Interplay of Inositol Pyrophosphate Pathway and Iron-Sulfur Cluster Biogenesis

Inaugural Dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Bianca Georgiana Axinte

from Düsseldorf, Germany

Düsseldorf, July 2025

From the Institute of Physical Biology of Heinrich-Heine-Universität Düsseldorf

Printed with approval of the Faculty of Mathematics and Natural Sciences at the Heinrich-Heine-Universität Düsseldorf

Primary advisor: Prof. Dr. Ingrid Span
Secondary advisor: apl.-Prof. Dr. Ursula Fleig

Day of the oral exam: 28.11.2025

TABLE OF CONTENTS

To the ones that watch from beyond.

Constantin

Elena

Dalooa

Abbreviations

Abbreviation	Meaning
1,5-IP_8	1,5-Bisdiphosphoinositol tetrakisphosphate
1-IP_7	1-Diphosphoinositol pentakisphosphate
2x YT	2-Fold yeast tryptone medium
5-IP_7	5-Diphosphoinositol pentakisphosphate
ADP	Adenosine diphosphate
APS	Ammonium persulfate
ATP	Adenosine triphosphate
AU	Absorbance units
BIPY	2,2-Bipyridyl
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
CCAN	Constitutive Centromere-Associated Network
CD	Circular dichroism
CIA	Cytosolic iron-sulfur cluster assembly machinery
CV	Column volume
DIPP	Diphosphoinositol polyphosphate phosphohydrolase
DNA	Desoxyribonucleic acid
DTT	1,4-Dithiothreitol
EAS	Electronic absorption spectroscopy
EDTA	Etylenediaminetetraacetic acid
EPR	Electron paramagnetic resonance spectroscopy
eq	Equivalents
FAD	Flavin adenine dinucleotide
Fe-S	Iron-sulfur cluster
FMN	Flavin adenine mononucleotide
GSH	Glutathione
GST	Glutathione S-transferase
HiPIP	High potential iron-sulfur proteins
His-tag	Histidine tag
HT	High tension voltage
ICP-MS	Inductively coupled plasma mass spectrometry
ICP-MS	Inductive-coupled plasma mass spectrometry
IMAC	Immobilized metal ion affinity chromatography
IMPase	Inositol monophosphatase

INPP4	Inositolpolyphos-phate-4-phosphatase
INPP5	Inositolpoly-phosphate-5-phosphatase
IP	Inositol phosphates
IP$_1$	Inositol-3-phosphate
IP$_2$	Inositol-1,4-bisphosphate
IP$_3$	Inositol-1,4,5-triphosphate
IP3K	Inositol-triphosphate 3-kinase
IP$_4$	Inositol-1,3,4,5-tetrakisphosphate, inositol-1,3,4,6-tetrakisphosphate, inositol-3,4,5,6-tetrakisphosphate
IP$_5$	Inositol-1,3,4,5,6-pentakisphosphate
IP$_6$	*Myo*-inositol, 4-diphosphoinositol-1,3,5,6-tetrakisphosphate
IP6K	Inositol-hexakisphosphate kinase
IPMK	Inositol polyphosphate multikinase
IPP	Inositol pyrophosphate
IPPK	Inositol-pentakisphosphate 2-kinase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ISC	Iron-sulfur cluster assembly machinery
ITPK1	Inositol-tetrakisphosphate 1-kinase
LB	Lysogeny broth
LMW	Low-molecular weight
LUCA	Last universal common ancestor
MBC	
MINPP	Multiple inositol polyphosphate phosphatase
MIPS	*Myo*-inositol-1-phosphate synthase
MRW	Mean residue weight
MS	Mass spectrometry
MT	Microtubule
MWCO	Molecular weight cut off
NaDT	Sodium dithionite
NTA	Nitrile tetra acetic acid
NTPase	Nucleoside triphosphatase
NUDT	Nucleoside diphosphate linked to some other moiety type phosphatases
OD	Optical density
PBS	Phosphate buffered saline
PBS	Phosphate buffered saline

PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PFA-DSP	Plant fungi atypical dual specificity phosphatases
PHO	Phosphate cascade
PHR	Phosphate starvation response transcription factors,
PIP_2	Phosphatidylinositol 4,5-bisphosphate
PIP_3	Phosphatidylinositol 3,4,5-trisphosphate
PI-PLC	Phosphatidylinositol specific phospholipase C
PMVK	Phosphomevalonate kinase
polyP	Inorganic polypohosphates
PPIP5K	Diphosphoinositol pentakisphos-phate-kinase
PTEN	Phosphatase and tensin homolog
PTM	Post-translation modification
RNA	Ribonucleic acid
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SPX	Syg1/Pho81/Xpr1
SUF	Sulfur-utilization factor
TB	Terrific broth
TBZ	Thiabendazole
TCEP	Tris(2-carboxyethyl)phosphin
TEMED	N,N,N′,N′-Tetramethylethylenediamine
TEV	Tobacco etch virus
TGS	Tris-glycine buffered saline
Vtc	Vacuolar transporter chaperone

Zusammenfassung

Inositolphosphaten (IPs), einschließlich Inositolpyrophosphaten (IPPs), sind vielseitige kleine Molekülbotenstoffe, die eine entscheidende Rolle bei zellulären Entscheidungsprozessen spielen. IPs sind an verschiedenen biologischen Prozessen beteiligt, z. B. an der Insulinsignalübertragung, der Tumorzellmotilität, der Apoptose und der Phosphathomöostase. Das Protein Asp1 ist ein Mitglied der hochkonservierten PPIP5K-Familie und spielt eine wichtige Rolle bei der Regulierung des intrazellulären IPP-Spiegels. Es handelt sich um ein Enzym mit zwei Domänen, wobei die N-terminale Domäne als Kinase und die C-terminale Domäne als Pyrophosphatase fungiert. Kürzlich wurde festgestellt, dass die Pyrophosphatase-Domäne einen Eisen-Schwefel (Fe-S) Cluster bindet. Strukturelle Vorhersagen und spektroskopische Untersuchungen haben jedoch widersprüchliche Informationen über die Struktur und den Ort des Clusters ergeben. Dieser Aspekt ist sehr faszinierend, da die PPIP5K-Familie hoch konserviert ist und Asp1 als funktionelles und strukturelles Modell für die gesamte Familie angesehen werden könnte, was neue Perspektiven für zelluläre Funktionen und katalytische Aktivitäten eröffnen würde.

