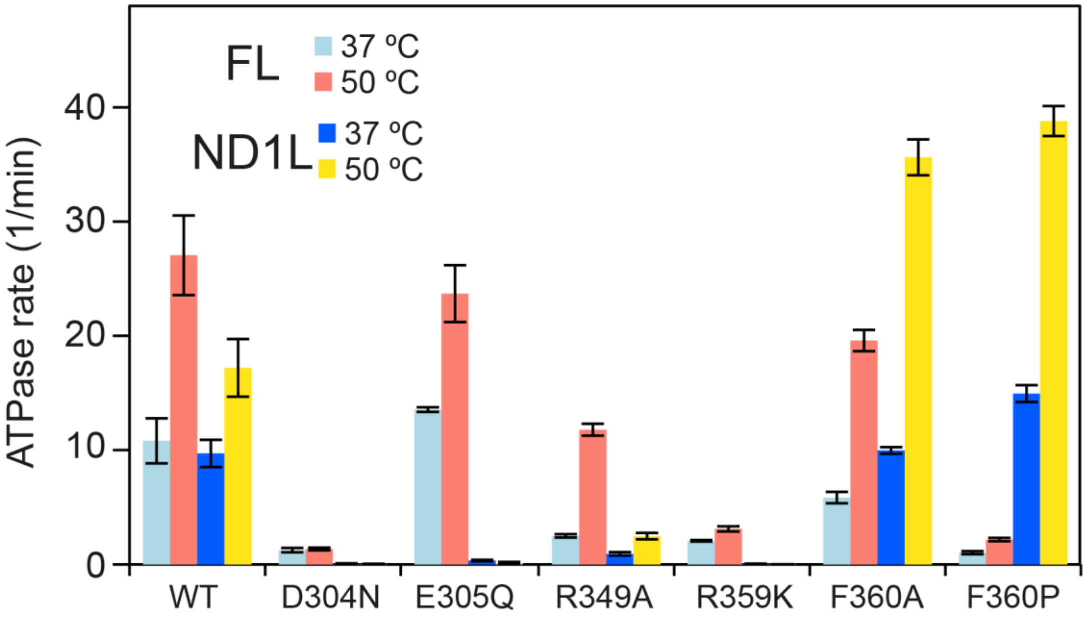


Figure 72. ATPase rates of full length p97 mutants. Under ATP turnover, all ND1L mutants (the state assignment of D304N is ambiguous) form ADP·P_i states. In the full length context only the E305Q mutant exhibits an enhanced ATPase rate, while the other mutants show reduced ATPase activities.

This hypothesis begs the question of whether other D1-ADP·P_i trapping mutations could also hyperstimulate D2 ATPase activity. In order to answer this question, we measured the ATPase rate of full-length mutants that form ADP·P_i states as ND1L constructs (Figure 72). However, in contrast to FL-E305Q, the ATPase rates of both FL-R359K and FL-R349A are significantly reduced, implying that ADP·P_i trapping alone in D1 is insufficient to hyperactivate the D2 domain. In an alternative interpretation, hyperstimulation of the D2 domain is insufficient to achieve a fast total ATPase rate in the assay if hydrolysis is coordinated between both domains, requiring hydrolysis in the D1 domain as well – which is stalled.

Surprisingly, measurements of the hyperactive F360[A/P] mutants also show reduced ATPase rates in the full-length context. This suggests that the rotamer switch of F360 – a key feature of the ADP·P_i state – has larger implications for the D1-D2 communication.

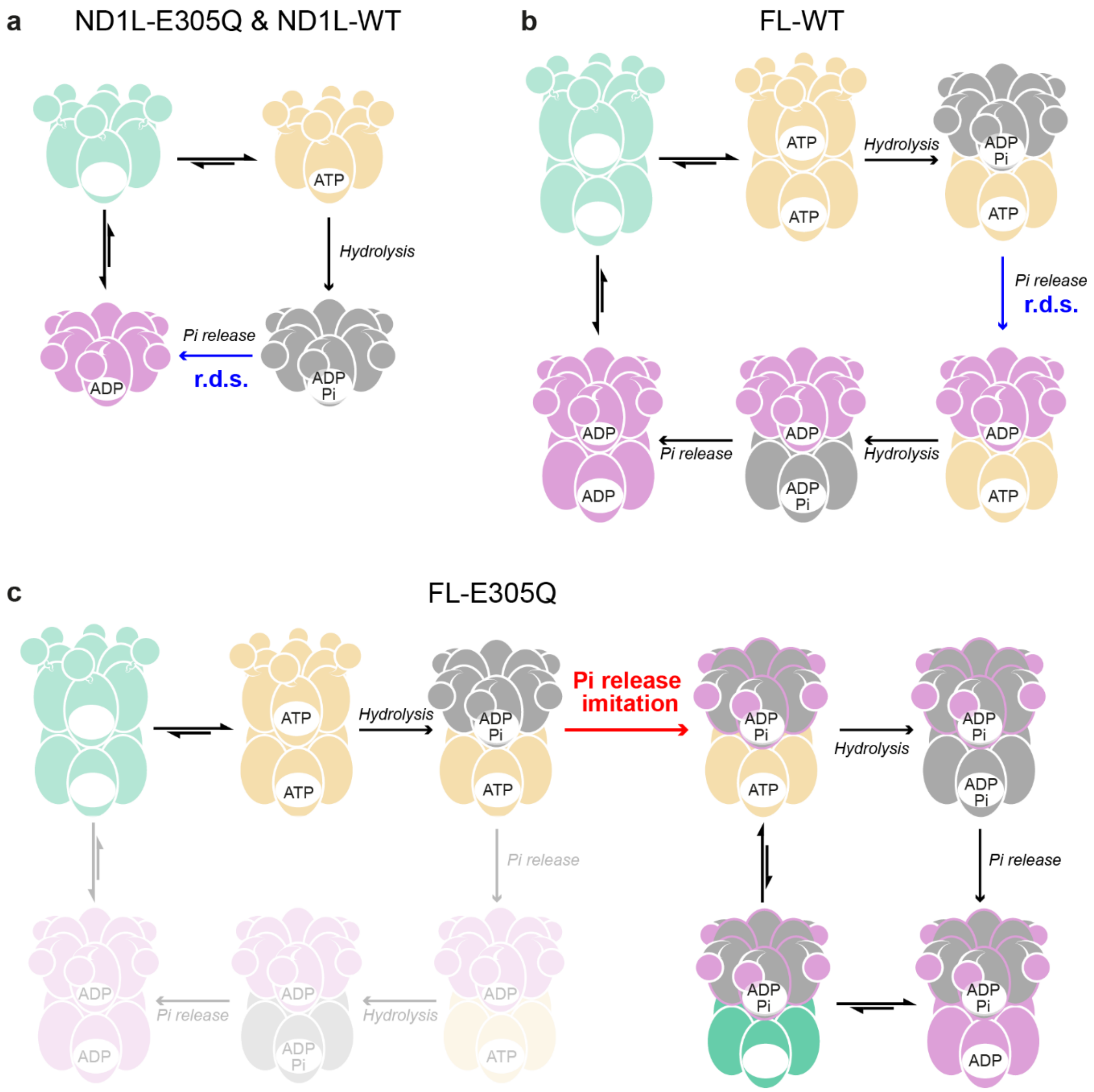


Figure 73. ATP hydrolysis model of wildtype and mutant full length p97. a) Under ATP turnover, both ND1L-wt and ND1L-E305Q form stable ADP·P_i states, although the ATPase activities are vastly different. In both constructs, phosphate release is the rate-determining step (r.d.s). **b**) In the full length wild-type p97 the rate determining step is the P_i release from the D1 domain. Most likely, afterwards ATP is hydrolysed in D2, until both binding pockets are filled with ADP, which is exchanged through one or several intermediate apo states with new ATP. **c**) In the full length E305Q mutant, the D1 domain is significantly stalled in its activity. Instead, the sensor loop of the D1 domain imitates a P_i release signal, hyperstimulating the ATPase activity of the D2 domain. While the normal reaction pathway is still accessible (desaturated), it is less likely to occur.

5.6.4 Kinetic analysis of the ATPase reaction using ATPγS

General analysis of ATPase measurements

So far, we have only considered the protein state of p97 during the ATPase reaction but have neglected the kinetic analysis. The nucleotide and PEP concentration is monitored by ^{1}H 1Ds over the course of the reaction. Using peak integration, one can determine the present concentrations in the solution and use this data to model p97 kinetics. In Figure 74, an example of a typical reaction is shown. Concentrations of PEP and all nucleotides in solution can be determined by integrating and fitting the respective peaks in ^{1}H 1D spectra over time. Respective regions are indicated in Figure 74a. For PEP, the sum integral of both peaks of the first 1D spectrum is set to 100%. Concentrations of nucleotides are calculated from their respective ratios, assuming a constant total concentration (Figure 74b).

A qualitative analysis of the reaction progression allows several observations. Initially, PEP is consumed at an approximately constant rate, indicating steady-state kinetics (Figure 74b,c). (The quantification of PEP consumption rate is explained below). During that phase, ADP concentration is kept below the detection limit. However, over time, the rate of PEP loss steadily declines. Moreover, ADP starts to emerge in the solution, although a substantial concentration of PEP can still be present [55]. After PEP is completely depleted, ATP is rapidly converted to ADP.

These observations can be explained by two separate issues. First, during ATP hydrolysis, inorganic phosphate is continuously liberated and accumulates over time. This small molecule acts as an allosteric inhibitor for the pyruvate kinase [108,109]. (ATP is also an inhibitor of the pyruvate kinase, but not at the low concentrations used in the NMR experiment [140]). Initially irrelevant, over time, the rising phosphate concentration in the solution starts to inhibit ATP regeneration, leading to a small prevalence of ADP in the solution. ADP, however, is efficient at stalling p97 [55], resulting in reduced ATP hydrolysis and PEP consumption rates and, finally, the stabilisation at a plateau.

The severity of these issues can be partially reduced through reaction setup optimisation. By decreasing the starting PEP concentration, the total runtime of the ATPase reaction is reduced, but the amount of phosphate is capped at less inhibitive concentrations. Additionally, to lessen the "workload" on the pyruvate kinase, the ATPase activity of p97 can be decreased by reducing the p97 and Mg^{2+} concentrations, increasing the NaCl concentration and lowering the temperature [56]. Following the optimised reaction setup, the rate of PEP hydrolysis becomes steadier and less exponential.

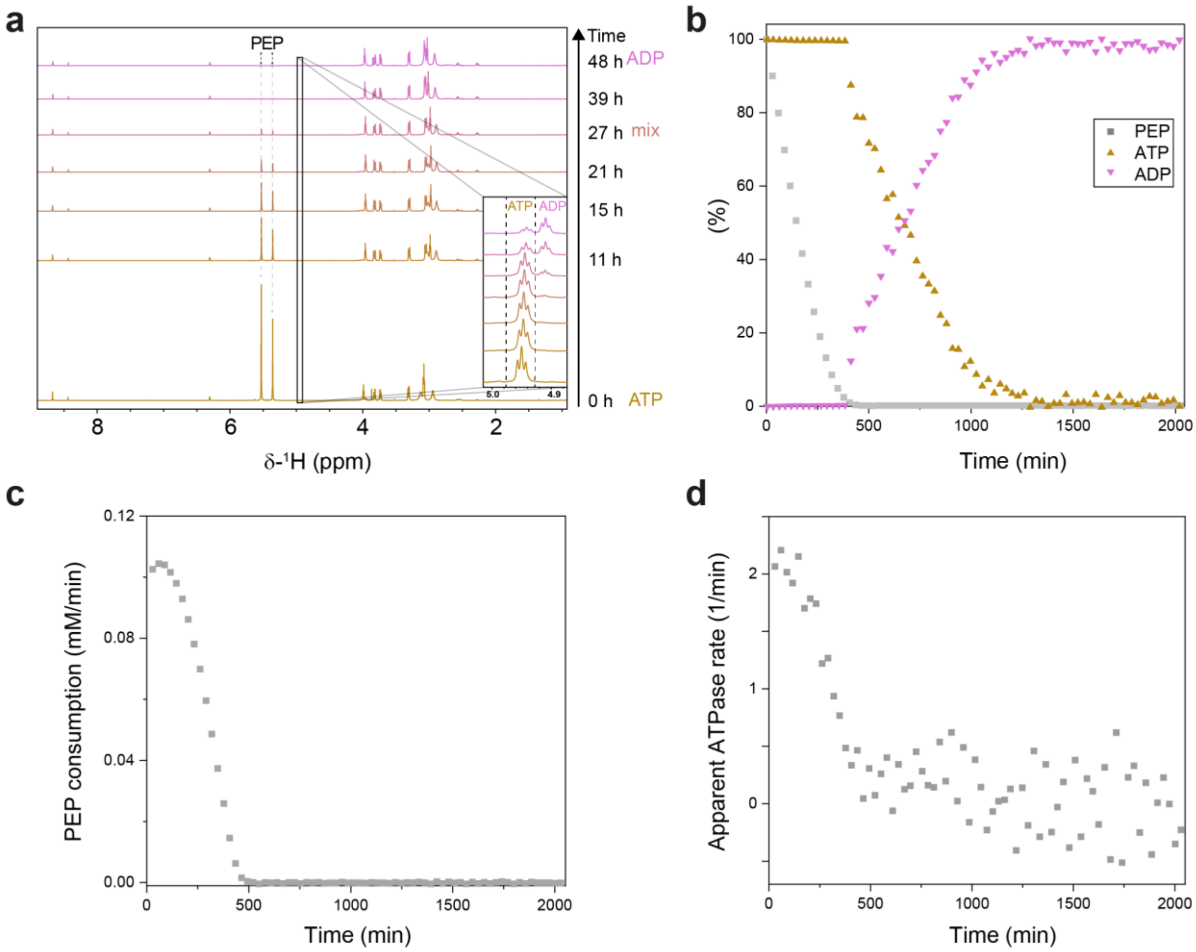


Figure 74. Analysis of reaction kinetics under ATP turnover by ND1L-wt. Small molecule concentrations are determined from ^{1}H 1Ds over the course of the ATPase measurement. Only three components are monitored: PEP, ATP and ADP. **a**) Representative ^{1}H 1D showing PEP depletion and conversion of ATP into ADP. The figure is adapted from ref. [55] **b**) Small molecule concentrations during ATP turnover. PEP concentration initially decreases in a linear fashion but then continuously slows down. While PEP is present, no traces of ADP are detected in the solution. After PEP is completely depleted, ADP concentration rapidly. **c**) Plot of the PEP consumption rate over time. Following an initially constant rate, the rate falls steadily to zero until no PEP is left in solution. **d**) Plot of the apparent ATP turnover rate over time. Plots shown in panels **c**) and **d**) are determined from the small molecule concentrations in the solution shown in **b**). Results shown in panels **b-d**) originate from another measurement than the one shown in **a**). The measurement was conducted at 40°C.

Quantification of ATPase real-time kinetics

Reaction monitoring can be used as a form of real-time kinetics. Pure PEP consumption rate is determined from the slope of the PEP concentration, [PEP], between two measurement points in time *i* and *f* with respect to time *t:*

$$v_{consumption}^{PEP} = -\frac{d[PEP]}{dt} = -\frac{[PEP]_f - [PEP]_i}{t_f - t_i}$$

Due to the constant regeneration of ATP from ADP by PEP, the change in ATP concentration does not fully reflect the ATPase activity of p97. When determining the rate of ATP turnover, PEP turnover has to be taken into account. Thus, the calculation of ATP hydrolysis rate, $v_{hydrolysis}^{ATP}$, was based on the derivatives of both the concentrations of ATP and PEP between two consecutive measurement points:

$$v_{hydrolysis}^{ATP} = -\left(\frac{d[ATP]}{dt} + \frac{d[PEP]}{dt}\right) = -\left(\frac{[ATP]_f - [ATP]_i}{t_f - t_i} + \frac{[PEP]_f - [PEP]_i}{t_f - t_i}\right)$$

with [ATP] or [PEP] being the respective concentrations.

Finally, ATPase rates were set in relation to the concentration of p97, [p97], to obtain turnover constants:

$$\boldsymbol{k_{Pi-release}^{ATP} = \frac{v_{hydrolysis}^{ATP}}{[p97]}}$$

Considering that the rate-determining step of the reaction is phosphate release [55], the hydrolysis constants of ATP turnover are denoted as $k_{Pi-release}^{ATP}$, was. Based on these equations, the ATPase rate of ND1L-wt is estimated to reach ~2 min^{-1} (Figure 74d). This estimation is smaller than the values determined for ND1L-wt at 37°C with the ATPase assay by approximately a factor of five (Section 5.2.2). This emphasises that even in an optimised reaction setup, p97 might not be able to reach its maximal ATPase rate. Reasons for this behaviour could possibly be attributed to the inhibition of p97 by low-level (below detection limit) concentrations of ADP in solution either due to slow ADP regeneration or diffusion effects inside the NMR tube or due to the usage of thawed protein stocks: experience with the ATPase assay shows that p97 tends to partially lose its ATPase activity upon freezing at -80°C. Furthermore, measurements shown in Figure 74 were conducted in D_2O, whose isotope effect might reduce the hydrolysis rate [141].

NMR measurements of p97 in the presence of ATPγS and the regeneration system

This approach can be extended to measurements of ATPγS in the presence of the regeneration system. ATPγS is a slowly hydrolysable analogue of ATP, and in its presence, the ATPase activity of p97 is hampered. By combining ATPγS and the regeneration system, one can vary the ratio between ATPγS and ATP over the course of the reaction, leading to different ratios of bound ATPγS and ATP in the binding pockets of the hexameric ring. Combining the results of the measured kinetics with the protein NMR measurements could give insight into the interprotomer coordination governing nucleotide hydrolysis. The approach of mixing active and inhibited subunits has been previously used to study AAA ATPases [142,143] but has been rarely used on p97 due to the stability of the hexameric assemblies [40].

The usage of ATPγS instead of ATP introduces some complications and surprising results. General results have been outlined previously in Chapter 4. The main observation of the measurement – the detection of the ADP·P$_i$ state – will be discussed in detail below. In the following, the quantification of the reaction kinetics is explained.

The presence of ATPγS at the beginning of the reaction introduces a third nucleotide species in the solution. Commercially available ATPγS is contaminated with ADP, which is immediately converted to ATP by the regeneration system. Over the course of the measurement, ATPγS is continuously hydrolysed, leading to an increase in ATP concentration in the solution (Figure 75a). These changes complicate the analysis. Due to the lack of overlap-free peaks representing the individual nucleotides, peak area ratios of two different regions were set in relation to each other. H8-proton peaks were fitted to deduce ATPγS concentration based on the ratio of the respective peak areas. Starting from the point when ADP was detectable in the spectra, its concentration was determined from the integral ratios of ribose peaks. Prior to that, ADP concentration was set to 0.

Thus, the ratio of respective nucleotides in solution, $\{nucleotide\}$, can be described by

$$\{ADP\} = \frac{integral(ADP, ribose)}{integral(all\ signals, ribose)} \times 100\%$$

$$\{ATP\gamma S\} = \frac{integral(ATP\gamma S, H8)}{integral(all\ signals, H8)} \times 100\%$$

$$\{ATP\} = 100\% - \{ATP\gamma S\} - \{ADP\}$$

Integrals were scaled with empirically determined factors (19 for ATP/ATPγS and 40 for ADP) to adjust for the effect water suppression on ribose peaks.

Finally, the concentration of respective nucleotides, [*nucleotide*], was determined from their ratios and assuming a constant total concentration:

$$[ATP] = \frac{\{ATP\} \times 5mM}{100\%}$$

The determination of the individual true nucleotide concentrations from relative integrals has the advantage of ignoring the loss of peak intensity due to the deuteration of the H8 position over time [144].

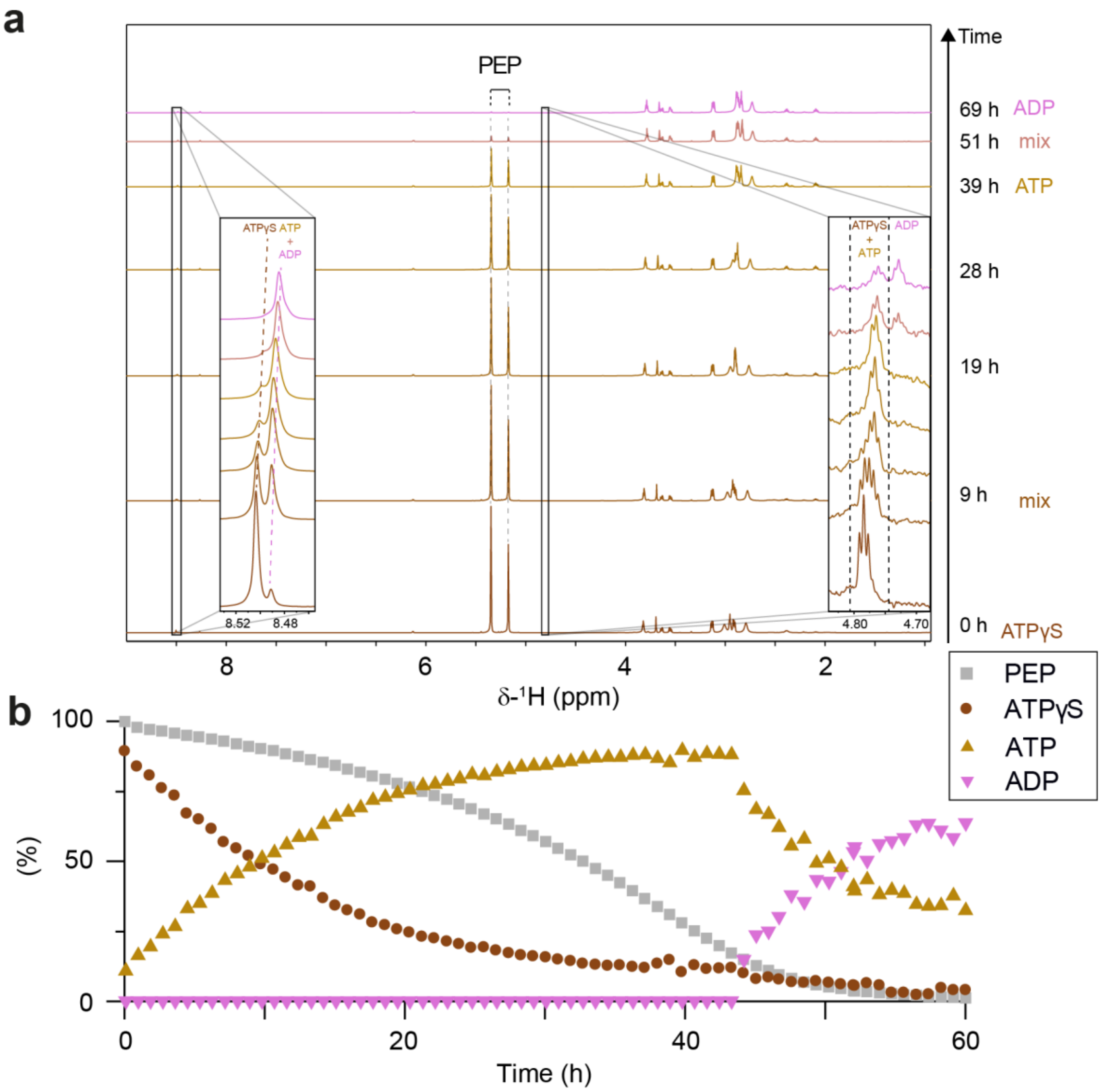


Figure 75. ATPase reaction of ND1L-wt with ATPγS. Small molecule concentrations are determined from ^{1}H 1Ds over the course of the ATPase measurement. Four components are monitored: PEP, ATPγS ATP and ADP. **a**) Representative ^{1}H 1D showing PEP depletion and conversion of ATPγS into ATP and later into ADP. **b**) Small molecule concentrations during ATP turnover. ATPγS concentration traces a seemingly exponential decay, leading to an increase in ATP concentration over time. At the same time, PEP consumption rate gradually increases until ADP starts to emerge, leading to additional p97 stalling. The measurement was conducted at 50°C.

From the nucleotide concentrations, the local rate of ATPγS hydrolysis, $v_{hydrolysis}^{ATP\gamma S}$, can be determined analogous to the rate of ATP hydrolysis, but without the PEP component (ATPγS hydrolysis rate calculation is not dependent on PEP consumption):

$$v_{hydrolysis}^{ATP\gamma S} = \frac{-d[ATP\gamma S]}{dt} = \frac{-[ATP\gamma S]_f - [ATP\gamma S]_i}{t_f - t_i}$$

and the hydrolysis constant $k_{hydrolysis}^{ATP\gamma S}$

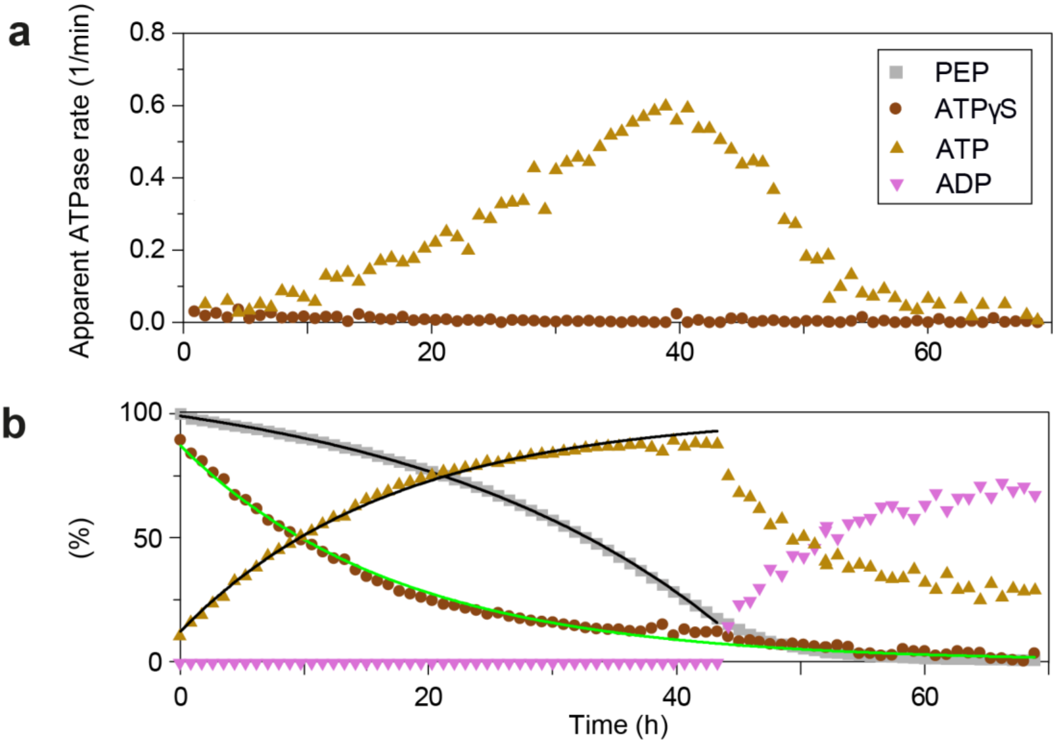
$$\mathrm{k}_{hydrolysis}^{\mathrm{ATP\gamma S}} = \frac{\mathrm{v}_{\mathrm{hydrolysis}}^{\mathrm{ATP\gamma S}}}{[\mathrm{p97}]}$$

The definition of such a hydrolysis constant, however, is of limited value since a major contribution to ATPγS degradation arises from (thermal) autohydrolysis, as will be shown below. This definition is only used to compare values with the observed ATP hydrolysis constant.

Analysis of the reaction

ATPγS concentration in solution decreases over time following a seemingly exponential-like decay. In consequence, ATP concentration increases following a reversed curve. In contrast to the ATP measurement, PEP consumption speeds up over the course of the measurement instead of slowing down (Figure 76b). After approx. 42 h, ADP starts to emerge in solution, which is accompanied by a reduction in PEP consumption rate. The emerging ADP starts to stall p97 further, leading to a slow conversion of ATP to ADP to a probable equilibrium around ~65% ADP.

The acceleration of PEP consumption is reflective of the continuous increase in the ATP turnover rate (Figure 76a). This is an expected consequence of the inhibitive properties of ATPγS and its replacement by ATP over the course of the measurement. Figure 76a shows that the values of hydrolysis constants of ATPγS and ATP are initially comparable but quickly drift apart as the rate of ATP turnover increases, while the rate of ATPγS keeps dropping. The ATP turnover rate reaches its maximum at ~0.6 min^{-1} and decreases thereafter as ADP starts to build up in solution. This limit is significantly below the ATPase rate determined by the ATPase assay and even slower than the rate measured in the reference ATPase NMR experiment (~2 min^{-1} at 40°C). Logically, the assumption is that in an unlimited ATPase measurement unrestricted by the PEP amount, the final ATPase rate (after all ATPγS is hydrolysed) would equate to the reference rate. Figure 76a suggests that the ATP turnover rate reached a stable plateau after ~40 h, but it is also conceivable that at that point, low ADP concentrations (below the detection limit) already started to stall p97.



c

Equation	$y(t) = y_0 + A \cdot e^{Rt}$		
Molecule	PEP	ATPγS	ATP
y_0	117.7 ± 0.7	0	100
A	-18.7 ± 0.5	87.2 ± 0.8	-87.9 ± 1.0
R	0.039 ± 0.0005	-0.057 ± 0.001	-0.058 ± 0.001
R^2	0.999	0.992	0.991

Figure 76. Kinetic analysis of the ATPase reaction of ND1L-wt with ATPγS. a) Plot of the apparent hydrolysis constants over time. Initially, both ATP and ATPγS are hydrolysed at similar rates. While the hydrolysis rate of ATPγS keeps decreasing over time, the ATP turnover rate reaches a maximum of ~0.6 min^{-1} after ~40 h and starts to decrease due to emergence of ADP. **b**) ATPγS hydrolysis rate follows an exponential decay. The small molecule ratios in solution, shown in Figure 75b, are modelled by exponential decay of the form $y(t) = y_0 + Ae^{Rt}$. Until ADP appears in solution, the growth of ATP ratio is an exact reverse of the ATPγS decay owing to the way both ratios are calculated. PEP consumption rate can be approximated by an exponential rate until the emergence of ADP. **c**) Fitting parameters of curves shown in **b**).

In Figure 76b, the changes of small molecule ratios in solution have been modelled by the single exponential decay of the form $\mathrm{y(t) = y_0 + Ae^{Rt}}$. The similarity between the decay/growth rates of ATPγS and ATP is not a coincidence but a consequence of the calculation of nucleotide ratios. Up to the point when the ADP concentration is taken into account, the exponential growth of ATP is the exact reverse of the ATPγS decay. In contrast, the concentration of PEP is determined independently from the nucleotides. The exponential increase in the PEP consumption rate can most likely be attributed to the substitution of the competitive inhibitor ATPγS by the substrate ATP, leading to faster overall turnover rates.

Control measurements

The success of the simple exponential model of ATPγS behaviour in solution raises the question of whether p97 is even involved in the hydrolysis or whether the exponential decay is indicative of (thermal) autohydrolysis. We conducted several control measurements in which we compared the rates of ATPγS decay with or without the regeneration system or p97. The experimental setup was designed to minimise the variability between the different experiments, i.e. the nucleotide and Mg^{2+} concentrations and the ionic strength concentration were kept identical in all four measurements (Table 14). The concentration of ATPγS over time in these control experiments is shown in Figure 77. All ATPγS decay curves could be fitted with single exponentials of the form $y(t) = Ae^{-Rt}$, but with different rates.

ATPγS decayed at identical rates in the control measurements alone or with p97, implying that p97 does not contribute to the hydrolysis, which instead is purely driven by autohydrolysis. In the presence of the regeneration system, ATPγS decays at a slightly enhanced rate. However, whether the enhanced hydrolysis rate can be attributed to interactions with the kinase or with PEP remains unclear. The rate of decay increased even further when both the regeneration system and p97 were present in the solution. This implies that p97 facilitates ATPγS hydrolysis, but only when ATP is present. This interpretation supports the hypothesis of a coordinated mechanism of ATP hydrolysis around the hexamer assembly: initiated hydrolysis of ATP in one binding pocket would support hydrolysis in a neighbouring pocket with bound ATPγS.

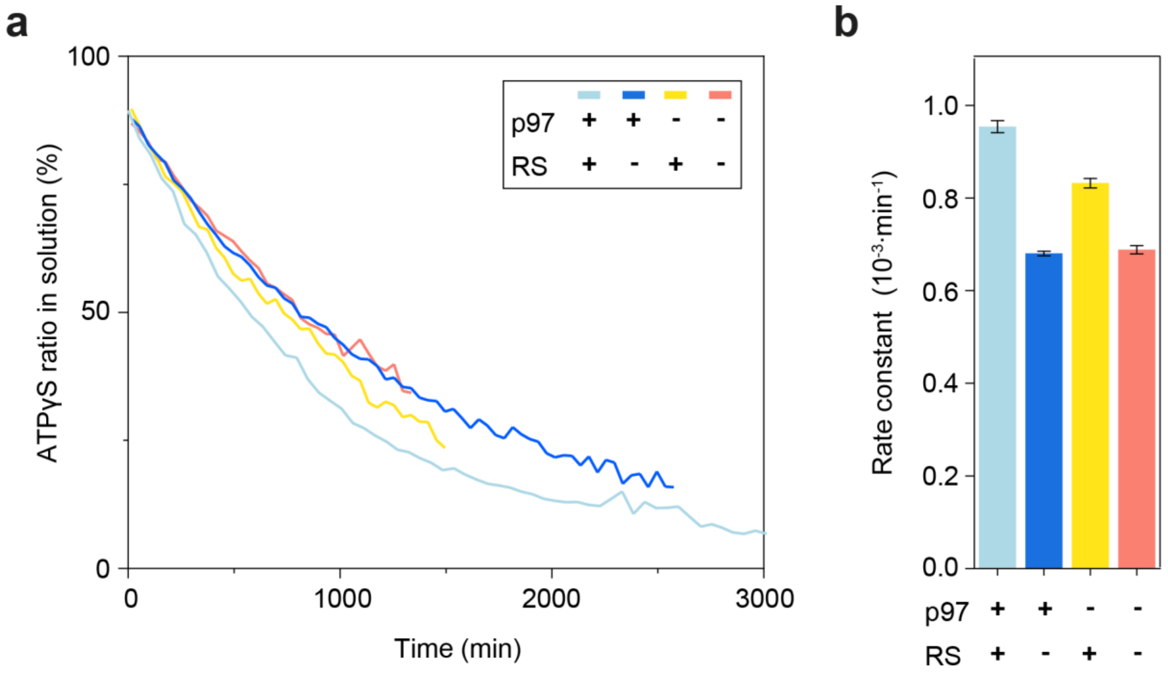


Figure 77. Comparison of ATPγS decay rates between control measurements. Changes in ATPγS concentration in solution over time have been associated with exponential decay. **a**) Ratio of ATPγS in solution over time in respective control measurements of ATPγS alone (red), with p97 (dark blue), with the regeneration system (RS) (yellow) or with both the regeneration system and p97 (light blue) are presented. ATPγS decays at identical rates both in the presence and absence of p97, implying that the decay originates from autohydrolysis. The degradation rate is increased in the presence of the regeneration system, confirming that regeneration system components facilitate ATPγS hydrolysis. The fastest rate is observed in the sample with both p97 and the regeneration system. **b**) Comparison of the rate constants from exponential fits of the form $y(t) = Ae^{-Rt}$ determined from the control measurements shown in **a**). In contrast, ATP is stable in solution at 50°C (data not shown).

Table 14. Experimental setup of the ATPγS control measurements. All measurements were conducted in the standard solution-state NMR buffer (25 mM HEPES pH 7.5, 25 mM NaCl, 100% D_2O). Abbreviations: PK = pyruvate kinase, NaP$_i$ = sodium phosphate pH adjusted to 7.5, R5P = ribose-5-phosphate

Conditions				
p97	**+**	**+**	**-**	**-**
Regeneration system	**+**	**-**	**+**	**-**

Buffer components	Concentrations			
p97	140 µM	140 µM	-	-
ATPγS	5 mM	5 mM	5 mM	5 mM
PEP	220 mM	-	220 mM	-
PK	40 units	-	40 units	-
NaP$_i$	-	220 mM	-	220 mM
$MgCl_2$	4 mM	4 mM	4 mM	4 mM
R5P	4 mM	4 mM	4 mM	4 mM
KCl	50 mM	50 mM	50 mM	50 mM

Protein spectra of ND1L-wt in the presence of ATPγS and the regeneration system

Dissociation constant of ATPγS for p97 have been reported to be approx. sixfold lower than that of ATP [47]. In the presence of the regeneration system and ATPγS, the nucleotide ratio ATPγS:ATP at the beginning of the measurement is ~90:10, rendering ATP binding unlikely. Considering that ATPγS has been used to characterise pre-hydrolysis states of p97 (Figure 17), one would expect that the prevalence of ATPγS at the beginning of the measurement would lead to a pre-hydrolysis-like protein state. However, instead, an ADP·P$_i$ state is formed with spectra barely differing from the ones in the presence of pure ATP (Figure 78a). It is also unreasonable to assume that p97 selectively binds ATP instead of ATPγS when both are present in solution: the initially low, but increasing PEP consumption over the course of the reaction confirms the enzymatic inhibition by ATPγS. One logical conclusion to be drawn from the spectra is that the presence of the minor ATP population supports the formation of the ADP·P$_i$ state – or rather of the ADP·SP$_i$ state. Similarly to the reference measurement, the rate-determining step of the ATPase cycle with ATPγS is the (thio-)phosphate release. This is reflected in three peak shifts: the ATPγS-like positions of I353 and M344 of the ADP·SP$_i$ state are reminiscent of the behaviour of slow mutants and correlate to deficient product release. Additionally, the peak position of I206 is shifted from the ADP·P$_i$-like position towards the

ADP reference, which could be attributed to a chemically different species – ADP·SP_i (Figure 78b,c).

These observations lead to the conclusion that the ND1L-wt hexamer is capable of hydrolysing ATPγS once hydrolysis of ATP has been initiated in one or several of its nucleotide binding pockets. Afterwards, an ADP·SP_i state is formed, and thiophosphate release is initiated. This process is seemingly less efficient than the release of inorganic phosphate.

While the rate-determining step (phosphate release) remains unchanged but obviously slowed down, it is unclear whether the rate of the hydrolysis step itself is also reduced. Still, this interpretation follows the kinetic studies of ATPγS decay, confirming that the nucleotide can indeed be processed by p97, thus slightly increasing its decay rate.