Ziel dieser Arbeit war es, neue Erkenntnisse über die Struktur, die Lage und die Rolle des Fe-S-Clusters in der Pyrophosphatasedomäne von Asp1 zu gewinnen. Frühere Studien standen vor der Herausforderung, genügend Proben für strukturelle und spektroskopische Untersuchungen herzustellen. Daher wurden neue Konstrukte für die heterologe Proteinproduktion entwickelt und bewertet. Es zeigte sich, dass die Verwendung einer für die Expression in *Echerichia coli* optimierten Sequenz die Proteinausbeute erheblich steigern konnte. Die Entfernung der 20 Aminosäuren am C-terminalen Ende führte jedoch nicht zu einer weiteren Steigerung der Proteinausbeute und -stabilität. Um den Ort des Clusters zu identifizieren, wurden Rückstände in der Nähe der vorgeschlagenen aktiven Stelle, die an der Bindung des Clusters beteiligt sein könnten, mutiert und ihre Rolle bei der Ligation untersucht. Auf der Grundlage von Alphafold2-Vorhersagen wurden die Aminosäurereste an den Positionen 397, 607, 643, 644, 663 und 868 durch ortsgerichtete Mutagenese ausgetauscht. Spektroskopische Untersuchungen ergaben, dass die Reste 397, 607 und 663 wahrscheinlich den Cluster koordinieren, während die Reste 643, 644 und 868 höchstwahrscheinlich nicht Teil der ersten Koordinationssphäre des Clusters sind. ^{57}Fe, das Asp1 enthält, wurde genutzt, um die Struktur des Fe-S-Clusters mithilfe der Mössbauer-Spektroskopie zu bestimmen. Die spektroskopischen Daten weisen die Form des Fe-S-Clusters in Asp1 nach. Schließlich wurde im Rahmen dieser Arbeit der oligomere Zustand von Asp1$^{395\text{-}920}$ in An- und Abwesenheit des Fe-S-Clusters untersucht. Die Ergebnisse bestätigen, dass Asp1 überwiegend als Dimer vorkommt, wie aus der Literatur bekannt, und dass es in einen monomeren Zustand übergeht, wenn vier Eisenatome pro Proteinmonomer gebunden sind. Darüber hinaus hat die Wahl der Reagenzien während der chemischen Rekonstitution oder der chromatographischen Experimente keinen Einfluss auf den oligomeren Zustand von Asp1. Insgesamt wirft diese Arbeit Licht auf die Möglichkeit, dass die Asp1-Pyrophosphatase-Domäne von *Schizosaccharomyces pombe* einen

[2Fe-2S]-Cluster in verschiedenen Regionen bindet, wie bisher angenommen, und dass eine Erhöhung des Eisengehalts den oligomeren Zustand von Asp1 beeinflusst.

Summary

Inositol phosphates (IPs), including inositol pyrophosphates (IPPs), are versatile small molecule messengers that play a critical role in cellular decision-making processes. IPs have been implicated in diverse biological processes such as insulin signaling, tumor cell motility, apoptosis, and phosphate homeostasis. The protein Asp1 is a member of the highly conserved PPIP5K family and plays an important role in modulating the intracellular IPP levels. It is a two-domain enzyme in which the N-terminal domain functions as a kinase and the C-terminal domain functions as a pyrophosphatase. Recently, it was found that the pyrophosphatase domain binds an iron-sulfur (Fe-S) cluster. However, conflicting information about the structure and location of the cluster has emerged from structural predictions and spectroscopic studies. This aspect is highly fascinating since the PPIP5K family is highly-conserved and Asp1 could be considered a functional and structural model for the entirety of the family, opening new perspectives for cellular functions and catalytic activities.

The aim of this work was to gain new insights into the structure, location, and role of the Fe-S cluster in the pyrophosphatase domain of Asp1. Previous studies have faced the challenge of producing sufficient samples for structural and spectroscopic studies. Therefore, new constructs for heterologous protein production were generated and evaluated. It was shown that using a codon optimized sequence for expression in *E. coli* could enhance the protein yield significantly but removing the 20 amino acids at the C-terminal end did not increase the protein yield and stability further. To identify the location of the cluster, residues that were close to the proposed active site and that could be involved in cluster binding were mutated and their role in ligation was investigated. Based on Alphafold2 predictions, the amino acid residues at position 397, 607, 643, 644, 663, and 868 were exchanged by site-directed mutagenesis. Spectroscopic studies suggested that residues 397, 607, and 663 are likely to coordinate the cluster, while residues 643, 644, and 868 are most likely not part of the first coordination sphere of the cluster. ^{57}Fe-containing Asp1 was produced to determine the structure of the Fe-S cluster using Mössbauer spectroscopy. The spectroscopic data identify the Fe-S cluster form in Asp1. Lastly, the oligomeric state of Asp1$^{395\text{-}920}$ was investigated as part of this thesis in the presence and absence of the Fe-S cluster. The results confirm that Asp1 is found predominantly as dimer as known from literature and that it transitions to a monomeric state when four iron atoms are bound per protein monomer. In addition, the choice of reagents during chemical reconstitution or chromatographic experiments does not influence the oligomeric state of Asp1. Overall, this thesis sheds light on the possibility that the *S. pombe* Asp1 pyrophosphatase domain binds an [2Fe-2S] cluster in different regions as previously assumed and that an increase in iron content influences the oligomeric state of Asp1.

1 Introduction

1.1 Inositol hexakisphosphate and diphosphoinositolpentakisphosphate kinase family (PPIP5K)

The highly conserved PPIP5K family is a class of enzymes that generate multi-phosphorylated compounds called IPPs alongside the IP_6 kinase family. IPPs produced by PPIP5Ks are well-investigated and comprise 1-diphosphoinositol pentakisphosphate (1-IP_7), 5-diphosphoinositol pentakisphosphate (5-IP_7), and 1,5-bisdiphosphoinositol tetrakisphosphate (1,5-IP_8). The defining feature of this IPP-synthesizing family is its dual domain structure and function, which requires an N-terminal domain with an adenosine triphosphate (ATP) grasp motif and a C-terminal pyrophosphatase domain containing histidine phosphatase signature amino acids. Interestingly, the histidine acid phosphatase motif is not fully conserved among PPIP5K homologues, which has structural implications. A PPIP5K ortholog in *S. pombe* has been identified under the name Asp1.[1,2] Its activity in the modulation of intracellular IPPs levels, especially 1,5-IP_8, has been linked to the existence of two autonomous domains which catalyze opposite reactions of phosphorylation and hydrolysis of one phosphate group at position 1 in the inositol ring (Figure 1).[3–5] The separation of Asp1 in structurally different domains implies a possibly complex regulation of the overall Asp1 activity, since the two domains have opposing activities. [3–5] Asp1 is able to use both *myo*-inositol (IP_6) and 5-IP_7 as substrates for the kinase domain in vivo (Figure 1).[4] However, since the rate constant of the conversion of IP_6 to 1-IP_7 is much lower than the one for the conversion of 5-IP_7 to 1,5-IP_8 the latter reaction is proposed to be physiologically more relevant.[4]

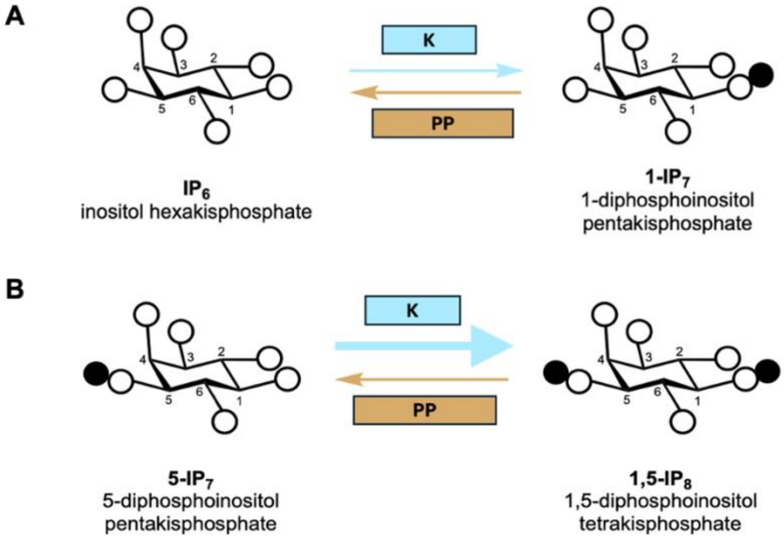


Figure 1: Phosphorylation and hydrolysis reactions catalyzed by the Asp1 kinase and phosphatase domains. A) Reversible phosphorylation reaction catalyzed by the Asp1 domains at C1 of IP_6 and 1-IP_7, respectively. B) Reversible phosphorylation reaction catalyzed by the Asp1 domains at C1 of 5-IP_7 and 1,5-IP_8, respectively. The Asp1 kinase and pyrophosphatase domains are shown separately as blue (kinase) and orange (pyrophosphatase) squares. The prominent arrow indicates that the production of 1,5-IP_8 is the faster of the two reactions. Thus, 1,5-IP_8 is more often synthesized than hydrolyzed in vivo. adapted after 3,4

The study of IPPs-modulating enzymes was marked significantly by the discovery of the 1-IP_7-producing Vip1 protein from *Saccharomyces cerevisiae* and the subsequent characterization of two large mammalian orthologs PPIP5K1 and PPIP5K2, which share 66% overall sequence identity and over 90% sequence identity in the kinase domain.[6–10]. These two human orthologs were described to be IP_6 and IP_7 kinases and together with the inositol-hexakisphosphate kinase (IP6K) they are responsible for generating 1,5-IP_8 which acts as a very potent signaling molecule.[11] PPIP5K/Vip1 orthologs from *Arabidopsis thaliana* VIH1 and VIH2 exhibit the same conserved dual-domain structure and are involved in the maintenance of steady-state and jasmonate-induced 1,5-IP_8 levels.[12] Moreover, it was described that IPPs as well as jasmonate are needed as binding partners to promote the assembly of the jasmonate receptor.[12] Taken together, these findings highlight that evolutionarily conserved PPIP5K enzymes among different organisms are central to the biosynthesis of 1,5-IP_8 and depict IPPs as pivotal signaling molecules that interact with proteins and hormone partners (e.g. jasmonate) to regulate divers cellular processes.

1.1.1 Origin of substrate for PPIP5K enzymes

PPIP5Ks are able to promote both synthesis and degradation of pyrophosphate groups of highly phosphorylated inositol molecules. The question that arises at this point is how is this grade of phosphorylation achieved in vivo in the first place? Soluble inositol phosphates with monophosphate groups at different positions in the hexose ring are referred to as inositol polyphosphates, whereas the ones with diphosphate functionalities at certain carbon atoms build an own class of soluble compounds themselves and are categorized as IPPs. Regarding the occurrence of inositol phosphates in nature, these are synthesized naturally by undergoing either the lipid-dependent pathway or the glucose-dependent pathway.