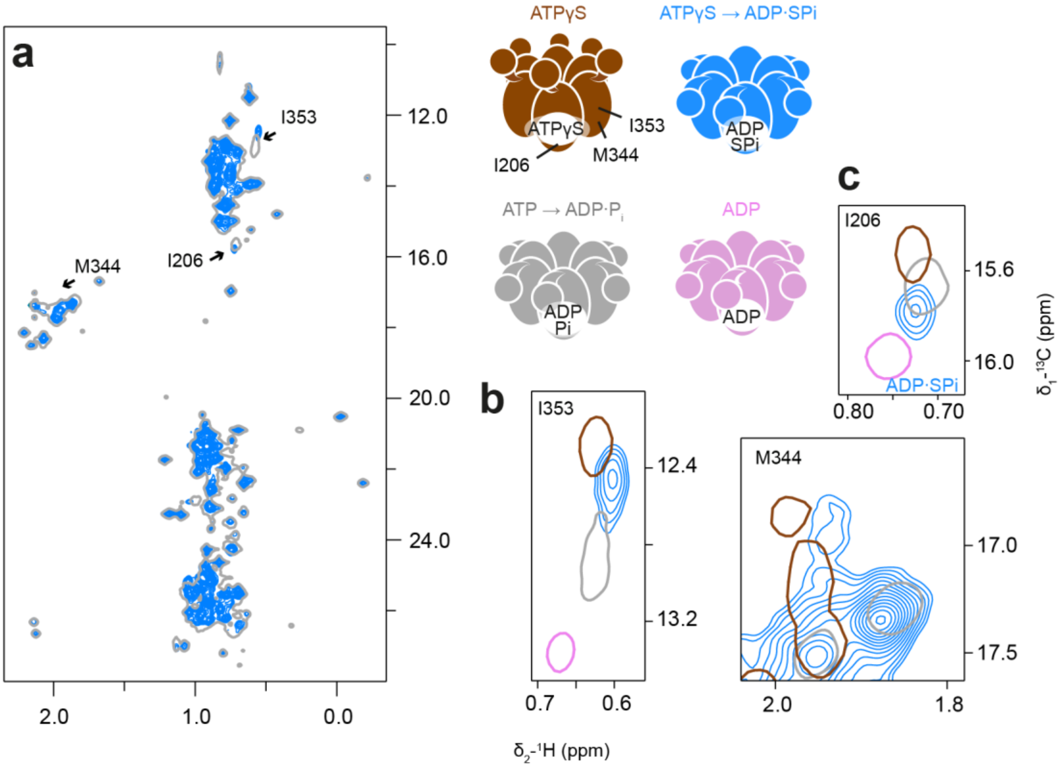


Figure 78. ND1L-wt forms an ADP·SP_i state in the presence of ATPγS and the regeneration system. Spectral regions of ^{1}H-^{13}C-HMQC spectra are shown. **a**) Spectra of the reference ADP·P_i state and the ADP·SP_i state overlap almost perfectly. **b**) Residues M344 and I353 report on the secondary structure of the sensor loop. In the ADP·SP_i state the loop adopts a pre-hydrolysis conformation, reminiscent of slow ND1L mutants. **c**) I206 informs on the type of nucleotide bound in the binding pocket. The ADP·SP_i peak is shifted towards the ADP reference, implying the presence of chemically similar, but distinct ADP·SP_i nucleotide states.

5.7 Discussion

5.7.1 Functional roles of nucleotide binding pocket residues

In this work, site-directed mutagenesis was used to investigate individual contributions of conserved amino acids inside the nucleotide binding pocket of p97. ND1L mutants were evaluated based on multiple characteristics: oligomeric state, nucleotide binding affinity, response of the N-terminal domain regarding the nucleotide state, ability to hydrolyse ATP, ability to form the ADP·P$_i$ state and efficiency of P$_i$ release. These characteristics were evaluated using biophysical methods, such as SEC, ITC, ATPase assay and NMR. Additionally, results from cryo-EM and MD were considered. The residues of interest were selected based on their roles in the protein sequence motifs or their importance, as observed in MD simulations and cryo-EM structures. The results of these assessments are summarised in Figure 79.

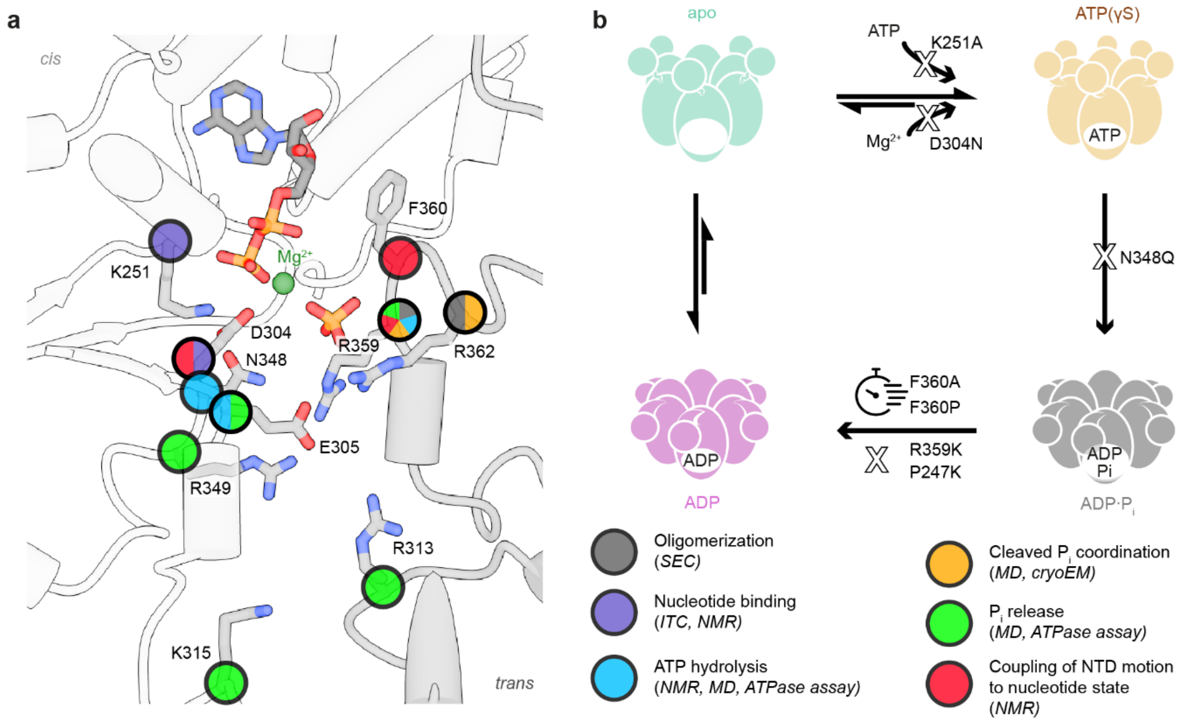


Figure 79. Functional summary of amino acids inside the p97 D1 nucleotide binding pocket. a) Summary of residue roles inside the nucleotide binding pocket. Residues are categorised by their role in oligomerisation (black), nucleotide binding (purple), ATP hydrolysis (blue), coordination of the cleaved P$_i$ (orange), P$_i$ release (green) and coupling of the NTD motion to the nucleotide state (red). The assessment was performed based on biophysical methods: SEC, ITC, NMR, MD, ATPase assay, cryo-EM. **b**) Overview of mutation induced defects on the ATPase cycle. Figure adapted from ref. [52].

Additionally, we investigated a possible pathway for the dissociation of the P_i ion after ATP hydrolysis, marking the transition between the ADP·P_i and ADP states. We found a positively charged channel leading from the nucleotide binding pocket directly to the central pore. The channel is lined with lysine and arginine residues, but mutagenesis studies suggest that only the Arg residues are central to phosphate ion transport.

5.7.2 ADP·P_i mimicking through phosphate binding

We investigated whether the ADP·P_i state could be artificially induced in p97 through the addition of exogenous phosphate. As previously reported [55], we failed to induce the ADP·P_i state in p97 starting from the ADP state. However, we were able to achieve an ADP·P_i-like state when we added high concentrations of inorganic phosphate (50-100 mM) to apo state ND1L. MD simulations show that the nucleotide binding pocket can adopt a similar conformation to the ADP·P_i state when two phosphate ions are bound [52]. Moreover, a similar trapping of two phosphates inside an apo binding pocket has been reported for the aspartate-β-semialdehyde dehydrogenase [74]. Worth noting is the fact that the magnesium ion itself is not required for the formation of the induced ADP·P_i state (Table 10). In MD simulations, monovalent cations from the surrounding solution were sufficient to balance the phosphate charges (Figure 60b) [52].

Additionally, we were able to induce the same protein state in p97 starting from the ATPγS-bound state, presumably through phosphate ions outcompeting ATPγS in binding inside the pocket. Again, this raises the question of why the same induced state cannot be reached starting from the ADP-bound state. A possible explanation is that since in the ADP-bound state, no Mg^{2+} is required, the metal ion leaves the nucleotide binding pocket, leading to an irreversible conformational rearrangement. This conformational rearrangement prevents phosphate binding inside the binding pocket, favouring ADP binding instead. This hypothesis is supported by the results of identical experiments performed with ND1L mutants in the apo states. These measurements suggest that the secondary structure of the P-loop could play a role in phosphate binding. Thus, the ADP·P_i state can be imitated starting from an empty binding pocket or one bound with ATP(γS). Another possibility is that phosphate binding can only occur starting from an NTD-up conformation (apo or ATP(γS) bound). The NTD conformation influencing the binding pocket structure is expected, considering that the type of bound nucleotides is normally telegraphed from the nucleotide binding pocket towards the NTD. Thus, a reverse connection is only natural.

When we performed the same experiment on the full-length p97, the NTD barely reacted to the high phosphate concentrations. Even increased phosphate concentrations up to 300 mM only lead to a minor repositioning of the N-terminal domain. The exact reason why the full-length protein fails to react to phosphate in a similar fashion as ND1L is unclear. Possibly, phosphate binding in the D2 binding pocket differs from that of D1 or the D2 nucleotide binding pocket remains empty altogether. Alternatively, the presence of the D2 domain impedes P_i binding in D1.

5.7.3 Sensor loop dynamics

In another study, we investigated the role of the so-called "sensor loop", a structural element (residues 348-362) directly connecting two neighbouring nucleotide binding pockets. This connection could be used to coordinate nucleotide processing between neighbouring protomers. However, so far, the role of the sensor loop has not been discussed in detail in the literature.
Presumably, this loop is one of the last structural elements to change its secondary structure from an ATPγS-like conformation in the ADP·P_i state to an ADP-like conformation after phosphate release. Indeed, in this work, we found a direct correlation between the sensor loop secondary structure and ATPase rate, possibly coupling structural rearrangement in the first half of the loop (residues 348-353) to P_i dissociation. We used the peak position of I353 to monitor the extent of secondary structure rearrangement in the loop, with I353 being a central residue of the loop itself. Surprisingly, we also observed a similar effect in the peak position of M344, a residue which is part of the β4 sheet leading to the sensor loop. However, this β-sheet does not undergo obvious secondary structure changes between the different protein states. For now, the exact reason for the conformational changes in the M344 side chain is unclear.
In order to study the importance of structural rearrangement, we mutated several residues that are part of the first half of the sensor loop. Mutations of residues R349, P350 and N351 to alanine reduced the ATPase rate to a similar extent but did not inactivate the protein. In contrast, triple full-length 349-351AAA mutants had no effect on ERAD activity, implying that these residues are not involved in the unfoldase function [45]. The effects evoked by these mutations may indeed differ from each other: for instance, R349 has been shown to take part in phosphate

release through its positive charge (Section 5.3) and has also been reported to be involved in communication with the D2 domain [45]. The double functionality of the arginine residue is reasonable since in the cryo-EM structure of full-length p97 under ATP turnover, the ADP·P_i state was only found in the D1 domain, while ATP was detected in the D2 binding pocket [52]. Thus, the phosphate release from the D1 domain needs to be communicated to the D2 domain to initiate ATP hydrolysis.

We attempted solution-state NMR to study the effect of the R349A mutation. However, the ND1L-R349A mutant does not display any major defects. Additionally, no obvious deviations from the wild-type behaviour were found for the P355A mutant, which was supposed to disrupt the structural conformation of the sensor loop. The mutant does not display any obvious defects – neither in the ATPase assay nor by NMR nor in unfoldase activity measurements [45], which is unsurprising, considering the lack of conservation of this residue (Figure 47).

Chapter 6: Materials and Methods

6.1 Molecular biology and protein purification

6.1.1 Construct design

Codon-optimised pET-28a+ plasmids encoding wild-type p97 (Uniprot P55072) for two different constructs (full-length p97 and ND1L p97) were supplied by authors of ref. [66]. These constructs contain an N-terminal His_6-tag, a short linker, and a TEV cleavage site (26 amino acids in total), followed by the proteins of interest (806 or 480 amino acids, respectively). The effective construct length after TEV cleavage is thus increased to 807 or 481 amino acids, respectively, due to the remaining glycine. Amino acid sequences for both constructs are shown in Supplementary Figure 1.

6.1.2 Cloning

Point mutations were introduced using site-directed mutagenesis. Forward and reverse primers were designed using an online tool, NEBaseChanger (NEB), and were purchased from IDT. Reaction enzymes were purchased in a kit (Q5 Site-Directed Mutagenesis Kit) from NEB. For the cloning procedure, the manufacturer's protocols for the usage of the kit were followed. Briefly, template DNA was amplified by the Q5 DNA polymerase in a thermal cycler reaction setup in the presence of primers containing the desired point mutation. Then, an enzyme mix was added to digest the template and to ligate the amplified product. Chemically competent *E.coli* (DH10β) were then transformed with the new plasmid and plated onto antibiotic (kanamycin) containing LB-agar plates. Selected colonies were grown in overnight cultures, harvested, and the plasmid extracted using a miniprep DNA purification kit. Successful mutations were confirmed by Sanger sequencing (Genewiz).

6.1.3 Protein expression

Non-labelled proteins were expressed following standard expression protocols in LB or TB media. ILVM-labelled proteins were produced by growing cells in M9 minimal media supplied with isotope-labelled carbon (D-glucose) and nitrogen (NH_4Cl) sources in D_2O. The protocol for the preparation of the M9 minimal media is found in Supplementary Table 4. Additionally, the medium was supplied with selectively labelled amino acids or amino acid precursors (see Supplementary Table 5).

Non-labelled proteins

E.coli BL21 (DE3) cells were transformed with plasmids containing the desired constructs and selected on kanamycin-containing agar plates. For pre-cultures, 5-10 mL LB medium supplied with kanamycin was inoculated with selected colonies and grown overnight at 37°C. The following morning, pre-cultures were diluted into main cultures of 1 L of LB or TB media (supplied with kanamycin) and grown at 37°C to an optical density value at 600 nm (OD600) of 0.6-0.9. The media was supplemented with 0.5 mM of IPTG and the proteins were expressed overnight at 18°C. In the morning, after the media had reached OD600 values of 3.0-4.0, the cells were harvested.

ILVM-labelled proteins

E.coli BL21 (DE3) Star cells were transformed with plasmids containing the desired constructs and selected on kanamycin-containing agar plates. For pre-cultures, 5-10 mL LB medium supplied with kanamycin was inoculated with selected colonies and grown for 4-5 hours at 37°C. At an OD600 of ~ 2, the cells were spun down and resuspended in 25 mL of perdeuterated M9 minimal media supplied with kanamycin and [$^2D_7$, $^{12}C_6$]-D-glucose. After an overnight adaptation period at 37°C, the culture was added to the main solution of 225 mL and grown at 37°C to OD600 = 0.8. Then, labelled methionine and Leu/Val precursors were added, and after an incubation period of 40 min, the medium was supplied with the last (Ile) precursor (Supplementary Table 5). After another 20 min, overnight expression was induced with 1 mM IPTG at 18°C. The next day, cells were harvested at OD600 values of 2.2-2.5.

6.1.4 Protein purification

Protein constructs were purified using a multistep purification approach. Briefly, cells were subjected to lysis, and the insoluble parts were separated by ultracentrifugation. The supernatant was purified by Ni-immobilised metal affinity chromatography (Ni-IMAC), and the His_6-tag was cleaved using TEV protease. After dialysis, ND1L constructs were concentrated, and the purity was improved in a polishing step using size exclusion chromatography (SEC). Full-length constructs were first passed through anion exchange chromatography (AIEX) since SEC alone was not sufficient to separate nucleic acid impurities from the protein for these constructs. Afterwards, eluted full-length p97 was concentrated and subjected to SEC as well. Finally, protein samples were concentrated to the desired volume.

Cell lysis

Cell pellets containing the overexpressed constructs were resolubilised in 30-50 mL p97 lysis buffer (100 mM Tris pH 7.4, 500 mM NaCl, 5 mM $MgCl_2$, 20 mM imidazole, 5 % glycerol, 2 mM BME, 1 mg/mL lysozyme, 150-300 units DNAse I, Protease Inhibitor mix HP Plus 1X; pre-cooled on ice) using a homogeniser. Resuspended cells were lysed using a French press over several runs until the suspension was transparent and non-viscous. Debris were centrifugated down at 40000 g for 30-60 min at 4°C. The supernatant was decanted and filtered through a 0.45 µm syringe filter.

Ni-IMAC and TEV cleavage

Approx. 8 mL of Ni-NTA resin were equilibrated with pre-cooled p97 wash buffer (50 mM HEPES pH 7.4, 150 mM KCl, 5 $MgCl_2$, 20 mM imidazole). The filtered supernatant was then passed through the column bed using gravity flow to ensure complete binding. The resin was washed extensively (10-20 CV) using pre-cooled p97 wash buffer until the flow-through solution stopped reacting in a Bradford assay (examined by eye). The bound protein was then eluted from the column using 30-50 mL of the pre-cooled p97 elution buffer (50 mM HEPES pH 7.4, 150 mM KCl, 5 $MgCl_2$, 300 mM imidazole). TEV protease (200 µL, 1 mg/mL) was then added to the combined elution fractions and the solution was transferred into a dialysis tubing with a 10 kDa cutoff. The filled tubing was then placed into 2 L of p97 TEV cleavage buffer (20 mM HEPES pH 7.5, 250 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 2 mM BME) for overnight dialysis at 4°C. The next morning, the solution was removed from the tubing and filtered through a 0.22 µm syringe filter.

Anion Exchange Chromatography

Prior to SEC, full-length p97 constructs were subjected to AIEX using the ÄKTA FPLC system with two buffers (p97 AIEX low salt buffer: 50 mM HEPES pH 7.0, 50 mM KCl; p97 AIEX high salt buffer: 50 mM HEPES pH 7.0, 1 M KCl). Specifically, a 5 mM HiTrap Q column was equilibrated with a pre-cooled low-salt buffer. The TEV-cleaved sample was then loaded onto the column while monitoring the UV absorbance in order to prevent column overloading and protein breakthrough. The column was consequently washed with 2 CVs of the low salt buffer and then with 5 CVs of a mixture containing 30 % high salt buffer, then 3 CVs with 50 %, 3 CVs with 75 % and finally 3 CVs with 100%. The protein was eluted in a volume of 3-4 mL at a concentration of 75 % of the high salt buffer, while nucleic acid contaminations were eluted at the 100 % high salt elution step.

Size exclusion chromatography

After dialysis, ND1L constructs were concentrated using ultracentrifugation concentration devices with cutoffs of 30 or 100 kDa to final volumes of 2-4 mL. Full-length constructs were concentrated with cutoffs of 100 kDa to similar final volumes after the AIEX step. Concentrated protein samples were injected into the ÄKTA FPLC system for subsequent SEC purification using a HiLoad 16/600 Superdex 200 pg column with the p97 SEC buffer (20 mM HEPES pH 7.5, 250 mM KCl, 1 mM $MgCl_2$). ND1L constructs eluted around 65 mL, full-length constructs around 55 mL.