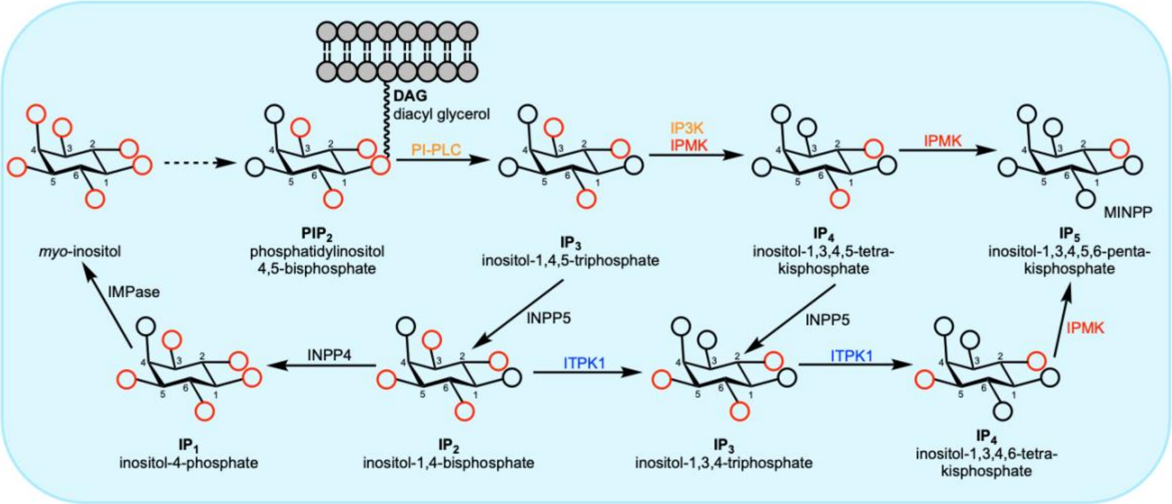


Figure 2: Overview of the lipid-dependent inositol polyphosphates synthetic pathway. The pathway leading to lipid-derived inositol phosphates is highlighted in blue. Red circles indicate ubiquitous hydroxyl groups, whereas black circles stand for phosphate groups. Dashed lines in grey indicate multistep that are not of focus in this figure. Abbreviated names of chemical substrates are shown in

bold black letters. Enzymes with kinase function are shown in different colours to underline substrate promiscuity specific for this enzyme class. Phosphatases are illustrated in black letters above the reaction arrow. PI-PLC = phosphatidylinositol specific phospholipase C, IPMK = inositol polyphosphate multikinase, IP3K = inositol-triphosphate 3-kinase, IMPase = inositol monophosphatase, INPP4 = inositolpolyphos-phate-4-phosphatase, INPP5 = inositolpolyphosphate-5-phosphatase, ITPK1 = inositol-tetrakisphosphate 1-kinase, MINPP = multiple inositol polyphosphate phosphatase. adapted after 13

The synthesis of the inositol pyrophosphates takes place in a successive manner under the activity of a multitude of enzymes (Figure 2). In the first pathway, the biosynthesis is initiated by the release of inositol-1,4,5-triphosphate (IP_3) from the receptor-bound phosphatidylinositol 4,5-bisphosphate (PIP_2) and a diacyl glycerol moiety via phosphatidylinositol specific phospholipase C (PI-PLC).[14–17] In the next steps, IP_3 can be further phosphorylated up to inositol-1,3,4,5,6-pentakisphosphate (IP_5) by successive activity of IPMK or IP3K, which act in a redundant manner.[18–20] The IP_3 and inositol-1,3,4,5-tetrakisphosphate (IP_4) intermediates built up to this point, as well as inositol-1,4-bisphosphate (IP_2) can be dephosphorylated by inositolpoly-phosphate-5-phosphatase (INPP5) or inositolpolyphosphate-4-phosphatase (INPP4), respectively, at any moment according to cellular needs, this leading to the release of IP_2, IP_3 or inositol-3-phosphate (IP_1) building blocks that can be further degraded to IP_1 and IP_6, or can be used for the synthesis of an IP_4 isomer (Figure 2).[21–23] The in vivo multi-step synthesis of IPPs not only provides different inositol metabolites as signaling molecules but enables alternate synthesis pathways via cross reactions. Thus, specific IPPs can be synthesized according to cellular needs and their biological function in more than one way.

From the IP_5 state another phosphorylation to inositol hexakisphosphate by the inositol-pentakisphosphate 2-kinase (IPPK) enables the production of the first fully phosphorylated inositol compound which can enter the synthetic pathway catalyzed by PPIP5K, IP6K and diphosphoinositol polyphosphate phosphohydrolase (DIPP) where pyrophosphorylated compounds are selectively produced and converted into each other.[24–28] In comparison with the many kinases mentioned above, DIPP belongs to the class of hydrolases and catalyzes the hydrolytic cleavage of β-phosphate groups from pyrophosphorylated IPP species.[27,28] In the turnover of inositol pyrophosphates catalyzed by *S. pombe* Asp1, the kinase domain of this protein facilitates specific phosphorylation reactions at position 1 of the hexose ring of IP_6 and 5-IP_7, this resulting in a new class of soluble diphosphoinositol pyrophosphates.

1.1.2 Biological function of PPIP5Ks

There are many IPPs which result from the activity of three IPP-processing enzymes. The critical role of such IPP compounds is highlighted by the existence of two alternating biosynthetic pathways to generate these signaling molecules. Due to their highly-energy content they play a role in many processes such as chemotaxis, dimorphic switch response, apoptosis and the regulation of endocytosis.[29–32] In numerous essential biological processes IPPs act as signaling molecules in intracellular pathways.[33–35] Studies investigating the biological function of IPPs demonstrated that

Asp1-generated IP_7 and IP_8 are key regulators of the dimorphic switch in *S. pombe*. This switch enables the yeast to transition from a unicellular to an invasive pseudohyphal growth form in response to environmental cues, which is a conserved mechanism of morphogenesis in fungi. [2] It was reported that IPPs regulate this cellular process via cyclic adenosine monophosphate protein (cAMP) kinase A pathway. The absence of 1,5-IP_8 impacts the cellular growth response whereas an increase further promotes the invasive growth. The dimorphic switch in *asp1*Δ is not activated even though the 5-IP_7 levels are significantly increased compared to wildtype.[2] This study clarifies, that 1,5-IP_8 produced by the PPIP5K family is predominantly relevant for the modulation of the dimorphic switch as a response mechanism to external stimuli beside their central role in the regulation of microtubule dynamics.