Protein storage

SEC purified samples were concentrated to concentrations of 20-100 mg/mL. 1 mM DTT and 5 % glycerol were added, and the samples were flash-frozen and stored at -80°C.

6.1.5 Apyrase digestion

During purification, p97 is co-purified with pre-bound ADP, which needs to be removed to attain the apo-protein. Apyrase digestion is performed in order to hydrolyse ADP to AMP, which can no longer be bound by p97. Since apyrase is able to hydrolyse ADP and ATP, it needs to be completely removed from p97 prior to further experiments. Failure to do so can lead to misinterpretation of biochemical assays due to irregular nucleotide hydrolysis (true for both ITC and ATPase assay experiments) or even to degradation of nucleotides in an NMR sample, potentially creating an undesired apo state.

Apyrase is a 50 kDa large enzyme which comes in several isoforms with varying ADPase and ATPase activities. While even monomeric full-length p97 (89 kDa) is significantly larger than the apyrase, monomeric p97-ND1L has a similar molecular weight (53 kDa). Depending on the initial protein concentration, overall protein mass, possible present point mutations and/or isotope labelling, ND1L constructs can elute from SEC columns in mixed oligomeric forms, including hexa-, di- and monomers. Some mutants lose their hexameric forms, specifically in apo states (see Chapter 5). For this reason, for some samples, it can become difficult to fully separate p97 from the apyrase via SEC runs. Apyrase can be removed to a high degree from natural abundance high-concentration samples, as long as sample losses are tolerated and only hexamer-containing fractions are collected. Lower-concentration samples tend to elute in non-uniform oligomeric forms, which is especially true for isotope-labelled samples. However, these samples are costly, and losses are unacceptable.

For these reasons, two protocols have been adopted for ADP removal. In the first protocol, an apyrase isoform with a high ADPase activity (supplier: Sigma-Aldrich) is used, which leads to a fast, efficient and complete ADP hydrolysis. Afterwards, strict fraction separation is required for complete apyrase removal. In the second protocol, an isoform with a low ADPase rate (supplier: NEB) is used. Due to this low enzymatic activity, a complete removal of the apyrase is not as crucial. The disadvantage of the latter protocol is that, in many cases, it leads to an incomplete ADP hydrolysis, i.e., to a mixed ADP and apo protein state. It is worth noting that since ATPγS is resistant to hydrolysis by the apyrase, experiments involving this nucleotide do not suffer from nucleotide hydrolysis. In such cases, an incomplete separation of the highly active apyrase may still be tolerated.

The protocol is outlined in the following: 4 mM $CaCl_2$, 4 mM DTT, and 0.5-2 µL of the apyrase solution are added to a p97 sample of 0.1-1 mL. The volume of added reagent and the incubation temperature and time depend on the exact setup:

1) In case the highly active apyrase (Sigma-Aldrich; 1 unit/µL) was used, the sample was incubated for 3 h at room temperature. For guidance, 0.5 µL of the enzyme solution was sufficient to hydrolyse ADP from 10 mg of p97 in ~2 h.

2) In case the slowly active apyrase (NEB; 0.5 unit/µL) was used, the sample was incubated overnight at 30°C.

Digested samples were then purified via size exclusion chromatography using the p97 SEC buffer on Superdex 200 pg or Superdex 200 Increase columns, depending on the amount and size of the protein.

6.2 Biophysical protein characterisation

6.2.1 Isothermal titration calorimetry

Nucleotide binding affinities were determined by isothermal titration calorimetry (ITC) using a MicroCal PEAQ-ITC (Malvern Panalytical) at 25°C. Freshly apyrase digested (NEB) p97 samples were purified on a Superdex 200 Increase 10/300 GL column and further diluted to 10-20 µM if required. Lyophilised nucleotides (ATPγS, ADP) were dissolved in the identical buffer to concentrations of 100-120 µM. Measurements were performed using a starting 0.4 µL injection followed by 2 µL injections (19 times) with 120 s spacing. Data were analysed using the MicroCal PEAQ-ITC Analysis software (v1.21). For binding calculations, a One Set of Sites binding model was chosen. K_d values were determined from two to three technical replicates. ITC measurements were largely performed by K. D. Leitl.

6.2.2 ATPase assay

Introduction

ATPase rates were determined in an NADH-coupled ATPase assay. In this assay, ATP hydrolysis is proportionally coupled to conversion of NADH to NAD^+ (Figure 80). The concentration of NADH can be monitored in real-time by its absorbance at 340 nm, as the absorbance of NAD^+ at this wavelength is negligible. The advantage of this assay is the tolerance towards the high concentrations of ATP and the possibility of running the assay in combination with an ATP regeneration system to prevent stalling of p97 by ADP [55]. On the other hand, the assay setup requires a significant number of pipetting steps, and the readout is relatively insensitive. The latter can be partially mitigated by analysing the data in a continuous fashion.

NADH concentration c_{NADH} was calculated from the absorbance at 340 nm (A_{340}) according to the Beer-Lambert law:

$$A_{340} = c_{NADH} \cdot \varepsilon_{NADH}^{340} \cdot d$$

with ε_{NADH}^{340} being the molar attenuation coefficient of NADH at 340 nm and *d* being the optical path length. In order to correctly determine the NADH concentration in the individual measurement, the path length needs to be corrected for the actual depth of the reaction solution, background absorbance and meniscus form, which in turn is affected by buffer composition, temperature and protein concentration [145]. This was achieved by performing additional measurements at 977 nm (absorbance maximum of pure water) and at 900 nm (for the instrument background) [146]. The limitations of this approach are that no reaction components absorb at 900 and 977 nm and that the visual path length remains clear.

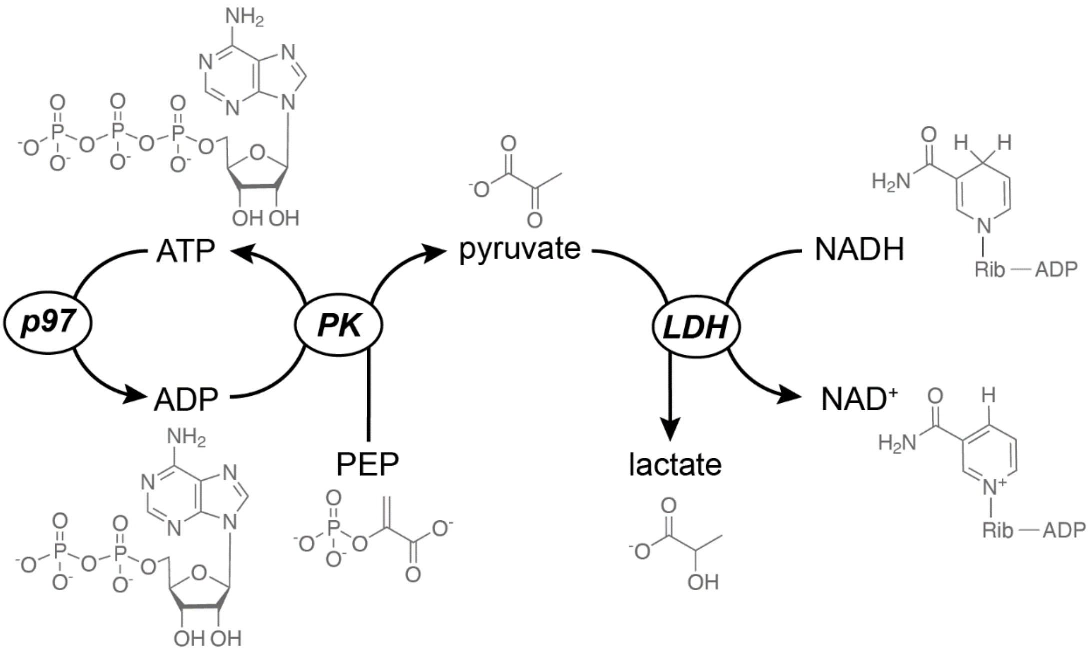


Figure 80. Working principle of the NADH-coupled ATPase assay. p97 hydrolyses ATP to ADP, which is immediately regenerated back to ATP by the pyruvate kinase (PK). In this secondary reaction PEP is transformed to pyruvate. The latter is then reduced to lactate by the lactate dehydrogenase (LDH) while NADH is converted to NAD^+.

Path length corrected absorbance $A_{340}(corrected)$ can then be calculated using absorbances at the two other wavelengths using a so-called K-factor [145]:

$$A_{340}(\text{corrected}) = A_{340}(\text{raw}) \cdot 10\,\frac{\text{mm}}{\text{pathlength}} = A_{340}(\text{raw}) \cdot \text{K-factor}/(A_{977} - A_{900}).$$

The K-factor for water is 0.186 at 37°C and 0.190 at 50°C, according to the manufacturer's manual.

Experimental implementation

The kinetic absorbance readout is performed in a temperature-controlled environment of a plate reader capable of measuring absorbances at 340, 900 and 977 nm (SpectraMax iD5, Molecular devices). The assay setup for a single measurement (120 µL) is shown in Table 15. Stock solutions of ATP, PEP and NADH were prepared by dissolving the respective powders into buffered solutions, pH adjusted to 7.5 and frozen at -20°C until further use. LDH was stored in the supplied ammonium sulphate suspension at 4°C prior to being diluted into H_2O to a concentration of 1 unit/µL directly before the measurement. Lyophilised powder of PK was dissolved in H_2O to a concentration of 1 unit/µL, aliquoted and stored at 4°C.

Table 15. Setup of the ATPase assay. Values for a single ATPase measurement with a volume of 120 µL are listed.

Component	Desired concentration [mM]	Stock concentration [mM]	Required stock volume [µL]
ATP[a]	2	50	4.8
PEP[a]	6	1420	0.5
NADH[a]	1	142	0.8
ATPase assay buffer, 10x[b]	1x	10x	12.0
H_2O			91.5
PK (from rabbit muscle)	1.2 units	1 unit/µL	1.2
LDH	1.2 units	1 unit/µL	1.2
p97	1-100 µM	10-1000 µM	12.0

[a]: Stocks of these components are buffered in 25 mM HEPES pH 7.5, 25 mM NaCl; [b]: Buffer composition: 250 mM HEPES pH 7.5, 250 mM NaCl, 40 mM $MgCl_2$, 500 mM KCl, 5 mM TCEP

In order to reduce pipetting errors, all components except ATP and p97 were combined in advance in a "master mix" solution, in which volumes of individual components are scaled up according to the planned number of reactions. For each reaction, 103.2 µL of the master mix was added into a single well of a transparent 96-well plate. Afterwards, p97 samples were added to the individual wells. The plate was then incubated at 37 or 50°C for 5 min to pre-equilibrate the reaction mixture while absorbance was recorded (see the beginning of the measurement in

Figure 81). Afterwards, ATP was added to the individual wells; the plate was sealed with a PCR-compatible adhesive film, and the measurement started. Caution is advised when using PCR-incompatible sealing tapes: these become cloudy over time and support condensation on the inner side of the film. This prevents accurate path length correction and should be avoided.

The assay was typically run at 37 or 50°C, respectively, for 60 min, although overnight measurements were used for very slow hydrolysing mutant constructs. Absorbances at 340, 900 and 977 nm were recorded in an interlaced fashion every 30 s (60 min measurement) or 20 min (overnight measurement), with the plate being shaken in between the recording.

Controls

ATPase rate measurements were determined in technical triplicates. Negative controls without p97 (with 12 µL of buffer instead) were run in triplicates in parallel. Positive controls involved a standard aliquoted wild-type ND1L sample with a known ATPase rate to check for systematic errors.

Data analysis

ATPase rate determination from absorbances was performed in several steps. First, raw absorbance data was converted to path length corrected absorbance values. Then, a region of constant NADH consumption was selected and approximated linearly. The slope *a* of the linear regression equates to

$$a = \frac{\Delta y}{\Delta x} = \frac{\Delta A}{\Delta t} = \frac{\Delta c \cdot \varepsilon \cdot d}{\Delta t}$$

with ΔA being the difference in absorbance and Δt being the difference in time. Using the molar attenuation coefficient of NADH ($\varepsilon_{NADH}^{340} = 6220\ \frac{L}{mol \cdot cm}$) [147] and a path length normalised distance $d = 1$ cm, this ratio was then converted to a difference in concentration of NADH:

$$\frac{\Delta A}{\varepsilon \cdot d \cdot \Delta t} = \frac{\Delta c(NADH)}{\Delta t}$$

Component degradation resulting from p97-independent reactions was then accounted for by subtracting the NADH consumption rates of the negative controls from the actual measurements:

$$\left(\frac{\Delta c(NADH)}{\Delta t}\right)_{corrected} = \left(\frac{\Delta c(NADH)}{\Delta t}\right)_{measurement} - \left(\frac{\Delta c(NADH)}{\Delta t}\right)_{Neg.Ctrl}$$

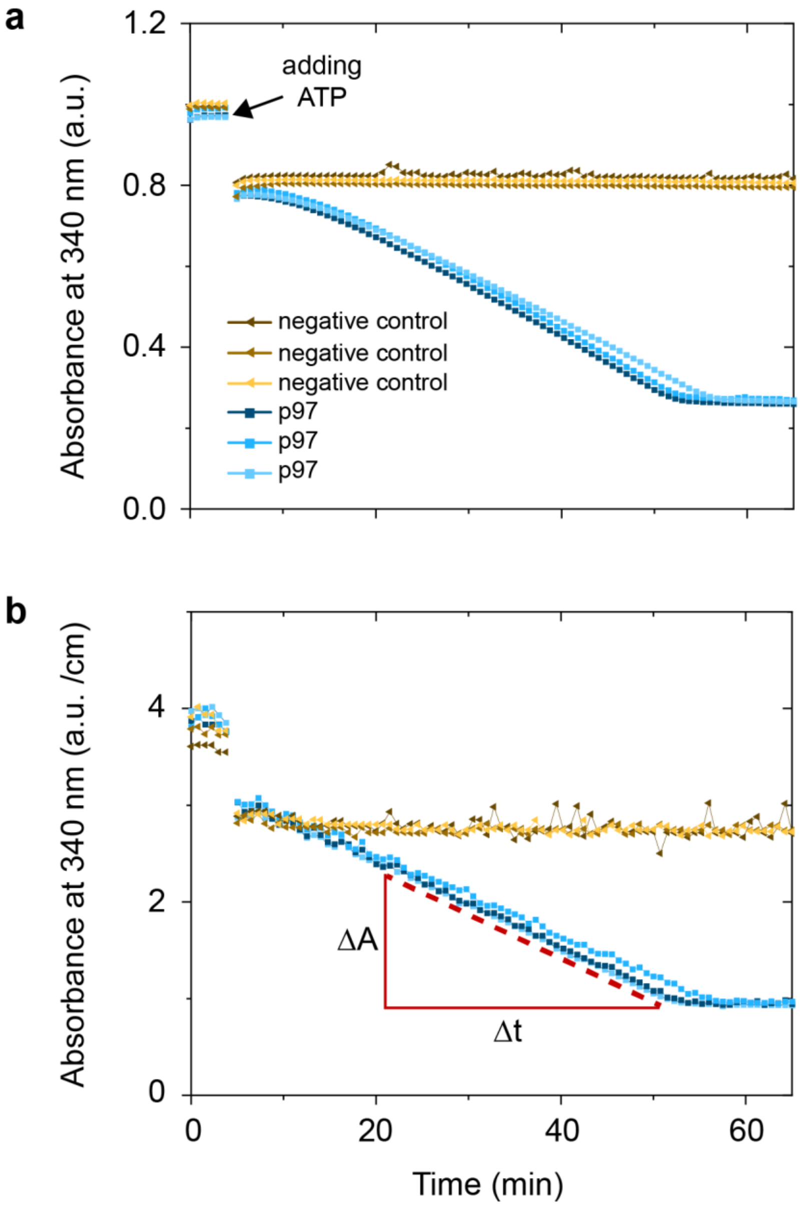


Figure 81. Example of an ATPase assay measurement at 37°C. Raw (**a**) and path length corrected (**b**) absorbances at 340 nm are presented. Curves of p97 containing samples are shown in shades of blue, while those of the negative control are shown in shades of brown. In the first 4 min, the samples are equilibrated. After the addition of ATP, the absorbance drops significantly in all samples. Afterwards, p97 containing samples show a fast decrease in absorbance over time, while the negative controls barely show any changes in absorbance. The non-linear decrease in absorbance in the first ten minutes after the addition of ATP is ascribed to the ongoing equilibration of the samples, now including ATP. After approximately 55 min, NADH is depleted, and a plateau is reached in the samples containing p97. Thus, only the region between 20 and 50 min can be used for linear fitting (indicated by a dashed line in red).

Since the consumption of NADH is directly proportional to the hydrolysis of ATP in a 1:1 ratio, the equation can be changed to

$$\left(\frac{\Delta c(NADH)}{\Delta t}\right)_{corrected} = \frac{\Delta c(ATP)}{\Delta t}$$

Finally, to calculate the p97 ATPase rate, the rate of ATP hydrolysis needed to be normalised to the actual p97 concentration ($c(p97)$) in the well:

$$r(ATPase) = \frac{\Delta c(ATP)}{\Delta t \cdot c(p97)}$$

6.2.3 Intersubunit crosslinking

Neighbouring subunits of p97 were crosslinked using the sulphhydryl-specific reagent *bis*-maleimidoethane (BMOE). BMOE is a homobifunctional maleimide crosslinker with a spacer arm length of 8 Å that forms a non-cleavable thioether linkage to cysteines (according to manufacturer’s instructions).

For the linkage, two point mutations (F360C and A413C) were introduced into ND1L-p97. These residues are located close to the interprotomer surface. On the *cis*-acting protomer side, A413 is located at the end of the helix $\alpha_{407\text{-}423}$, while on the *trans*-acting protomer side, F360 is a part of the arginine finger loop. The distance between the corresponding Cα atoms in different nucleotide states varies between 8.6 Å (ADP state, PDB ID: 5ftk)) and 9.1 Å (ADP·P_i state, PDB ID: 8ooi), which is significantly larger than the typical 4.5-5.5 Å Cα-Cα distance of the naturally occurring disulphide bonds [135]. For that reason, disulphide bond formation between these two mutant cysteines is unfavourable or might lead to strain in both protomers. Instead, crosslinking with BMOE can provide the necessary length to gap the distance. However, all native cysteines had to be removed through point mutations, to prevent undesired inter- and intraprotomer side reactions, creating the p97-ND1L-ΔCys-F360C-A413C construct, with ΔCys = C69V-C77V-C105A-C174A-C184W-C209V-C415A.

For crosslinking, p97 was apyrase digested to the apo state. Then, the sample was fully reduced by incubating the protein with 2 mM DTT for 2 h at 30°C. The reducing agent was removed by SEC on the Superdex 200 Increase using the p97 crosslinking buffer (20 mM HEPES pH 7.2, 250 mM KCl, 5 mM EDTA). Fractions eluting as hexamers were pooled and diluted to

20-60 µM. BMOE was then added in twofold excess from a 20 mM stock in DMSO (stored at -80°C), and the solution was incubated on ice for 2 h. The crosslinking reaction was quenched by adding 50 mM DTT and a further 15 min incubation period on ice. A second SEC run was performed under the same conditions to remove excess crosslinking reagent and to select the crosslinked hexamer fraction. Crosslinking was confirmed by reducing SDS-PAGE.