The MT skeleton is highly dynamic and has multiple roles in interphase such as transport of organelles. In mitosis MTs form the mitotic spindle required for chromosome segregation. In *S. pombe*, this mechanism is dependent on 1,5-IP_8 generated by Asp1. The mutation of the catalytic aspartate D333 in the Asp1 kinase domain generated a kinase-dead enzyme where delayed segregation of sister chromatids was observed during mitosis. This highlights the role of 1,5-IP_8 and the Asp1 protein in the modulation and stabilization of the MT cytoskeleton.[36]

In addition, PPIP5K were shown to be one of the three phosphatase families, alongside to plant fungi atypical dual specificity phosphatases (PFA-DSPs) and nucleoside diphosphate linked to some other moiety type (NUDT) phosphatases, that contribute to IPP catabolism for the regulation of plant growth and development.[37] Mutants of PPIP5K ortholog of *Marchantia polymorpha* exhibited severe growth and development defects due to altered IPP levels. Upon quantification, it was reported that the deletion of PPIP5K *M. polymorpha* Vip1 leads to elevated intracellular levels of 5-IP_7 and reduced levels of 1,5-IP_8, which consist with the enzymatic properties of PPIP5K kinase domain.[37] Furthermore, the variant of Vip1 of *M. polymorpha* exhibited higher levels of accumulated nitrate, this suggesting that IPPs modulate nitrate homeostasis, yet the involved IPP isomer remains elusive. Nonetheless, these studies now link the cellular IPP pools to nitrate homeostasis and suggest that PPIP5Ks together with PFA-DSP and NUDT phosphatases regulate plant growth and development collectively by modulating intracellular IPPs pools.[37]

1.1.3 Enzymatic activity of PPIP5Ks

The dual structure of the PPIP5K/Vip1 orthologs and their competing functionalities imply a highly regulated mechanism behind the in vivo function of these enzymes. The ATP-grasp fold of the N-terminal domain of PPIP5Ks is characteristic for proteins that perform ATP-hydrolysis. Among the PPIP5K orthologs this domain is responsible for phosphorylation reactions of 5-IP_7 at position 1 leading to the formation of a double pyrophosphorylated product 1,5-IP_8 (Figure 1). Interestingly, it was shown in vitro that an upregulated kinase domain can accept IP_6 as substrate for phosphorylation and thus generate 1-IP_7, but only after the substrate pool of 5-IP_7 is completely depleted. In vivo both reactions depicted in Figure 1 are possible, with reaction B possibly being the biologically more relevant by the kinase domain as underlined by the large arrow. This preference towards 5-IP_7 was quantified as rate of substrate phosphorylation and was investigated under competitive substrate conditions with excessive concentrations of IP_6.[4] Still, the first-order rate constant was significantly higher for 5-IP_7 than for IP_6, showing that the conversion of 5-IP_7 to 1,5-IP_8 is the favoured reaction in vivo.[4] Physiologically, 1,5-IP_8 makes only 10% of the intracellular levels of its precursor 5-IP_7 in mammalian cells.[24] It remains unknown to this day if this dual-function enzyme class is regulated only by 1,5-IP_8 or by the concentration difference between the cellular pools of 1,5-IP_8 and 5-IP_7. These regulation mechanisms remain elusive since it has been shown that the enzymes do not differentiate between 5-IP_7 and 1,5-IP_8 in vitro.

Furthermore, the literature to PPIP5K/Vip1 family of enzymes with competing kinase and phosphatase domains offers divergent proposals as to which reaction catalyzed by one of the domains is the more biologically relevant. To date it is known that PPIP5K/Vip1 orthologs can phosphorylate two IP species, IP_6 or 5-IP_7, with the phosphorylation of 5-IP_7 to 1,5-IP_8 possibly being the more relevant one in vivo.[38] Early studies performed in *S. cerevisiae* highlight the impact of Vip1-catalyzed phosphorylation of IP_6 to 1-IP_7, a reaction that is upregulated in vivo under phosphate starvation conditions.[39] However, more recent literature challenges this theory by reporting that a decrease in the extracellular phosphate pool leads to a significant decrease of the IP_7 levels.[4,40–43] The reduction of IP_7 levels could alternatively occur due to its conversion to the pyrophosphorylated metabolite 1,5-IP_8.[4,41,44] While in vitro studies and truncated versions of Asp1 have demonstrated kinase activity toward 5-IP_7 and, under certain conditions, towards IP_6, these results do not fully capture the regulatory complexity in vivo.[4] Recent publications have shown that the full-length Asp1 protein in *S. pombe* exhibits a preference for converting 5-IP_7 to 1,5-IP_8 under physiological conditions, rather than phosphorylating IP_6 to 1-IP_7.[4] These findings suggest that the primary in vivo function of Asp1 is the phosphorylation of 5-IP_7 to 1,5-IP_8, whereas the production of 1-IP_7 from IP_6 may be of limited physiological relevance, occurring mainly under specific stress conditions such as phosphate starvation.[4] Taken together, the most likely and biologically relevant source of 1-IP_7 in the cell is the dephosphorylation of 1,5-IP_8. In vitro, the phosphorylation of IP_6 to 1-IP_7 occurs at similar rates as the reaction of 5-IP_7 to 1,5-IP_8, but this does not reflect the in vivo regulation. Upon external stimuli, 1-IP_7 may either be degraded to IP_6 to support phosphate recycling or utilized for synthesis of 1,5-IP_8.