Estimation of the oligomeric state

Assessment of the oligomeric state of p97 constructs was performed via SEC runs of a Superdex 200 Increase using the p97 SEC buffer. The molecular weight of eluted species was estimated using an external column size calibration with known molecular weight standards (see Supplementary Figure 2).

6.3 NMR spectroscopy

6.3.1 Sample preparation

Solution-state NMR

Perdeuterated *proR*-ILVM-labelled p97 was buffer exchanged into the deuterated p97 solution-state NMR buffer (25 mM HEPES pH 7.5, 25 mM NaCl, 5 mM TCEP, 100 % D_2O) using ultracentrifugation concentration devices (30 or 100 kDa cutoff). The samples were concentrated to volumes of 50-130 uL such that after the addition of further optional components, the final concentration would be in the range of 50-250 µM. All reagents were added from stock solutions, which are listed in Tables 16 and 17. Stocks were prepared by dissolving the respective powders in the p97 solution-state NMR buffer and adjusting the pH to approximately 7.5 with NaOD. All stocks were stored at -20°C.

While apo state p97 was measured without the addition of any compounds, static nucleotide states were assessed by adding 5 mM of ATPγS, AMP-PNP or ADP. In the case of the first two nucleotide states, 4 mM of $MgCl_2$ was supplied as well. For ion titrations, apo state p97 samples were supplied with inorganic phosphate or phosphate mimics (arsenate, sulphate) and 4 mM $MgCl_2$.

Measurements of the ADP·P_i state were conducted using an ATP regeneration system following the protocol of Rydzek *et al.* [55]. Briefly, while ATP is hydrolysed to ADP by p97, pyruvate kinase is continuously regenerating ATP from ADP while consuming phosphoenolpyruvate (PEP). The amount of PK and PEP was varied depending on the expected ATPase rate of the p97 construct. Higher initial concentrations of PEP allowed for longer measurements under constant levels of ATP. The upper limit of PEP was dictated by the inhibition of PK by inorganic phosphate, which is constantly released during ATP hydrolysis [108,109]. Further reagents (ribose-5-phosphate, KCl) were supplied to increase the reaction rate of PK. A typical setup of the ADP·P_i state measurement is presented in Table 17.

Table 16. Reagent and stock solutions used in solution-state NMR experiments of reference states. All stocks were prepared in the p97 solution-state NMR buffer, and the pH was adjusted to 7.5.

Compound	Final concentration	Stock concentration
ATPγS	5 mM or 10 mM	100 mM
ADP	5 mM or 10 mM	500 mM
AMP-PNP	5 mM	100 mM
$MgCl_2$	4 mM	200 mM
Na_2HPO_4	10-300 mM	800 mM
Na_2HAsO_4	100 mM	300 mM
Na_2SO_4	100 mM	1 M

Table 17. Reagent and stock solutions used in solution-state NMR experiments under nucleotide turnover. Stocks were prepared in the p97 solution-state NMR buffer, and the pH was adjusted to 7.5.

Compound	Final concentration	Stock concentration
ATP/*Sp*-ATPαS[a]/ *Sp*-ADPβS[a]	5 mM/5 mM/2 mM	100 mM/60 mM/10 mM
$MgCl_2$	4 mM	200 mM
PEP	30-200 mM	1.42 M
Ribose-5-phosphate	4 mM	322 mM
KCl	50 mM	1 M
Pyruvate kinase (from *Bac. stearothermophilus*)	10-50 units	1 unit/µL

[a]: Stocks were purchased in pure H_2O and the stock volume reduced via lyophilisation.

All solution-state NMR samples were transferred into 3 mm NMR tubes for measurements.

Solid-state NMR

For solid-state experiments, p97 constructs with low ATPase rates were chosen. Non-labelled ND1L-E305Q or FL-E305Q-E578Q constructs were buffer exchanged into the protonated p97 solid-state NMR buffer (25 mM HEPES pH 7.4, 50 mM KCl, 5 mM TCEP, 1 mM DSS) and concentrated to volumes of 100-1000 µL. The experimental setup followed the same principles as with solution-state samples with minor differences. Most significantly, nucleotide concentration in the solution was increased to 10 mM to account for the higher protein concentrations in the rotor [8]. Additionally, rabbit muscle pyruvate kinase was used instead of the PK from *Bac. stearothermophilus* to increase the ATP regeneration rate at the low measurement temperatures. Finally, reagent stocks were preferably prepared in H_2O. Table 18 shows the typical compound concentration used in solid-state NMR measurements.

Sp-ATPβS was regenerated from *Sp*-ADPβS following the same buffer and compound composition (Table 18) for a total duration of two days at 4°C and at room temperature. Since the formation of *Sp*-ATPβS appears to inhibit pyruvate kinase activity, the reaction mixture was diluted with H_2O to a concentration of 2 mM *Sp*-ATPβS after the first day and from then on incubated at room temperature. Afterwards, the total volume was reduced via lyophilisation.

Table 18. Reagent and stock solutions used in solid-state NMR experiments. Stocks were prepared in the p97 solution-state NMR buffer, and the pH was adjusted to 7.5. Only one type of nucleotides was used per experiment.

Compound	Final concentration	Stock concentration
ATP[a]/ATPγS	10 mM	100 mM
Sp-ATPαS[a]/*Sp*-ATPβS[a]	10 mM	60 mM/52 mM
$MgCl_2$[a]	4 mM	1 M
PEP	30-150 mM	1.42 M
Ribose-5-phosphate[a]	4 mM	322 mM
KCl	50 mM	1 M
Pyruvate kinase (from rabbit muscle)[a]	20-80 units	1 unit/µL
DSS	1 mM	100 mM

[a]: Stocks were prepared in pure H_2O.

Protein samples were sedimented overnight into 1.3 mm or 3.2 mm MAS rotors using commercial filling tools under ultracentrifugation at 130,000 g. 3-5 mg of protein were required to fill a 1.3 mm rotor using additional sealing plugs inside the rotor. To fill a 3.2 mm rotor, at least 30 mg of protein were necessary. Due to size differences, full-length p97 sedimented faster, forming a viscous protein gel, while complete ND1L sedimentation often required 20-48 h and often produced a more liquid-like gel.

6.3.2 NMR measurements

Solution-state NMR spectroscopy

Solution-state NMR spectra were acquired on Avance III Bruker spectrometers operating at 600, 800, 900 or 950 MHz equipped with cryo-TCI probes. Temperatures were set to 310 K for apo state measurements and to 313 or 323 K for all other states, depending on mutant stability and the ATPase rate.

2D ^{1}H-^{13}C methyl TROSY spectra were recorded for the assessment of the protein state and 1D ^{1}H NMR spectra for the control of PEP and nucleotides in solution. Acquisition and processing parameters are listed in Tables 19 and 20. 2D spectra were acquired in a time series interlaced with the 1D spectra over a total measurement time of 3 h to 72 h. Identical spectra were added for an improved signal-to-noise ratio.

Table 19. Acquisition parameters of solution-state NMR measurements of p97.

Parameter	**2D ^{1}H-^{13}C methyl TROSY**	**1D ^{1}H**
Pulse sequence	^{1}H-^{13}C HMQC, phase-sensitive, using water suppression with WATERGATE W5 sequence, with ^{1}H decoupling during acquisition [14,148]	1D ^{1}H with water suppression using WATERGATE W5 pulse sequence with gradients [11]
Information gain	Assessment of the protein state	Assessment of the small molecule composition in solution
Size of FID [number of points]	3072 (^{1}H)/512 (^{13}C)	28844
Acquisition time [ms]	100 (^{1}H)/45 (^{13}C)	1000
Spectra width [ppm]	16 (^{1}H)/25 (^{13}C)	16
Transmitter frequency offset [ppm]	4.7 (^{1}H)/18.6 (^{13}C)	4.0/4.7
Interscan delay [s]	1	1
Number of scans	4	128
Total experimental time	45 min	4 min 30 s

Spectra were processed and analysed using Topspin 3.7. ^{1}H chemical shifts of small molecules were assigned based on reported chemical shifts [149] and/or identified experimentally. Methyl peak assignments of TROSY spectra were taken from Schütz *et al*. [66].

Table 20. Processing parameters of solution-state NMR measurements of p97.

Parameter	**2D ^{1}H-^{13}C methyl TROSY**	**1D ^{1}H**
Window function	sine squared	exponential
Window function parameter	cosine	2 Hz
Zero filling	twofold	twofold
Number of FID points used for processing	1536 (^{1}H)/512 (^{13}C)	28844

Solid-state NMR spectroscopy

Solid-state NMR spectra were aquired on an Avance III HD Bruker spectrometer operating at 800 MHz equipped with a wide bore probe in HPC mode. Sample temperatures were set to 5-20°C based on an internal calibration from the water frequency with respect to DSS. MAS rates of 45 kHz were set for 1.3 mm rotors and 17-18 kHz for 3.2 mm rotors.

Two types of pulse sequences were employed: directly pulsed ^{1}H-decoupled 1D ^{31}P NMR for detection of nucleotides in solution, and cross-polarization based ^{1}H-decoupled 1D ^{1}H-^{31}P 1D NMR for characterisation of protein-bound nucleotides. Acquisition and processing parameters are listed in Tables 21 and 22. Both spectra types were acquired in a time series interlaced with the 1D ^{1}H NMR spectra for reference over a total measurement time of 72 h to 160 h. Identical spectra were summed up for an improved signal-to-noise ratio.

Chemical shifts were calibrated internally to DSS. Spectra were processed and analysed using Topspin 3.7.

Table 21. Acquisition parameters of solid-state NMR measurements of p97.

Parameter	**1D ^{31}P**		**1D ^{1}H-^{31}P CP**	
Pulse sequence	1D ^{31}P with low power ^{1}H decoupling		^{1}H-^{31}P cross-polarization based 1D ^{31}P with low power ^{1}H decoupling	
Information gain	Assessment of the small molecule composition in solution		Assessment of the protein-bound nucleotides	
Rotor type	1.3 mm	3.2 mm	1.3 mm	3.2 mm
Size of FID [number of points]	2048-4096	2374	2048	800
Acquisition time [ms]	20-40	20	20	7
Spectra width [ppm]	154	184	154	184
Transmitter frequency offset [ppm]	0 (ATP) or 15 (ATPγP) (^{31}P) / 4.0 (^{1}H)			
Interscan delay [s]	8	3	1	2
Number of scans	256-2048	256-512	8192	4096
Total experimental time	1-4 h	0.5-1 h	2.5 h	
90° excitation pulse [µs]	2.0 (^{31}P)	4.5 (^{31}P)	1.5 (^{1}H)	2.2 (^{1}H)
Cross-polarization transfer			ramp 70-100%, 30 kHz (^{31}P)/ 93 kHz (^{1}H)	ramp 70-100%, 55 kHz (^{31}P)/ 100 kHz (^{1}H)
Cross-polarization contact time [µs]			1500 µs	
^{1}H decoupling during acquisition	slTPPM, 11 kHz	spinal64, 60 kHz	slTPPM, 11 kHz	spinal64, 80 kHz

Table 22. Processing parameters of solid-state NMR measurements of p97.

Parameter	**1D ^{31}P**	**1D ^{1}H-^{31}P CP**
Window function	exponential	
Window function parameter	50-100 Hz	100-200 Hz
Zero filling	twofold	twofold

6.3.3 ^{31}P nucleotide shift measurements

Sample preparation

^{31}P chemical shifts of ATP and ADP in solution were recorded in several sets of experiments. To compare the effects of ionic strength on ^{31}P chemical shifts, two buffers, termed ^{31}P nucleotide shift low and high salt conditions, were chosen (see Table 23). Non-labelled nucleotides were dissolved to a final concentration of 5 mM in the respective buffers with volumes of 500 µL and measured in 5 mm NMR tubes. The experiments were repeated with varying concentrations of $MgCl_2$, ranging from 0 to 10 mM in 2 mM steps.

Table 23. Composition of buffers used for determination of nucleotide shifts. Note the varying concentrations of $MgCl_2$: X = 0, 2, 4, 6, 8, 10 mM.

Buffer identity	Composition
^{31}P nucleotide shift low salt conditions	5 mM nucleotide, 10 mM HEPES, X mM $MgCl_2$, 1 mM DSS, 5% D_2O, pH 7.5
^{31}P nucleotide shift high salt conditions	5 mM nucleotide, 10 mM HEPES, X mM $MgCl_2$, 1 mM DSS, 5% D_2O, 25 mM NaCl, 50 mM KCl, pH 7.5

NMR measurements

NMR spectra were recorded on a Bruker spectrometer operating at 400 MHz equipped with a broadband BBO probe. The temperature was set to 298 K. All spectra were referenced internally to DSS.

Pulse programs used for the determination of chemical shifts, as well as their descriptions and functions, are listed in Table 24. Mainly, direct excitation 1D ^{31}P with proton decoupling was used to identify ^{31}P chemical shifts.

Table 24. Pulse sequences used for determination of nucleotide chemical shifts in solution.

Pulse program	**Description**	**Use**
1D ^{1}H	1D ^{1}H with water suppression using WATERGATE W5 pulse sequence with gradients [11]	Chemical shift referencing to DSS
1D ^{31}P	1D ^{31}P with low power ^{1}H decoupling	Identification of ^{31}P chemical shifts

Acquisition and processing parameters are listed in Tables 25 and 26.

Table 25. Acquisition parameters for measurements of nucleotide chemical shifts.

Parameter	**1D ^{1}H**	**1D ^{31}P**
Spectrometer ^{1}H frequency [MHz]	400	
Probe	BBO	
Acquisition time [ms]	1000	125
Spectra width [ppm]	16	100
Transmitter frequency offset [ppm]	4.7	-5
Interscan delay [s]	1	5
Number of scans	16	128
Total experimental time	30 s	11 min

Spectra were processed and analysed using Topspin 4.3. Peaks were assigned based on reported ^{31}P chemical shifts [89].

Table 26. Processing parameters for measurements of nucleotide chemical shifts.

Parameter	**1D ^{1}H**	**1D ^{31}P**
Window function	exponential	exponential
Window function parameter	2 Hz	5 Hz
Zero filling	twofold	fourfold

Conclusion and outlook

In this thesis, we studied the AAA+ ATPase p97 to elucidate the characteristics of the post-hydrolysis ADP·P_i state.

We employed solution and solid-state NMR techniques in combination with biophysical and biochemical methods and were able to answer key questions regarding this state.

We find that the ADP·P_i state is a symmetric, long-lived protein state with phosphate dissociation being the rate-determining step in the ATPase cycle. Among all hexamers, the nucleotides are bound symmetrically, but the cleaved phosphate is found in several heterogeneous environments. Aside from two stable conformations, phosphate ions experience additional protein interactions when they are transported out of the binding pocket through a positively charged channel towards the central pore. We find that this process is coupled to secondary structure rearrangements in the sensor loop – a short linker region directly connecting neighbouring nucleotide binding pockets. We argue that this sensor loop plays a key role in intersubunit communication and transduces information regarding phosphate release from one protomer to another.

Using mutational analysis, we investigated the contributions of several conserved motif residues to the stability of the ADP·P_i state. We analysed the mutant characteristics following a defined pipeline, assessing potential defects in each step. We found that most mutants were still able to form the ADP·P_i state, but usually with reduced P_i release rates.

Finally, we tested several ATP analogues in regard to their ability to mimic the native nucleotides. While AMP-PCP, *Sp*-ATPαS and ATPβS turned out to be poor ATP mimics, with p97 displaying non-native behaviour upon binding of these analogues, ATPγS and AMP-PNP were found to be good ATP substitutes to stabilise the pre-hydrolysis state of p97.

Overall, we believe that this study helped to elucidate the characteristics of the native post-hydrolysis ADP·P_i state. Still, several questions remain open. For once, it is unknown whether the NTD-down movement and ATP hydrolysis occur simultaneously or successively during the ATPase cycle, and in the latter case which one triggers the other. Furthermore, it remains unclear whether the NTD-down movement is a prerequisite for nucleotide hydrolysis or whether these two processes can occur independently.

Although we could induce an ADP·P_i-like state in ND1L-p97 through incubation of the apo protein with inorganic phosphate, full-length p97 is less willing to react in the same manner. Moreover, the ADP·P_i state cannot be induced starting from ADP-bound ND1L. The reason for these observations remains ambiguous, but it hints at additional communication pathways which remain to be explored.

Our overall approach should be of value for other biological systems as well. Considering that the ADP·P_i state of p97 was only observed under active ATP turnover, we believe that studies of other proteins could reveal new, previously hidden or unobservable protein and nucleotide states, if they were investigated in the presence of a regeneration system. This is especially true for systems with a short-lived, unstable pre-hydrolysis state or where the imitation of the cellular environment (i.e. high ATP, but low ADP concentrations) is required.

Furthermore, our work could help to position the ADP·P_i state as a central intermediate in the ATP hydrolysis cycle, motivating studies of the ADP·P_i state of other ATPases or even expand its relevance to NTPases in general. Conversely, previously published nucleotide structures could be retroactively re-categorised as ADP·P_i states.

Other biological systems could also benefit from the accumulated experience regarding nucleotide analogues, both from the side of biological relevance and practical usefulness.

Finally, we emphasise the contribution of NMR to this investigation. In a collaborative effort, the p97 ADP·P_i state was studied by cryo-EM, MD and NMR, each method contributing valuable insight into the intricacies of this structural state on an atomic level. The major contribution of NMR to this project was to observe the heterogeneous nature of the nucleotide binding pocket, which likely would have been overlooked in cryo-EM studies. This highlights the value of NMR for structural biology.

Acknowledgement

This doctoral thesis would not have been possible without the contributions of numerous people directly or indirectly involved.
First and foremost, I would like to express my thanks to Prof. Dr Anne K. Schütz for giving me the opportunity to work in her group. These past years have seen their fair share of successes and hardships and several project changes. Yet, I was able to reach the end of this journey thanks to her close guidance, ultrafast responses and a keen eye for the essentials.

I wish to thank examination committee members Prof. Dr Martin Zacharias and Prof. Dr Bernd Reif for their valuable time and contribution.

I am grateful to the Technical University of Munich, the Helmholtz Munich and the Ludwig Maximilian University of Munich for providing access to NMR spectrometers and laboratories. On that note, I would like to give my thanks to the head of the Bavarian NMR Center, Prof. Dr Michael Sattler.

No science institute runs on its own. My thanks go to Dr Riddhiman Sarkar, PD Dr Gerd Gemmecker, Dr Sam Asami, Dr Matthias Brandl and Dr Julian Holzinger for the theoretical and practical support regarding NMR and spectrometers in general; to Gülden Yilmaz and Dr Franziska von Hammerstein for keeping the labs organised and running; to Dr Winfried Meining, Christoph Singer and Bernhard Veith for the IT support; and to Karen Biniossek and Waltraud Wolfson for the help with paperwork.

I would like to acknowledge all my collaborators for sticking with me through both easy and challenging projects. My appreciation goes to Dr Carsten Peters and Dr Christoph Kaiser for the great discussions and all the help surrounding cryo-EM and negative stain TEM.

I wish to express my thanks to the Sattler lab for taking me in and teaching me the basics of protein biochemistry. Similarly, my thanks go to the members of Reif and Hagn labs for all the help, discussions and helpful discussions throughout the years.

I am indebted to all the colleagues directly or indirectly involved in my PhD projects. Specifically, I want to thank Dr Pavlo Bielytskyi, Simon Rydzek, Kira Leitl and Maxim Drömer for the great teamwork. Of course, I am equally grateful to the rest of the Schütz lab. This journey would have been so much harder without Zahra Harati Taji, Alexis Clavier, Dr Jelena Habjanic, Zeynab Rezaei, Martina Stobbe and Linda Nguyen. A special thanks to my interns Tim, Adrian, Florian, and Niels for their contribution to my projects.

Finally, I want to thank the people who truly made me who I am today – my parents and siblings. Without their support and encouragement, I wouldn’t have made it this far. Last but not least, my endless love goes to Katya for always having my back, believing in me, and sharing the highs and lows of the past years.

List of abbreviations

Abbreviation	Full name
ADP	Adenosine diphosphate
AIEX	Anion exchange chromatography
AMP	Adenosine monophosphate
AMP-PNP	Adenosine-5'-[(β,γ)-imido]triphosphate
APBS	Adaptive Poisson-Boltzmann Solver
ATP	Adenosine triphosphate
ATPγS	Adenosine-5'-(γ-thio)-triphosphate
BME	2-mercaptoethanol
BMOE	*Bis*-(maleimido)ethane
Cryo-EM	Cryogenic electron microscopy
CSA	Chemical shift anisotropy
CSP	Chemical shift perturbation
CV	Column volume
DMSO	Dimethyl sulphoxide
DSS	Sodium 3-(trimethylsilyl)propane-1-sulphonate
DTT	DL-Dithiothreitol
E.coli	Escherichia coli
EDTA	Ethylene diamine tetraacetic acid
FID	Free induction decay
FPLC	Fast protein liquid chromatography
HEPES	2-[4-(2-Hydroxyethyl)piperazin-1-yl]ethane-1-sulphonic acid
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ISS	Intersubunit signalling
ITC	Isothermal titration calorimetry
LB	Lysogeny broth
LDH	L-Lactate dehydrogenase
MAS	Magic angle spinning
MD	Molecular dynamics
NAD^+	Nicotinamide adenine dinucleotide, oxidised
NADH	Nicotinamide adenine dinucleotide, reduced

Ni-NTA	Nickel nitriloacetic acid
NMR	Nuclear magnetic resonance
OD	Optical density
PEP	Phosphoenolpyruvate
PK	Pyruvate kinase
R5P	Ribose-5-phosphate
SEC	Size exclusion chromatography
TB	Terrific broth
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco Etch virus
TROSY	Transverse relaxation optimized spectroscopy
TSA	Thermal shift assay

Bibliography

1. Reif B, Ashbrook SE, Emsley L, Hong M. Solid-state NMR spectroscopy. Nat Rev Methods Prim. 2021;1. doi:10.1038/s43586-020-00002-1
2. Alderson TR, Kay LE. NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. Cell. 2021;184: 577–595. doi:10.1016/j.cell.2020.12.034
3. Harris RK, Becker ED, Cabral De Menezes SM, Goodfellow R, Granger P. NMR nomenclature: Nuclear spin properties and conventions for chemical shifts (IUPAC recommendations 2001). Concepts Magn Reson Part A Bridg Educ Res. 2002;14: 326–346. doi:10.1002/cmr.10035
4. Koole NJ, De Koning AJ, De Bie MJA. 31P relaxation in some organophosphorus compounds. J Magn Reson. 1977;25: 375–378. doi:10.1016/0022-2364(77)90033-6
5. Becker ED, Shoup RR. 13C nmr spectroscopy: relaxation times of 13C and methods for sensitivity enhancement. Pure Appl Chem. 1972;32: 51–66. doi:10.1351/pac197232010051
6. Cavanagh J, Fairbrother WJ, Palmer AG, Rance M, Skelton NJ. Protein NMR Spectroscopy. 2nd ed. Amsterdam: Elsevier; 2007. doi:10.1016/B978-0-12-164491-8.X5000-3
7. Campbell ID. The evolution of protein NMR. Biomed Spectrosc Imaging. 2013;2: 245–264. doi:10.3233/BSI-130055
8. Bertini I, Luchinat C, Parigi G, Ravera E. SedNMR: On the edge between solution and solid-state NMR. Acc Chem Res. 2013;46: 2059–2069. doi:10.1021/ar300342f
9. Duer MJ. Introduction to Solid-State NMR Spectroscopy. 1st ed. Oxford: Blackwell Science; 2004.
10. Hong M, Schmidt-Rohr K, Nanz D. Study of phospholipid structure by 1H, 13C, and 31P dipolar couplings from two-dimensional NMR. Biophys J. 1995;69: 1939–1950. doi:10.1016/S0006-3495(95)80064-0
11. Liu M, Mao XA, Ye C, Huang H, Nicholson JK, Lindon JC. Improved Watergate Pulse Sequences for Solvent Suppression in NMR Spectroscopy. J Magn Reson. 1998;132: 125–129. doi:10.1006/jmre.1998.1405
12. Schütz S, Sprangers R. Methyl TROSY spectroscopy: A versatile NMR approach to study challenging biological systems. Prog Nucl Magn Reson Spectrosc. 2020;116: 56–84. doi:10.1016/j.pnmrs.2019.09.004

13. Pervushin K, Riek R, Wider G, Wüthrich K. Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy. Proceedings of the National Academy of Sciences of the United States of America. 1997. pp. 12366–12371.
14. Ollerenshaw JE, Tugarinov V, Kay LE. Methyl TROSY: Explanation and experimental verification. Magn Reson Chem. 2003;41: 843–852. doi:10.1002/mrc.1256
15. Janin J, Miller S, Chothia C. Surface, subunit interfaces and interior of oligomeric proteins. J Mol Biol. 1988;204: 155–164. doi:10.1016/0022-2836(88)90606-7
16. Saraste M, Sibbald PR, Wittinghofer A. The P-loop — a common motif in ATP- and GTP-binding proteins. Trends Biochem Sci. 1990;15: 430–434. doi:10.1016/0968-0004(90)90281-F
17. Longo LM, Jabłońska J, Vyas P, Kanade M, Kolodny R, Ben-Tal N, et al. On the emergence of P-Loop NTPase and Rossmann enzymes from a Beta-Alpha-Beta ancestral fragment. Elife. 2020;9: 1–16. doi:10.7554/eLife.64415
18. Erzberger JP, Berger JM. EVOLUTIONARY RELATIONSHIPS AND STRUCTURAL MECHANISMS OF AAA+ PROTEINS. Annu Rev Biophys Biomol Struct. 2006;35: 93–114. doi:10.1146/annurev.biophys.35.040405.101933
19. Seraphim T V., Houry WA. AAA+ proteins. Curr Biol. 2020;30: R251–R257. doi:10.1016/j.cub.2020.01.044
20. Khan YA, White KI, Brunger AT. The AAA+ superfamily: a review of the structural and mechanistic principles of these molecular machines. Crit Rev Biochem Mol Biol. 2022;57: 156–187. doi:10.1080/10409238.2021.1979460
21. Frickey T, Lupas AN. Phylogenetic analysis of AAA proteins. J Struct Biol. 2004;146: 2–10. doi:10.1016/j.jsb.2003.11.020
22. Puchades C, Sandate CR, Lander GC. The molecular principles governing the activity and functional diversity of AAA+ proteins. Nat Rev Mol Cell Biol. 2020;21: 43–58. doi:10.1038/s41580-019-0183-6
23. Barthelme D, Sauer RT. Origin and Functional Evolution of the Cdc48/p97/VCP AAA+ Protein Unfolding and Remodeling Machine. J Mol Biol. 2016;428: 1861–1869. doi:10.1016/j.jmb.2015.11.015
24. Valimehr S, Sethi A, Shukla M, Bhattacharyya S, Kazemi M, Rouiller I. Molecular Mechanisms Driving and Regulating the AAA+ ATPase VCP/p97, an Important Therapeutic Target for Treating Cancer, Neurological and Infectious Diseases. Biomolecules. 2023;13: 737. doi:10.3390/biom13050737

25. Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, Szymborska A, et al. The quantitative proteome of a human cell line. Mol Syst Biol. 2011;7: 1–8. doi:10.1038/msb.2011.82
26. Chu S, Xie X, Payan C, Stochaj U. Valosin containing protein (VCP): initiator, modifier, and potential drug target for neurodegenerative diseases. Mol Neurodegener. 2023;18: 52. doi:10.1186/s13024-023-00639-y
27. Ye Y, Tang WK, Zhang T, Xia D. A Mighty “Protein Extractor” of the Cell: Structure and Function of the p97/CDC48 ATPase. Front Mol Biosci. 2017;4: 1–20. doi:10.3389/fmolb.2017.00039
28. Braxton JR, Southworth DR. Structural insights of the p97/VCP AAA+ ATPase: How adapter interactions coordinate diverse cellular functionality. J Biol Chem. 2023;299: 105182. doi:10.1016/j.jbc.2023.105182
29. Fessart D, Marza E, Taouji S, Delom F, Chevet E. P97/CDC-48: Proteostasis control in tumor cell biology. Cancer Lett. 2013;337: 26–34. doi:10.1016/j.canlet.2013.05.030
30. Vekaria PH, Home T, Weir S, Schoenen FJ, Rao R. Targeting p97 to Disrupt Protein Homeostasis in Cancer. Front Oncol. 2016;6: 1–8. doi:10.3389/fonc.2016.00181
31. Pöhler R, Krahn JH, van den Boom J, Dobrynin G, Kaschani F, Eggenweiler H, et al. A Non-Competitive Inhibitor of VCP/p97 and VPS4 Reveals Conserved Allosteric Circuits in Type I and II AAA ATPases. Angew Chemie Int Ed. 2018;57: 1576–1580. doi:10.1002/anie.201711429
32. Banerjee S, Bartesaghi A, Merk A, Rao P, Bulfer SL, Yan Y, et al. 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. Science. 2016;351: 871–875. doi:10.1126/science.aad7974
33. Stach L, Morgan RM, Makhlouf L, Douangamath A, von Delft F, Zhang X, et al. Crystal structure of the catalytic D2 domain of the AAA+ ATPase p97 reveals a putative helical split-washer-type mechanism for substrate unfolding. FEBS Lett. 2020;594: 933–943. doi:10.1002/1873-3468.13667
34. Koller KJ, Brownstein MJ. Use of a cDNA clone to identify a supposed precursor protein containing valosin. Nature. 1987;325: 542–545. doi:10.1038/325542a0
35. Briggs LC, Baldwin GS, Miyata N, Kondo H, Zhang X, Freemont PS. Analysis of Nucleotide Binding to P97 Reveals the Properties of a Tandem AAA Hexameric ATPase. J Biol Chem. 2008;283: 13745–13752. doi:10.1074/jbc.M709632200
36. Pan M, Yu Y, Ai H, Zheng Q, Xie Y, Liu L, et al. Mechanistic insight into substrate processing and allosteric inhibition of human p97. Nat Struct Mol Biol. 2021;28: 614–

625. doi:10.1038/s41594-021-00617-2

37. Nandi P, Li S, Columbres RCA, Wang F, Williams DR, Poh Y-P, et al. Structural and Functional Analysis of Disease-Linked p97 ATPase Mutant Complexes. Int J Mol Sci. 2021;22: 8079. doi:10.3390/ijms22158079

38. Hänzelmann P, Schindelin H. Characterization of an Additional Binding Surface on the p97 N-Terminal Domain Involved in Bipartite Cofactor Interactions. Structure. 2016;24: 140–147. doi:10.1016/j.str.2015.10.027

39. Xia D, Tang WK, Ye Y. Structure and function of the AAA+ ATPase p97/Cdc48p. Gene. 2016;583: 64–77. doi:10.1016/j.gene.2016.02.042

40. Wang Q, Song C, Li C-CH. Hexamerization of p97-VCP is promoted by ATP binding to the D1 domain and required for ATPase and biological activities. Biochem Biophys Res Commun. 2003;300: 253–260. doi:10.1016/S0006-291X(02)02840-1

41. Hänzelmann P, Schindelin H. Structural Basis of ATP Hydrolysis and Intersubunit Signaling in the AAA+ ATPase p97. Structure. 2016;24: 127–139. doi:10.1016/j.str.2015.10.026

42. Yu G, Bai Y, Li K, Amarasinghe O, Jiang W, Zhang Z-Y. Cryo-electron microscopy structures of VCP/p97 reveal a new mechanism of oligomerization regulation. iScience. 2021;24: 103310. doi:10.1016/j.isci.2021.103310

43. Gao H, Li F, Ji Z, Shi Z, Li Y, Yu H. Cryo-EM structures of human p97 double hexamer capture potentiated ATPase-competent state. Cell Discov. 2022;8: 19. doi:10.1038/s41421-022-00379-1

44. Niwa H, Ewens CA, Tsang C, Yeung HO, Zhang X, Freemont PS. The Role of the N-Domain in the ATPase Activity of the Mammalian AAA ATPase p97/VCP. J Biol Chem. 2012;287: 8561–8570. doi:10.1074/jbc.M111.302778

45. Li G, Huang C, Zhao G, Lennarz WJ, Ennarz WJL. Interprotomer motion-transmission mechanism for the hexameric AAA ATPase p97. Proc Natl Acad Sci USA. 2012;109: 3737–3741. doi:10.1073/pnas.1200255109

46. Huang C, Li G, Lennarz WJ. Dynamic flexibility of the ATPase p97 is important for its interprotomer motion transmission. Proc Natl Acad Sci USA. 2012;109: 9792–9797. doi:10.1073/pnas.1205853109

47. Chou T-F, Bulfer SL, Weihl CC, Li K, Lis LG, Walters MA, et al. Specific Inhibition of p97/VCP ATPase and Kinetic Analysis Demonstrate Interaction between D1 and D2 ATPase Domains. J Mol Biol. 2014;426: 2886–2899. doi:10.1016/j.jmb.2014.05.022

48. Walker JE, Saraste M, Runswick MJ, Gay NJ. 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. 1982;1: 945–951. doi:10.1002/j.1460-2075.1982.tb01276.x

49. Wendler P, Ciniawsky S, Kock M, Kube S. Structure and function of the AAA+ nucleotide binding pocket. Biochim Biophys Acta - Mol Cell Res. 2012;1823: 2–14. doi:10.1016/j.bbamcr.2011.06.014
50. Gates SN, Martin A. Stairway to translocation: AAA+ motor structures reveal the mechanisms of ATP-dependent substrate translocation. Protein Sci. 2020;29: 407–419. doi:10.1002/pro.3743
51. delToro D, Ortiz D, Ordyan M, Sippy J, Oh C-S, Keller N, et al. Walker-A Motif Acts to Coordinate ATP Hydrolysis with Motor Output in Viral DNA Packaging. J Mol Biol. 2016;428: 2709–2729. doi:10.1016/j.jmb.2016.04.029
52. Shein M, Hitzenberger M, Cheng TC, Rout SR, Leitl KD, Sato Y, et al. Characterizing ATP processing by the AAA+ protein p97 at the atomic level. Nat Chem. 2024;16: 363–372. doi:10.1038/s41557-024-01440-0
53. Zhang S, Mao Y. AAA+ ATPases in Protein Degradation: Structures, Functions and Mechanisms. Biomolecules. 2020;10: 629. doi:10.3390/biom10040629
54. Tang WK, Li D, Li C, Esser L, Dai R, Guo L, et al. A novel ATP-dependent conformation in p97 N–D1 fragment revealed by crystal structures of disease-related mutants. EMBO J. 2010;29: 2217–2229. doi:10.1038/emboj.2010.104
55. Rydzek S, Shein M, Bielytskyi P, Schütz AK. Observation of a Transient Reaction Intermediate Illuminates the Mechanochemical Cycle of the AAA-ATPase p97. J Am Chem Soc. 2020;142: 14472–14480. doi:10.1021/jacs.0c03180
56. Song C, Wang Q, Li C-CH. ATPase Activity of p97-Valosin-containing Protein (VCP). J Biol Chem. 2003;278: 3648–3655. doi:10.1074/jbc.M208422200
57. Zhang X, Wigley DB. The “glutamate switch” provides a link between ATPase activity and ligand binding in AAA+ proteins. Nat Struct Mol Biol. 2008;15: 1223–1227. doi:10.1038/nsmb.1501
58. Neuwald AF, Aravind L, Spouge JL, Koonin E V. AAA + : A Class of Chaperone-Like ATPases Associated with the Assembly, Operation, and Disassembly of Protein Complexes. Genome Res. 1999;9: 27–43. doi:10.1101/gr.9.1.27
59. Zhang X, Shaw A, Bates PA, Newman RH, Gowen B, Orlova E, et al. Structure of the AAA ATPase p97. Mol Cell. 2000;6: 1473–1484. doi:10.1016/S1097-2765(00)00143-X

60. Gerega A, Rockel B, Peters J, Tamura T, Baumeister W, Zwickl P. VAT, the Thermoplasma Homolog of Mammalian p97/VCP, Is an N Domain-regulated Protein Unfoldase. J Biol Chem. 2005;280: 42856–42862. doi:10.1074/jbc.M510592200

61. Ogura T, Whiteheart SW, Wilkinson AJ. Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. J Struct Biol. 2004;146: 106–112. doi:10.1016/j.jsb.2003.11.008

62. Augustin S, Gerdes F, Lee S, Tsai FTF, Langer T, Tatsuta T. An Intersubunit Signaling Network Coordinates ATP Hydrolysis by m-AAA Proteases. Mol Cell. 2009;35: 574–585. doi:10.1016/j.molcel.2009.07.018

63. Tang WK, Xia D. Altered Intersubunit Communication Is the Molecular Basis for Functional Defects of Pathogenic p97 Mutants. J Biol Chem. 2013;288: 36624–36635. doi:10.1074/jbc.M113.488924

64. Guedes Aguiar B, Padmanabhan PK, Dumas C, Papadopoulou B. Valosin-containing protein VCP/p97 is essential for the intracellular development of Leishmania and its survival under heat stress. Cell Microbiol. 2018;20: e12867. doi:10.1111/cmi.12867

65. Chia WS, Chia DX, Rao F, Bar Nun S, Geifman Shochat S. ATP Binding to p97/VCP D1 Domain Regulates Selective Recruitment of Adaptors to Its Proximal N-Domain. Kursula I, editor. PLoS One. 2012;7: e50490. doi:10.1371/journal.pone.0050490

66. Schuetz AK, Kay LE. A Dynamic molecular basis for malfunction in disease mutants of p97/VCP. Elife. 2016;5: 1–25. doi:10.7554/eLife.20143

67. Noi K, Yamamoto D, Nishikori S, Arita-Morioka K, Kato T, Ando T, et al. High-Speed Atomic Force Microscopic Observation of ATP-Dependent Rotation of the AAA+ Chaperone p97. Structure. 2013;21: 1992–2002. doi:10.1016/j.str.2013.08.017

68. Taylor EW, Lymn RW, Moll G. Myosin-product complex and its effect on the steady-state rate of nucleoside triphosphate hydrolysis. Biochemistry. 1970;9: 2984–2991. doi:10.1021/bi00817a008

69. Flaherty KM, Wilbanks SM, DeLuca-Flaherty C, McKay DB. Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. II. Structure of the active site with ADP or ATP bound to wild type and mutant ATPase fragment. J Biol Chem. 1994;269: 12899–12907. doi:10.1016/S0021-9258(18)99961-8

70. Reynolds MJ, Hachicho C, Carl AG, Gong R, Alushin GM. Bending forces and nucleotide state jointly regulate F-actin structure. Nature. 2022;611: 380–386. doi:10.1038/s41586-022-05366-w

71. Heissler SM, Arora AS, Billington N, Sellers JR, Chinthalapudi K. Cryo-EM structure

of the autoinhibited state of myosin-2. Sci Adv. 2021;7: 1–9. doi:10.1126/sciadv.abk3273