The second homologous domain found among PPIP5Ks is the pyrophosphatase domain which shares motifs with the phytase subgroup of histidine acid phosphatases (Figure 3).[5] The amino acids involved in the catalytic reaction can be identified from the name of the enzyme class and consist in two histidine residues (RHxxR and HD).[45] During the phosphatase reaction of histidine acid phosphatases, the transfer of a phosphate group to the first histidine in the RHxxR motif is facilitated by the surrounding two arginine residues. These aid by balancing the large electronegative density of the phosphate group with their own positive charges. The second histidine from the HD motif is involved in the catalytic reaction at the active site and the accompanying aspartate is the residues that ultimately donates a proton and support the release of the phosphate group from the binding pocket.[45]

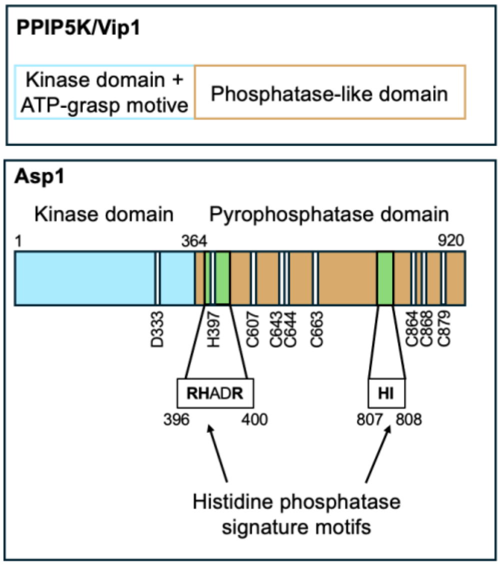


Figure 3: Overview of the dual-structure domain shared among PPIP5K orthologs. The kinase domain is depicted in blue, whereas the pyrophosphatase domain in orange. Catalytic amino acids (D333, H397) and potential residues involved in the coordination of the Fe-S cluster (cysteines) are highlighted as white lines. The conserved histidine phosphatase signature motifs are marked with green in the protein sequence.

However, members of the PPIP5K/Vip1 family do not present the typical HD motif since the proton-donating aspartate is substituted by either valine, alanine, or isoleucine (Figure 3).[11] This was the reason why the C-terminal pyrophosphatase domain of PPIP5K/Vip1 proteins was erroneously described as inactive for a long time.[46] The first function of the histidine acid phosphatase-like domain was identified as part of a study regarding the role of the Vip1 kinase family in the regulation of polarized growth and MT dynamics in fungi.[47] Here the analyzed single *S. pombe* Asp1 mutants were a kinase-deficient one lacking the D333 catalytic aspartate and a phosphatase-deficient one where H397 of the first histidine motif was substituted.

Moreover, the phosphorylation of IP_6 to 1-IP_7 was pursued in vitro using wildtype and mutated Asp1 variants and different combinations of the constructs. These experiments indicated that $Asp1^{H397A}$ has a much higher 1-IP_7 output, this correlating with a defective phosphatase-like domain due to the H397A mutation.[47] At the same time, an in vitro assay with purified protein showed that the addition of the pyrophosphatase domain to full-length Asp1 reduced the overall in vitro 1-IP_7 output whereas the use of a mutated version ($Asp1^{365\text{-}920\ H397A}$) did not cause any change.

These results highlight two possible modes of regulation of the two domains. The pyrophosphatase domain could either modulate the function of kinase domain directly, thus decreasing the IPPs levels. However, it is as possible that the pyrophosphatase domain of Asp1 exhibits enzymatic activity and uses the IPPs produced by the kinase domain as substrate. Overall, these results were the first evidence that *S. pombe* Asp1 has C-terminal 1-phosphatase activity and that the substrates of the histidine acid phosphatase-like domain are N-terminal generated IPPs.[47]

1.1.4 Mode of action of PPIP5Ks

The mode of action of these proteins can be very diverse and it can be brought in connection with conformational flipping of IP_8, IP_8 receptor activity, RNA processing, IP_8-mediated modulation of PIP_3 signaling, immunity response in plants and protein phosphorylation. It is widely known that five of the substituents of IPPs adopt an equatorial conformation inside the inositol ring and that only the 2-phosphate has an axial orientation. It has been reported that under harsh conditions which often imply a highly acidic milieu and elevated levels of magnesium the conformational equilibrium shifts in the favour of a flip in the opposing 5-axial and 1-equatorial state.[48,49] Such a conformational change is often facilitated by the increase in the number of phosphate groups which are space demanding. It has been observed that 30% of 1,5-IP_8 is flipped under physiological conditions. Interestingly, the fact that this flipped conformation cannot be observed for IP_7 could be indicative for ligand specificity towards more phosphorylated IPPs.[50]

One mode of action of PPIP5Ks is best described by the IP_8 receptor activity of Syg1/Pho81/Xpr1 (SPX) domains. Here the interaction between SPX domain of the vacuolar transporter chaperone (Vtc) complex and IPPs plays a central role. The Vtc complex identified in *S. cerevisiae* contains four known subunits. *S. cerevisiae* Vtc1 comprises only three helices and makes up a small part of the complex.[51,52] Some of the Vtc proteins (Vtc2, Vtc3 and Vtc4) comprise a C-terminal transmembrane helix and N-terminal SPX domain. Most of the proteins involved in phosphate homeostasis contain SPX domains in order to regulate phosphate transport and activate phosphate storage via the Vtc complex.[53,54] These processes are modulated by 1,5-IP_8 that binds directly to SPX domain based on their negative charge. [55] This highlights the fact that the SPX-Vtc4-Vtc2/3 complex is a multifactorial system that is affected by the presence of the pyrophosphorylated 1,5-IP_8 molecule.[56]

The SPX domains of the Vtc4-Vtc2/3 complex regulate vacuolar polyphosphate synthesis by forming a transmembrane channel that remains closed through SPX-SPX interactions. [55,57] Binding of 1,5-IP_8 to a conserved lysine residue in the SPX domains disrupts this interaction and opens the channel, enabling ATP transport and polyP synthesis. [55,57] In *S. pombe*, the Asp1D333A mutant lacks kinase activity and cannot produce 1,5-IP_8, leading to impaired channel opening and reduced polyP synthesis, despite elevated 5-IP_7 levels. [58] This suggests that 1,5-IP_8 is essential for SPX-mediated activation of polyP synthesis and highlights its central role in phosphate homeostasis. [55,58]

Furthermore, it has been reported that 1,5-IP_8 can enhance SPX interactions with target proteins. In plants, the activation of phosphate starvation induced genes is conducted by dimeric phosphate starvation response (PHR) transcription factors, which are not genetically repressed any longer in the presence of elevated IP_8 levels.[59–61] Structural studies conducted by two groups favor a model where IPPs bind allosterically to SPX domains to promote the formation of a 1:1 SPX1:PHR complex and at the same time the uncoupling of the PHR:PHR dimer. Immunoprecipitation experiments further confirmed that a IPP binding site exists outside of the SPX1-PHR interface and that 1,5-IP_8 allosterically regulates SPX/PHR dimerization.[59–61]