72. Drakou CE, Malekkou A, Hayes JM, Lederer CW, Leonidas DD, Oikonomakos NG, et al. hCINAP is an atypical mammalian nuclear adenylate kinase with an ATPase motif: Structural and functional studies. Proteins Struct Funct Bioinforma. 2012;80: 206–220. doi:10.1002/prot.23186

73. Sobti M, Ueno H, Noji H, Stewart AG. The six steps of the complete F1-ATPase rotary catalytic cycle. Nat Commun. 2021;12: 4690. doi:10.1038/s41467-021-25029-0

74. Faehnle CR, Blanco J, Viola RE. Structural basis for discrimination between oxyanion substrates or inhibitors in aspartate-β-semialdehyde dehydrogenase. Acta Crystallogr Sect D Biol Crystallogr. 2004;60: 2320–2324. doi:10.1107/S0907444904026411

75. Tang WK, Xia D. Role of the D1-D2 Linker of Human VCP/p97 in the Asymmetry and ATPase Activity of the D1-domain. Sci Rep. 2016;6: 20037. doi:10.1038/srep20037

76. Tang WK, Zhang T, Ye Y, Xia D. Structural basis for nucleotide-modulated p97 association with the ER membrane. Cell Discov. 2017;3: 17045. doi:10.1038/celldisc.2017.45

77. Pan M, Zheng Q, Yu Y, Ai H, Xie Y, Zeng X, et al. Seesaw conformations of Npl4 in the human p97 complex and the inhibitory mechanism of a disulfiram derivative. Nat Commun. 2021;12: 121. doi:10.1038/s41467-020-20359-x

78. Szántó JK, Dietschreit JCB, Shein M, Schütz AK, Ochsenfeld C. Systematic QM/MM Study for Predicting 31 P NMR Chemical Shifts of Adenosine Nucleotides in Solution and Stages of ATP Hydrolysis in a Protein Environment. J Chem Theory Comput. 2024;20: 2433–2444. doi:10.1021/acs.jctc.3c01280

79. Lassila JK, Zalatan JG, Herschlag D. Biological phosphoryl-transfer reactions: Understanding mechanism and catalysis. Annual Review of Biochemistry. 2011. doi:10.1146/annurev-biochem-060409-092741

80. Kamerlin SCL, Sharma PK, Prasad RB, Warshel A. Why nature really chose phosphate. Q Rev Biophys. 2013;46: 1–132. doi:10.1017/S0033583512000157

81. Kiani FA, Fischer S. Comparing the catalytic strategy of ATP hydrolysis in biomolecular motors. Phys Chem Chem Phys. 2016;18: 20219–20233. doi:10.1039/c6cp01364c

82. Luo D, Xu T, Watson RP, Scherer-Becker D, Sampath A, Jahnke W, et al. Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. EMBO J. 2008;27:

3209–3219. doi:10.1038/emboj.2008.232

83. Gout E, Rébeillé F, Douce R, Bligny R. Interplay of Mg2+, ADP, and ATP in the cytosol and mitochondria: Unravelling the role of Mg2+ in cell respiration. Proc Natl Acad Sci USA. 2014;111: E4560–E4567. doi:10.1073/pnas.1406251111

84. Greiner J V., Glonek T. Intracellular ATP Concentration and Implication for Cellular Evolution. Biology. 2021;10: 1166. doi:10.3390/biology10111166

85. Clarke K, Kashiwaya Y, Todd King M, Gates D, Keon CA, Cross HR, et al. The β/α peak height ratio of ATP. A measure of free [Mg2+] using 31P NMR. J Biol Chem. 1996;271: 21142–21150. doi:10.1074/jbc.271.35.21142

86. Veech RL, Gates DN, Crutchfield C, Gitomer WL, Kashiwaya Y, King MT, et al. Metabolic Hyperpolarization of Liver by Ethanol: The Importance of Mg2+ and H+ in Determining Impermeant Intracellular Anionic Charge and Energy of Metabolic Reactions. Alcohol Clin Exp Res. 1994;18: 1040–1056. doi:10.1111/j.1530-0277.1994.tb00081.x

87. Yoza N, Ueda N, Nakashima S. pH-dependence of 31P-NMR spectroscopic parameters of monofluorophosphate, phosphate, hypophosphate, phosphonate, phosphinate and their dimers and trimers. Fresenius J Anal Chem. 1994;348: 633–638. doi:10.1007/BF00325563

88. Pecoraro VL, Hermes JD, Cleland WW. Stability constants of magnesium and cadmium complexes of adenine nucleotides and thionucleotides and rate constants for formation and dissociation of magnesium-ATP and magnesium-ADP. Biochemistry. 1984;23: 5262–5271. doi:10.1021/bi00317a026

89. Jaffe EK, Cohn M. 31P Nuclear Magnetic Resonance Spectra of the Thiophosphate Analogues of Adenine Nucleotides; Effects of pH and Mg2+ Binding. Biochemistry. 1978;17: 652–657. doi:10.1021/bi00597a014

90. Glonek T. 31P NMR of Mg-ATP in dilute solutions: Complexation and exchange. Int J Biochem. 1992;24: 1533–1559. doi:10.1016/0020-711X(92)90171-V

91. Roberts JKM, Wade-Jardetzky N, Jardetzky O. Intracellular pH Measurements by 31P Nuclear Magnetic Resonance. Influence of Factors Other Than pH on 31P Chemical Shifts. Biochemistry. 1981;20: 5389–5394. doi:10.1021/bi00522a006

92. Mohan MS, Rechnitz GA. Ion-electrode study of magnesium(II)-ATP and manganese(II)-ATP association. Arch Biochem Biophys. 1974;162: 194–199. doi:10.1016/0003-9861(74)90118-0

93. Kettlun AM, Espinosa V, García L, Valenzuela MA. Potato tuber isoapyrases:

Substrate specificity, affinity labeling, and proteolytic susceptibility. Phytochemistry. 2005;66: 975–982. doi:10.1016/j.phytochem.2005.03.015

94. Buelens FP, Leonov H, De Groot BL, Grubmüller H. ATP-Magnesium Coordination: Protein Structure-Based Force Field Evaluation and Corrections. J Chem Theory Comput. 2021;17: 1922–1930. doi:10.1021/acs.jctc.0c01205

95. Cohn M, Hughes TR. Nuclear Magnetic Resonance Spectra of Adenosine Di- and Triphosphate. J Biol Chem. 1962;237: 176–181. doi:10.1016/s0021-9258(18)81382-5

96. Jaffe EK, Cohn M. Divalent cation-dependent stereospecificity of adenosine 5‘-O-(2-thiotriphosphate) in the hexokinase and pyruvate kinase reactions. The absolute stereochemistry of the diastereoisomers of adenosine 5‘-O-(2-thiotriphosphate). J Biol Chem. 1978;253: 4823–4825. doi:10.1016/S0021-9258(17)34617-3

97. Huang SL, Tsai MD. Does the Magnesium(II) Ion Interact with the α-Phosphate of Adenosine Triphosphate? An Investigation by Oxygen-17 Nuclear Magnetic Resonance. Biochemistry. 1982;21: 951–959. doi:10.1021/bi00534a021

98. Halawani D, LeBlanc AC, Rouiller I, Michnick SW, Servant MJ, Latterich M. Hereditary Inclusion Body Myopathy-Linked p97/VCP Mutations in the NH 2 Domain and the D1 Ring Modulate p97/VCP ATPase Activity and D2 Ring Conformation. Mol Cell Biol. 2009;29: 4484–4494. doi:10.1128/MCB.00252-09

99. Lacabanne D, Wiegand T, Wili N, Kozlova MI, Cadalbert R, Klose D, et al. ATP Analogues for Structural Investigations: Case Studies of a DnaB Helicase and an ABC Transporter. Molecules. 2020;25. doi:10.3390/MOLECULES25225268

100. Iuga A, Spoerner M, Ader C, Brunner E, Kalbitzer HR. Rapid assignment of solution 31P NMR spectra of large proteins by solid-state spectroscopy. Biochem Biophys Res Commun. 2006;346: 301–305. doi:10.1016/j.bbrc.2006.05.116

101. Lacabanne D, Wiegand T, Di Cesare M, Orelle C, Ernst M, Jault JM, et al. Solid-State NMR Reveals Asymmetric ATP Hydrolysis in the Multidrug ABC Transporter BmrA. J Am Chem Soc. 2022;144: 12431–12442. doi:10.1021/jacs.2c04287

102. Twomey EC, Ji Z, Wales TE, Bodnar NO, Ficarro SB, Marto JA, et al. Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. Science. 2019;365. doi:10.1126/science.aax1033

103. Wiegand T, Cadalbert R, Lacabanne D, Timmins J, Terradot L, Böckmann A, et al. The conformational changes coupling ATP hydrolysis and translocation in a bacterial DnaB helicase. Nat Commun. 2019;10: 1–11. doi:10.1038/s41467-018-07968-3

104. Iuga A, Spoerner M, Kalbitzer HR, Brunner E. Solid-state 31P NMR spectroscopy of

microcrystals of the Ras protein and its effector loop mutants: Comparison between crystalline and solution state. J Mol Biol. 2004;342: 1033–1040. doi:10.1016/j.jmb.2004.07.077

105. Wiegand T, Cadalbert R, Gardiennet C, Timmins J, Terradot L, Böckmann A, et al. Monitoring ssDNA Binding to the DnaB Helicase from Helicobacter pylori by Solid-State NMR Spectroscopy. Angew Chemie - Int Ed. 2016;55: 14164–14168. doi:10.1002/anie.201607295

106. Fürtig B, Richter C, Wöhnert J, Schwalbe H. NMR spectroscopy of RNA. ChemBioChem. 2003;4: 936–962. doi:10.1002/cbic.200300700

107. Mayar M, De Roo N, Hoos P, Van Duynhoven J. 31P NMR Quantification of Phospholipids and Lysophospholipids in Food Emulsions. J Agric Food Chem. 2020;68: 5009–5017. doi:10.1021/acs.jafc.0c00404

108. Yamada T, Carlsson J. Glucose 6 phosphate dependent pyruvate kinase in Streptococcus mutans. J Bacteriol. 1975;124: 562–563. doi:10.1128/jb.124.1.562-563.1975

109. Waygood EB, Rayman MK, Sanwal BD. The control of pyruvate kinases of Escherichia coli. II. Effectors and regulatory properties of the enzyme activated by ribose 5 phosphate. Can J Biochem. 1975;53: 444–454. doi:10.1139/o75-061

110. Marchanka A, Simon B, Althoff-Ospelt G, Carlomagno T. RNA structure determination by solid-state NMR spectroscopy. Nat Commun. 2015;6. doi:10.1038/ncomms8024

111. Tsai M-D, Huang SL, Kozlowski JF, Chang CC. Applicability of the phosphorus-31 (oxygen-17) nuclear magnetic resonance method in the study of enzyme mechanism involving phosphorus. Biochemistry. 1980;19: 3531–3536. doi:10.1021/bi00556a018

112. Taji ZH, Bielytskyi P, Shein M, Sani MA, Seitz S, Schütz AK. Transient RNA Interactions Leave a Covalent Imprint on a Viral Capsid Protein. J Am Chem Soc. 2022;144: 8536–8550. doi:10.1021/JACS.1C12439

113. Boudet J, Devillier JC, Wiegand T, Salmon L, Meier BH, Lipps G, et al. A Small Helical Bundle Prepares Primer Synthesis by Binding Two Nucleotides that Enhance Sequence-Specific Recognition of the DNA Template. Cell. 2019;176: 154-166.e13. doi:10.1016/j.cell.2018.11.031

114. Eckstein F. NUCLEOSIDE PHOSPHOROTHIOATES. Annu Rev Biochem. 1985;54: 367–402. doi:10.1146/annurev.bi.54.070185.002055

115. Zhu K, Cai Y, Si X, Ye Z, Gao Y, Liu C, et al. The phosphorylation and

dephosphorylation switch of VCP/p97 regulates the architecture of centrosome and spindle. Cell Death Differ. 2022;29: 2070–2088. doi:10.1038/s41418-022-01000-4

116. Davies JM, Brunger AT, Weis WI. Improved Structures of Full-Length p97, an AAA ATPase: Implications for Mechanisms of Nucleotide-Dependent Conformational Change. Structure. 2008;16: 715–726. doi:10.1016/j.str.2008.02.010

117. Connolly BA, Eckstein F. Structures of the mono- and divalent metal nucleotide complexes in the myosin ATPase. J Biol Chem. 1981;256: 9450–9456. doi:10.1016/s0021-9258(19)68783-1

118. Eckstein F, Goody RS. Synthesis and properties of diastereoisomers of adenosine 5’-(O-1-thiotriphosphate) and adenosine 5’-(O-2-thiotriphosphate). Biochemistry. 1976;15: 1685–1691. doi:10.1021/bi00653a015

119. Jaffe EK, Cohn M. Diastereomers of the nucleoside phosphorothioates as probes of the structure of the metal nucleotide substrates and of the nucleotide binding site of yeast hexokinase. J Biol Chem. 1979;254: 10839–10845. doi:10.1016/S0021-9258(19)86597-3

120. Wennefors CK, Dobrikov MI, Xu Z, Li P, Shaw BR. Stereospecificity, substrate, and inhibitory properties of nucleoside diphosphate analogs for creatine and pyruvate kinases. Bioorg Chem. 2008;36: 169–177. doi:10.1016/j.bioorg.2008.03.001

121. Romaniuk PJ, Eckstein F. Structure of the metal-nucleotide complex in the acetate kinase reaction. A study with gamma-32P-labeled phosphorothioate analogs of ATP. J Biol Chem. 1981;256: 7322–7328. doi:10.1016/s0021-9258(19)68965-9

122. Weiss GA, Watanabe CK, Zhong A, Goddard A, Sidhu SS. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. Proc Natl Acad Sci USA. 2000;97: 8950–8954. doi:10.1073/pnas.160252097

123. Ye Y, Meyer HH, Rapoport TA. Function of the p97–Ufd1–Npl4 complex in retrotranslocation from the ER to the cytosol. J Cell Biol. 2003;162: 71–84. doi:10.1083/jcb.200302169

124. Blythe EE, Olson KC, Chau V, Deshaies RJ. Ubiquitin- and ATP-dependent unfoldase activity of P97/VCP•NPLOC4•UFD1L is enhanced by a mutation that causes multisystem proteinopathy. Proc Natl Acad Sci USA. 2017;114: E4380–E4388. doi:10.1073/pnas.1706205114

125. Nishikori S, Esaki M, Yamanaka K, Sugimoto S, Ogura T. Positive Cooperativity of the p97 AAA ATPase Is Critical for Essential Functions. J Biol Chem. 2011;286: 15815–15820. doi:10.1074/jbc.M110.201400

126. DeLaBarre B, Christianson JC, Kopito RR, Brunger AT. Central Pore Residues Mediate the p97/VCP Activity Required for ERAD. Mol Cell. 2006;22: 451–462. doi:10.1016/j.molcel.2006.03.036

127. Wang Q, Song C, Irizarry L, Dai R, Zhang X, Li C-CH. Multifunctional Roles of the Conserved Arg Residues in the Second Region of Homology of p97/Valosin-containing Protein. J Biol Chem. 2005;280: 40515–40523. doi:10.1074/jbc.M509636200

128. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7: 1–6. doi:10.1038/msb.2011.75

129. Morgan AA, Rubenstein E. Proline: The Distribution, Frequency, Positioning, and Common Functional Roles of Proline and Polyproline Sequences in the Human Proteome. Casarini DE, editor. PLoS One. 2013;8: e53785. doi:10.1371/journal.pone.0053785

130. Wiederstein JL, Nolte H, Günther S, Piller T, Baraldo M, Kostin S, et al. Skeletal Muscle-Specific Methyltransferase METTL21C Trimethylates p97 and Regulates Autophagy-Associated Protein Breakdown. Cell Rep. 2018;23: 1342–1356. doi:10.1016/j.celrep.2018.03.136

131. He C, Chen J, Wang H, Wan Y, Zhou J, Dan Z, et al. Crystal structures of rice hexokinase 6 with a series of substrates shed light on its enzymatic mechanism. Biochem Biophys Res Commun. 2019;515: 614–620. doi:10.1016/j.bbrc.2019.05.139

132. Oosterheert W, Blanc FEC, Roy A, Belyy A, Sanders MB, Hofnagel O, et al. Molecular mechanisms of inorganic-phosphate release from the core and barbed end of actin filaments. Nat Struct Mol Biol. 2023;30: 1774–1785. doi:10.1038/s41594-023-01101-9

133. Fraústo da Silva JJR, Williams RJP. The Biological Chemistry of the Elements: The Inorganic Chemistry of Life. 2nd ed. Oxford: Oxford University Press; 2001. doi:10.1093/oso/9780198508472.003.0001

134. Schütz AK, Rennella E, Kay LE. Exploiting conformational plasticity in the AAA+ protein VCP/p97 to modify function. Proc Natl Acad Sci USA. 2017;114: E6822–E6829. doi:10.1073/pnas.1707974114

135. Katz BA, Kossiakoff A. The crystallographically determined structures of atypical strained disulfides engineered into subtilisin. J Biol Chem. 1986;261: 15480–15485. doi:10.1016/s0021-9258(18)66737-7

136. Heinig M, Frishman D. STRIDE: A web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. 2004;32: 500–502. doi:10.1093/nar/gkh429

137. Hansen DF, Neudecker P, Kay LE. Determination of isoleucine side-chain conformations in ground and excited states of proteins from chemical shifts. J Am Chem Soc. 2010;132: 7589–7591. doi:10.1021/ja102090z

138. Bodnar NO, Rapoport TA. Molecular Mechanism of Substrate Processing by the Cdc48 ATPase Complex. Cell. 2017;169: 722-735.e9. doi:10.1016/j.cell.2017.04.020

139. Park S, Rancour DM, Bednarek SY. Protein Domain-Domain Interactions and Requirements for the Negative Regulation of Arabidopsis CDC48/p97 by the Plant Ubiquitin Regulatory X (UBX) Domain-containing Protein, PUX1. J Biol Chem. 2007;282: 5217–5224. doi:10.1074/jbc.M609042200

140. Holmsen H, Storm E. The adenosine triphosphate inhibition of the pyruvate kinase reaction and its dependence on the total magnesium ion concentration. Biochem J. 1969;112: 303–316. doi:10.1042/bj1120303

141. Urbauer JL, Dorgan LJ, Schuster SM. Effects of deuterium on the kinetics of beef heart mitochondrial ATPase. Arch Biochem Biophys. 1984;231: 498–502. doi:10.1016/0003-9861(84)90413-2

142. Le DT-V, Eckert T, Woehlke G. Computer Simulation of Assembly and Co-operativity of Hexameric AAA ATPases. Salsbury F, editor. PLoS One. 2013;8: e67815. doi:10.1371/journal.pone.0067815

143. Eckert T, Link S, Le DT, Sobczak J-P, Gieseke A, Richter K, et al. Subunit Interactions and Cooperativity in the Microtubule-severing AAA ATPase Spastin. J Biol Chem. 2012;287: 26278–26290. doi:10.1074/jbc.M111.291898

144. Huang X, Yu P, Leproust E, Gao X. An efficient and economic site-specific deuteration strategy for NMR studies of homologous oligonucleotide repeat sequences. Nucleic Acids Res. 1997;25: 4758–4763. doi:10.1093/nar/25.23.4758

145. SP&A Application Laboratory. Microplate-Based Pathlength Correction Method for Photometric DNA Quantification. 2015 p. 4. Available: https://assets.thermofisher.com/TFS-Assets/LCD/Application-Notes/AN-SkanIT-Microplate-Based-Pathlength-Correction-Technical-Note-EN.pdf

146. Nachabé R, Hendriks BHW, Desjardins AE, van der Voort M, van der Mark MB, Sterenborg HJCM. Estimation of lipid and water concentrations in scattering media with diffuse optical spectroscopy from 900 to 1600 nm. J Biomed Opt. 2010;15:

037015. doi:10.1117/1.3454392

147. Veskoukis AS, Margaritelis N V., Kyparos A, Paschalis V, Nikolaidis MG. Spectrophotometric assays for measuring redox biomarkers in blood and tissues: the NADPH network. Redox Rep. 2018;23: 47–56. doi:10.1080/13510002.2017.1392695

148. Sklenar V, Piotto M, Leppik R, Saudek V. Gradient-Tailored Water Suppression for 1H-15N HSQC Experiments Optimized to Retain Full Sensitivity. J Magn Reson Ser A. 1993;102: 241–245. doi:10.1006/jmra.1993.1098

149. Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B, et al. HMDB: A knowledgebase for the human metabolome. Nucleic Acids Res. 2009;37: 603–610. doi:10.1093/nar/gkn810

150. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, et al. The CCPN data model for NMR spectroscopy: Development of a software pipeline. Proteins Struct Funct Genet. 2005;59: 687–696. doi:10.1002/prot.20449

Appendix

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die bei der TUM School of Natural Sciences der Technischen Universität München zur Promotionsprüfung vorgelegte Arbeit mit dem Titel: *NMR-based studies of nucleotide states encountered during the ATP hydrolysis cycle of the p97 enzyme* an der Technischen Universität München, Bayerisches NMR Zentrum, Lehrstuhl für Biomolekular NMR Spektroskopie und an der Ludwig-Maximilians-Universität München, Fakultät für Chemie und Pharmazie, Lehrstuhl für NMR Spektroskopie unter der Anleitung und Betreuung durch Prof. Dr. Anne K. Schütz ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 6 Ab. 6 und 7 Satz 2 angebotenen Hilfsmittel benutzt habe.