The SPX recently became a central aspect for regulation by IPPs in context with the phosphate (PHO) cascade. Pho81 is the phosphate system positive regulatory protein 81 of *S. cerevisiae*, which is part of the PHO cascade and contains a SPX domain. There is agreement that the activity of Pho81 is regulated by IPPs yet the nature of the IPP species and the binding site remained elusive. More recent work identified Lys54 as an IPP binding residue in the Pho81 SPX domain and that 1,5-IP$_8$ acts as a central regulatory ligand for SPX domains.[62–65] Further experiments identified interaction of Pho81 SPX domains with 5-IP$_7$ as ligand to be possible as well. The current hypothesis is that the Pho81 SPX domains' selectivity towards 1,5-IP$_8$ is flexible, and that function can be rescued by compensating low intracellular 1,5-IP$_8$ levels with the physiologically tenfold higher 5-IP$_7$ levels. [62–65]

However, this is only one of the known modes of action for PPIP5Ks and the IPPs underlying their function, since they appear to be involved in the pyrophosphorylation and binding of different proteins. The in vitro protein phosphorylation by 5-IP$_7$ and 1,5-IP$_8$ is a chemical reaction driven by mass actions effects and the rate of phosphorylation depends on the concentration of the phosphate-donating molecules.[66–68] This reaction involves the cleavage of a β-phosphate from 5-IP$_7$ and 1,5-IP$_8$ and its transfer to a pre-existing phosphoserine or phosphothreonine residue. Here Mg^{2+} is required as a cofactor an acidic environment is preferred, even though this reaction is not catalyzed by any enzyme. [66–68] Interestingly, most of the examples for protein phosphorylation with biological relevance have been attributed to 5-IP$_7$.[69] It is possible that this molecule outcompetes 1,5-IP$_8$ for pyrophosphorylation since the intracellular physiological 5-IP$_7$ levels are tenfold higher. IPPs were not only found to bind known interaction partners such as IP6K1, IP6K2, PPIP5K2, DIPP1 and DIPP2 with high affinity but also the motor protein KIF14 (by 5-IP$_7$) and phosphomevalonate kinase (PMVK, by 1,5-IP$_8$).[70] Furthermore, it was reported in vitro that IP$_7$ is responsible for the physiological phosphorylation of mammalian and yeast proteins in a way that depicts pyrophosphorylation as overall process.[71,72] Target proteins must be primed by a phosphate group from ATP to undergo pyrophosphorylation by IP$_7$ and appear to be more resistant to phosphatases after this posttranslational modification (PTM).[72] This study supports the role of IPPs as phosphorylating agents, which was determined much earlier by measuring the standard free energy of hydrolysis of the phosphor anhydride bond of 5-IP$_7$ and 1,5-IP$_8$ which is estimated to be higher and lower than the values for adenosine diphosphate (ADP) and ATP, respectively.[73,74] This effect is enhanced by the strong electrostatic repulsion and steric constraints of the vicinal pyrophosphate functional groups of 1,5-IP$_8$ and lead to a free hydrolysis energy rivaling the highest value known for ATP.[75] More recent studies using a combination of radiolabeling and mass spectrometry (MS) methods further affirmed that PPIP5K-generated IPPs promote intracellular protein pyrophosphorylation.[76] There is a strong dependence of endogenous phosphorylation on the IPPs, with 5-IP$_7$ as the key phosphate donor. The role of 1,5-IP$_8$ in pyrophospharylation was questioned and a comparison of wildtype and PPIP5K-defficient cell lines did not show a significant difference in the pyrophosphorylation levels.[76] This once again shows that 1,5-IP$_8$ is not a major regulator of protein pyrophosphorylation, but this can also be due to the naturally lower physiological levels of this IPP species in the cell in comparison to

5-IP_7 which is ten times more concentrated.[76] In summary, PPIP5K-generated IPPs not only maintain intracellular phosphate homeostasis by modulating SPX-SPX interactions and polyP synthesis but they also act binding partners and phosphorylation agents for various proteins.

1.2 Structure and function of Asp1

1.2.1 Biological function of Asp1

Asp1, a member of the PPIP5K enzyme family, has functional roles that extend beyond the regulation of IPP metabolism, notably in morphogenesis, chromosome segregation and phosphate homeostasis. In the context of cellular morphogenesis, Asp1-generated IPPs influences the dimorphic switch in *S. pombe*, a process by which the yeast alters its growth form in response to environmental or internal cues such as nutrient limitation, temperature changes, or specific signaling molecules. [77–80]

Asp1 also plays a pivotal role in chromosome segregation, particularly through the regulation of intracellular 1,5-IP_8 levels. Depletion of 1,5-IP_8 results in severe chromosome transmission defects, including aneuploidy and polyploidy. [36,81] Interestingly, 1,5-IP_8 affects kinetochore architecture my modulating the composition of the Constitutive Centromere-Associated Network (CCAN) complex. [82] For example, the reduced levels of 1,5-IP_8 lead to increased kinetochore association of specific proteins such as Mal2 and Fta2, whereas other CCAN components show the opposite response. [82] This seemingly contradictory effect suggests that 1,5-IP_8 does not uniformly enhance or inhibit kinetochore binding but instead fine-tunes the complex's composition to support accurate chromosome segregation. [82]

Asp1 of *S. pombe* plays a central role in regulating phosphate homeostasis through its inositol pyrophosphate signaling activities. [83] Intracellular levels of IPPs, particularly 1,5-IP_8 directly influence the transcriptional phosphate starvation response: loss of Asp1 function results in hyper-repression of PHO regulon genes even under phosphate-replete conditions, while elevated 1,5-IP_8 leads to de-repression of the PHO regulon genes.[83] Mechanistically, phosphate acts allosterically to inhibit Asp1's phosphatase activity, thereby promoting net synthesis of 1,5-IP_8 and reinforcing phosphate signaling. [83] Additionally, depletion of extracellular phosphate rapidly decreases cellular 1,5-IP_8 levels, suggesting a feedback loop in which Asp1 dynamically adjusts IPP pools in response to environmental phosphate. [83] Taken together, these findings establish Asp1 as a biochemical sensor that couples phosphate availability to gene expression, via tight control over intracellular 1,5-IP_8 concentrations and downstream transcriptional regulators. [83]