Ich habe keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht, oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

Ich habe die Dissertation in dieser oder ähnlicher Form in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.

Ich habe den angestrebten Doktorgrad noch nicht erworben und bin nicht in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.

Die öffentlich zugängliche Promotionsordnung der TUM ist mir bekannt, insbesondere habe ich die Bedeutung von § 28 (Nichtigkeit der Promotion) und § 29 (Entzug des Doktorgrades) zur Kenntnis genommen. Ich bin mir der Konsequenzen einer falschen Eidesstattlichen Erklärung bewusst.

Mit der Aufnahme meiner Personenbezogenen Daten in die Alumni-Datei bei der TUM bin ich einverstanden.

München, 20.01.2025

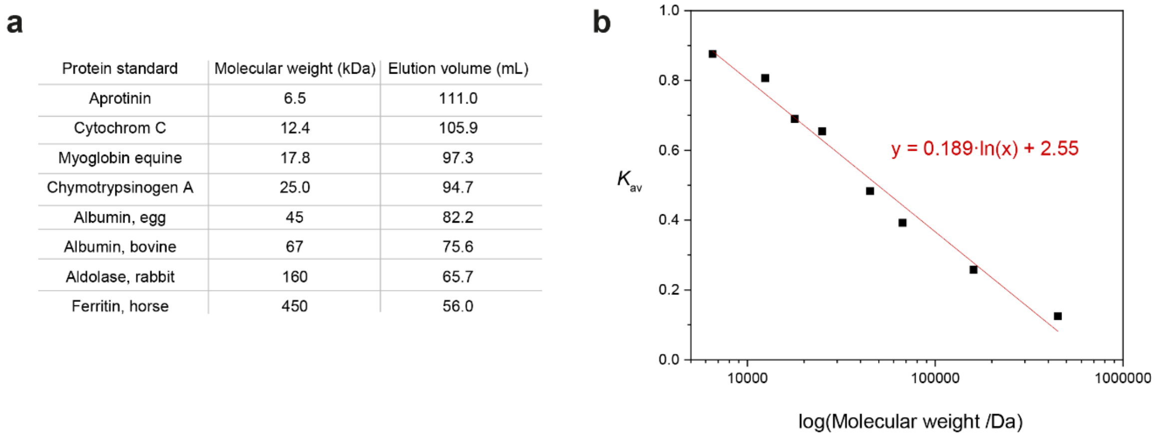
Full-length p97

MGSSHHHHHH SSGDYDIPRE NLYFQGMASG ADSKGDDLST AILKQKNRPN RLIVDEAINE
DNSVVSLSQP KMDELQLFRG DTVLLKGKKR REAVCIVLSD DTCSDEKIRM NRVVRNNLRV
RLGDVISIQP CPDVKYGKRI HVLPIDDTVE GITGNLFEVY LKPYFLEAYR PIRKGDIFLV
RGGMRAVEFK VVETDPSPYC IVAPDTVIHC EGEPIKREDE EESLNEVGYD DIGGCRKQLA
QIKEMVELPL RHPALFKAIG VKPPRGILLY GPPGTGKTLI ARAVANETGA FFFLINGPEI
MSKLAGESES NLRKAFEEAE KNAPAIIFID ELDAIAPKRE KTHGEVERRI VSQLLTLMDG
LKQRAHVIVM AATNRPNSID PALRRFGRFD REVDIGIPDA TGRLEILQIH TKNMKLADDV
DLEQVANETH GHVGADLAAL CSEAALQAIR KKMDLIDLED ETIDAEVMNS LAVTMDDFRW
ALSQSNPSAL RETVVEVPQV TWEDIGGLED VKRELQELVQ YPVEHPDKFL KFGMTPSKGV
LFYGPPGCGK TLLAKAIANE CQANFISIKG PELLTMWFGE SEANVREIFD KARQAAPCVL
FFDELDSIAK ARGGNIGDGG GAADRVINQI LTEMDGMSTK KNVFIIGATN RPDIIDPAIL
RPGRLDQLIY IPLPDEKSRV AILKANLRKS PVAKDVDLEF LAKMTNGFSG ADLTEICQRA
CKLAIRESIE SEIRRERERQ TNPSAMEVEE DDPVPEIRRD HFEEAMRFAR RSVSDNDIRK
YEMFAQTLQQ SRGFGSFRFP SGNQGGAGPS QGSGGGTGGS VYTEDNDDDL YG

ND1L

MGSSHHHHHH SSGDYDIPRE NLYFQGMASG ADSKGDDLST AILKQKNRPN RLIVDEAINE
DNSVVSLSQP KMDELQLFRG DTVLLKGKKR REAVCIVLSD DTCSDEKIRM NRVVRNNLRV
RLGDVISIQP CPDVKYGKRI HVLPIDDTVE GITGNLFEVY LKPYFLEAYR PIRKGDIFLV
RGGMRAVEFK VVETDPSPYC IVAPDTVIHC EGEPIKREDE EESLNEVGYD DIGGCRKQLA
QIKEMVELPL RHPALFKAIG VKPPRGILLY GPPGTGKTLI ARAVANETGA FFFLINGPEI
MSKLAGESES NLRKAFEEAE KNAPAIIFID ELDAIAPKRE KTHGEVERRI VSQLLTLMDG
LKQRAHVIVM AATNRPNSID PALRRFGRFD REVDIGIPDA TGRLEILQIH TKNMKLADDV
DLEQVANETH GHVGADLAAL CSEAALQAIR KKMDLIDLED ETIDAEVMNS LAVTMDDFRW
ALSQSNPSAL RETVVEVPQV TWEDIG

Supplementary Figure 1. Amino acid sequences of p97 constructs used in this work.



Protein standard	Molecular weight (kDa)	Elution volume (mL)
Aprotinin	6.5	111.0
Cytochrom C	12.4	105.9
Myoglobin equine	17.8	97.3
Chymotrypsinogen A	25.0	94.7
Albumin, egg	45	82.2
Albumin, bovine	67	75.6
Aldolase, rabbit	160	65.7
Ferritin, horse	450	56.0

Supplementary Figure 2. Calibration of the S200 column for molecular weight estimation. a) Protein standards, their molecular weights and their elution volume on a HiLoad 16/600 Superdex 200 pg. With a column volume $\boldsymbol{V_c}$ of 120 mL and a void volume $\boldsymbol{V_0}$ of 46.1 mL, the gel-phase distribution coefficient $\boldsymbol{K_{av}}$ can be calculated: $\boldsymbol{K_{av}} = \frac{\boldsymbol{V_e - V_0}}{\boldsymbol{V_c - V_0}}$. b) Plotting $\boldsymbol{K_{av}}$ against the logarithm of the molecular weight yields a linear fit. Unknown molecular weights of biomolecular entities based on their elution volumes can be estimated from that linear regression.

Chemicals and reagents

Supplementary Table 1. Chemicals and reagents used in this thesis.

Chemicals and reagents	Supplier
[$^{13}C_6$]-D-glucose	Sigma-Aldrich
[$^{2}D_7$, $^{12}C_6$]-D-glucose	Cambridge Isotopes
^{15}N-ammonium chloride, $^{15}NH_4Cl$	Cambridge Isotopes
2-(D_3)-methyl,2,4-($^{13}C_2$)-acetolactate	nmrbio
2-keto-4-^{13}C,3,3-D_2-butyrate	Sigma-Aldrich
2-mercaptoethanol	Sigma-Aldrich
ADP	Sigma-Aldrich
Ammonium chloride, $^{15}N_4Cl$	Cambridge Isotopes
AMP-PNP	Jena Bioscience
Apyrase (from *S. tuberosum*, extracted)	Sigma-Aldrich
Apyrase (from *S. tuberosum*, recombinant)	New England Biolabs, NEB
ATP	SERVA
ATPγS	Jena Bioscience
Biotin	SERVA
Bis-maleimidoethane, BMOE	Thermo Fisher Scientific
Boronic acid, H_3BO_3	Acros Organics
Bradford reagent, 5x	SERVA
Calcium chloride, $CaCl_2$	AppliChem
Cobalt chloride, $CoCl_2$-2 H_2O	Acros Organics
Copper (II) chloride, $CuCl_2$-2 H_2O	AppliChem
D_2O	Sigma-Aldrich
Dimethyl sulphoxide, DMSO	Sigma-Aldrich
Disodium hydrogenarsenate, Na_2HAsO_4	Sigma-Aldrich
Disodium hydrogenphosphate, Na_2HPO_4	Sigma-Aldrich
DL-Dithiothreitol, DTT	SERVA
DNAse I	SERVA
DSS	Sigma-Aldrich
E. coli (DH10B, BL21)	EMBL
Ethylene diamine tetraacetic acid, EDTA	Sigma-Aldrich
Glycerol, 86 %	Carl Roth
HEPES	Sigma-Aldrich
Imidazole	Carl Roth
IPTG	SERVA
Iron (III) chloride, $FeCl_3$	Grüssing
Kanamycin sulphate	SERVA
LB	Carl Roth
LB agar	Carl Roth
L-Lactate dehydrogenase	Roche
Lysozyme from chicken egg white	SERVA
Magnesium chloride, $MgCl_2$	Grüssing

Magnesium sulphate, $MgSO_4$	Grüssing
Manganese chloride, $MnCl_2$-4 H_2O	Sigma-Aldrich
NADH	Sigma-Aldrich
Nickel nitriloacetic acid, Ni-NTA	SERVA
Phosphoenolpyruvate, PEP	Sigma-Aldrich
Potassium chloride, KCl	Sigma-Aldrich
Potassium dihydrogenphosphate, KH_2PO_4	SERVA
Primers	Integrated DNA Technologies, IDT
Protease Inhibitor mix HP Plus	SERVA
Protein molecular weight standards	SERVA
Protein Thermal Shift Dye Kit	Thermo Fisher Scientific
Pyruvate kinase (from Bac. stearothermophilus)	Sigma-Aldrich
Pyruvate kinase (from rabbit muscle)	Sigma-Aldrich
Q5 Site-Directed Mutagenesis Kit	New England Biolabs, NEB
Ribose-5-phosphate	Sigma-Aldrich
Sodium chloride, NaCl	Sigma Aldrich
Sodium deuteroxide, NaOD	Sigma-Aldrich
Sodium hydroxide, NaOH	Sigma-Aldrich
Sodium sulphate, Na_2SO_4	Sigma-Aldrich
TB	Carl Roth
Thiamin	SERVA
Tris(2-carboxyethyl)phosphine, TCEP	SERVA
Tris/Tris HCl	Carl Roth
U-$^{13}C^{15}N$-ATP	Sigma-Aldrich
Wizard Plus SV Miniprep DNA purification kit	Promega
Zinc chloride, $ZnCl_2$	Grüssing
ε_1-[$^{13}CH_3$]-methionine	Cambridge Isotopes

Consumables and equipment

Supplementary Table 2. Consumables used in this thesis.

Consumables	Supplier
Ultracentrifugation concentration device, *Amicon*	Merck Millipore
Dialysis tubing, *SnakeSkin*	Thermo Fisher Scientific
NMR tubes, 3 mm	Norell
NMR tubes, 5 mm	Sigma-Aldrich
Optical Adhesive Film for well plates, *MicroAmp*	Thermo fisher Scientific
MAS rotors, various sizes	Bruker
1.3 mm sealing plugs	Bruker

Supplementary Table 3. Laboratory equipment used in this thesis.

Equipment	Name
Thermal cycler	*MiniAmp*, Thermo Fisher Scientific
Homogeniser	*Silent Crusher*, Heidolph
French press	*HTU Digi-F-Press*, G. Heinemann
Standing ultracentrifuge	Thermo Fisher Scientific
HiLoad 16/600 Superdex 200 pg	Cytiva
Superdex 200 Increase 10/300 GL	Cytiva
HiTrap Q column	Cytiva
Table centrifuge	*5910 Ri*, Eppendorf
Table centrifuge	*Micro Star 21R*, VWR
Real-Time PCR	*QuantStudio 3*, Thermo Fisher Scientific
Isothermal titration calorimetry (ITC)	*MicroCal PEAQ-ITC*, Palvern Panalytical
Plate reader	*SpectraMax iD5*, Molecular devices
Ultracentrifuge	Beckman-Coulter
Rotor filling tools	Giotto Biotech
Thermal mixer	*ThermoMixer*, Eppendorf

p97 ILVM expression protocol

Supplementary Table 4. Preparation protocol for 250 mL M9 minimal medium.

Component	Chemical	Stock concentration	Final concentration	Amount to add
D_2O				
M9 salt solution in D_2O, pH 7.2	Na_2HPO_4	15X	33.7 mM	16.7 mL
	KH_2PO_4		22.0 mM	
	NaCl		8.55 mM	
	$^{15}NH_4Cl$		9.35 mM	
D-glucose	$[^{2}D_7, ^{12}C_6]$-D-glucose		3 g/L	750 mg
antibiotic in D_2O	kanamycin sulphate			250 µL
$MgSO_4$ in D_2O		1 M	1 mM	250 µL
$CaCl_2$ in D_2O		1 M	0.6 mM	150 µL
biotin in D_2O		10 mg/mL	10 µg/mL	250 µL
thiamine in D_2O		10 mg/mL	10 µg/mL	250 µL
trace elements in D2O, pH 7.5	EDTA	100X	13.4 mM	2.5 mL
	$FeCl_3$-6 H_2O		3.1 mM	
	$ZnCl_2$		0.62 mM	
	$CuCl_2$-2 H_2O		76 µM	
	$CoCl_2$-2 H_2O		42 µM	
	H_3PO_3		162 µM	
	$MgCl_2$-4 H_2O		8.1 µM	

Supplementary Table 5. ILVM precursors.

Desired labelling	Precursor chemical	Final concentration	Requirement for adding precursor to the medium
I-δ_1-$[^{13}CH_3]$ isoleucine	2-keto-4-^{13}C,3,3-D_2-butyrate	60 mg/L	40 min after the other precursors
L-δ_1(*proR*)-$[^{13}CH_3, ^{12}CD_3]$ leucine	2-(D_3)-methyl,2,4-($^{13}C_2$)-acetolactate	1 aliquot per 250 mL medium	at OD~0.8
V-γ_1(*proR*)-$[^{13}CH_3, ^{12}CD_3]$ valine			
M-ε_1-$[^{13}CH_3]$ methionine	M-ε_1-$[^{13}CH_3]$ methionine	100 mg/L	at OD~0.8

Buffers

Supplementary Table 6. Buffers and their composition used in this thesis.

Buffer name	Components
p97 lysis buffer	100 mM Tris pH 7.4, 500 mM NaCl, 5 mM $MgCl_2$, 20 mM imidazole, 5 % glycerol, 2 mM BME, 1 mg/mL lysozyme, 150-300 units DNAse I, Protease Inhibitor mix HP Plus 1X
p97 wash buffer	50 mM HEPES pH 7.4, 150 mM KCl, 5 $MgCl_2$, 20 mM imidazole
p97 elution buffer	50 mM HEPES pH 7.4, 150 mM KCl, 5 $MgCl_2$, 300 mM imidazole
p97 TEV cleavage buffer	20 mM HEPES pH 7.5, 250 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 2 mM BME
p97 SEC buffer	20 mM HEPES pH 7.5, 250 mM KCl, 1 mM $MgCl_2$
p97 AIEX low salt buffer	50 mM HEPES pH 7.0, 50 mM KCl
p97 AIEX high salt buffer	50 mM HEPES pH 7.0, 1 M KCl
p97 solid-state NMR buffer	25 mM HEPES pH 7.4, 50 mM KCl, 5 mM TCEP, 1 mM DSS
p97 solution-state NMR buffer	25 mM HEPES pH 7.5, 25 mM NaCl, 5 mM TCEP, 100 % D_2O
ATPase assay buffer, 10X	250 mM HEPES pH 7.5, 250 mM NaCl, 40 mM $MgCl_2$, 500 mM KCl, 5 mM TCEP
p97 crosslinking buffer	20 mM HEPES pH 7.2, 250 mM KCl, 5 mM EDTA
^{31}P nucleotide shift minimal ionic conditions	10 mM HEPES, 10 mM $MgCl_2$, 1 mM DSS, 5% D_2O, pH 7.5
^{31}P nucleotide shift high ionic conditions	10 mM HEPES, 10 mM $MgCl_2$, 1 mM DSS, 5% D_2O, 25 mM NaCl, 50 mM KCl, pH 7.5
^{13}C-^{15}N nucleotide shift minimal ionic conditions	10 mM nucleotide, 10 mM HEPES, 0.5 mM Tris^{a}, 1 mM DSS, 1 mM $MgCl_2$^{b}, 5% D_2O, pH 7.5
^{13}C-^{15}N nucleotide shift high ionic conditions	10 mM nucleotide, 10 mM HEPES, 0.5 mM Tris^{a}, 1 mM DSS, 4 mM $MgCl_2$, 5% D_2O, 25 mM NaCl, 50 mM KCl, pH 7.5

[a]: from the ATP solution; [b]: only present in the ADP measurement

Software

Supplementary Table 7. Software used in this thesis.

Software name	Purpose	Developer
Topspin, v 3.7 & v 4.3	Acquisition, processing and analysis of NMR spectra	Bruker
CcpNMR Analysis, v 2.4.2	Analysis of 2D NMR spectra	CCPN [150]
Mnova, v 11.0	Analysis of 1D NMR spectra	Mestrelab