1.2.2 Enzymatic role of Asp1

An interesting aspect regarding the Asp1 enzymatic activity is the fundamental separation of the two domains and their activities. Various groups have reported a different division of the complete protein sequence of Asp1 which led to different assumptions about the correct length of each domain. One group reported the kinase domain of Asp1 comprises the first 371 amino acids and the pyrophosphatase domain makes out 60% of the protein sequence with its 549 remaining amino acids.[5] Based on sequence alignments and secondary structure predictions a second group produced a version of the Asp1 pyrophosphatase domain shorter by six amino acids.[3] The C-terminal domain of Asp1 was described to exhibit phosphatase activity, yet the unclarity regarding the separation of the two domains raises question about the regulation of the latter domain and if the 6 amino acids make any difference.

It is rather unusual that an enzyme can synthesize an important compound with one of its domains only for it to be directly degraded by the neighboring domain. Literature and multiple studies about other PPIP5K family members indicates that the two functionalities are strictly regulated. This is possibly the case for Asp1 as well but the exact regulation patterns of the two competing activities remain elusive.

As mentioned earlier, most of PPIP5K/Vip1 orthologs contain the signature motifs of this domain which are fully conserved. In case of Asp1 the amino acid at position 808 is exchanged by an isoleucine, which substitutes the expected proton-donating aspartate residue. Overall, the phosphatase activity against 1-IP_7 and 1,5-IP_8 was understood but was demonstrated only later for the first time for any of the PPIP5K members. The catalytic activity against these IPPs was confirmed for the first time using an in vitro test system that demonstrated that the C-terminal domain of said family members indeed acts as a phosphatase.[47]

A correlation on cellular basis was discovered between PPIP5K/Vip1-produced 1,5-IP_8 levels and microtubule stability, a phenomenon which was investigated in detail on the example of Asp1 from *S. pombe*. When overexpressing the wildtype and intact pyrophosphatase domain on a plasmid, the activity of this domain leads to depletion of the intracellular 1,5-IP_8. This causes sensitivity of the yeast strain harbouring the Asp1$^{365\text{-}920}$ plasmid against MT destabilizing agents, such as thiabendazole (TBZ) and methyl benzimidazole-2-yl carbamate (MBC). A strain overexpressing plasmid-born Asp1$^{365\text{-}920\ \mathrm{I808D}}$ is significantly less sensitive towards MT destabilizing agents, this indicating that Asp1$^{365\text{-}920\ \mathrm{I808D}}$ is inactive and cannot catalyze the in vivo degradation of 1,5-IP_8. This way the intracellular 1,5-IP_8 levels are higher than physiologically normal and ensure decreased sensitivity of the yeast strain with the relevant Asp1 version towards TBZ and MBC. Interestingly, the mutation of I808 to an aspartate to generate an Asp1 version with the correct signature motif of a histidine acid phosphatase results in the complete loss of C-terminal activity.

The so far available insights into the function and regulation of the Asp1 pyrophosphatase domain raise more questions about the actual mechanism according to which this protein conducts IPP catalysis. Insights into a potential structural model of the *S. pombe* Asp1 pyrophosphatase domain were reported more recently, whereas the studies in question present divergent hypothesis regarding the functional regulation of the pyrophosphatase activity due to the unusual presence of an Fe-S cluster as cofactor, which has been observed for the first time in a phosphatase domain.[5]

1.3 Iron-sulfur clusters in Nature

Iron-sulfur clusters have always played a crucial role since the beginnings of evolution and the existence of primordial catalysis.[84] Today these serve as relics of ancient sulfur-based and anaerobic metabolism.[84,85] It is known today that many ancient key proteins such as metalloproteins are descendants of the last universal common ancestor (LUCA). Early life organisms most probably were not equipped with complex metal-acquisition machineries but had to evolve in certain geothermal areas where necessary metals were so abundant that they were obtrusive to biological cells.[85] Examples are the complex structures of heme and chlorophyll that are centered around metal ions and are still highly relevant even in today's life and biology. It is assumed that these structures formed in milieus with very high pH values, for example a hydrothermal vent in the immediate surrounding of cooling volcanic lava. Machineries for metal supply as well as heat-shock proteins to facilitate correct folding and assembly are crucial to all organisms, yet their development occurred in the later stages of terrestrial life and not in the beginning of evolution.[85] At that time the atmosphere contained much lower amounts of oxygen and the oxidation of the Fe(II) to insoluble Fe(III) by molecular oxygen was not as problematic for the stability of Fe-S clusters as it is today. [85] Due to this beneficial atmospherical conditions at those early times iron-sulfur clusters proved to be key players and major stabilizing compounds for cellular metabolism during primordial catalysis. [85] Organisms adapted to these conditions by replacing iron atoms with other metal ions to fulfill the same cellular function, however up to date iron-sulfur proteins are still extremely common in all living organisms.

1.3.1 Biological role of iron-sulfur clusters

Iron-sulfur clusters play a key role in many biological processes such as photosynthesis, nitrogen fixation, sensing of iron and oxygen, oxidative respiration, hydrogen production, and desoxyribonucleic acid (DNA) repair.[86] The most important functions that they assume here are electron transfer, sensing of reactive oxygen species, redox and non-redox catalysis. The iron atoms inside a cluster are able to transport electrons from a donor to an acceptor, which possibly enters an active conformation upon its reduction.[87] Ferredoxins are the first class of Fe-S that were discovered and are well known to fulfill this function of electron transport. Their low reduction potentials (-0.5 and -0.15 V) and their [2Fe-2S] cluster located on the protein surface make them suitable for electron transport.[88] This way, Fe-S cluster modulate the activation of proteins. Due to their various redox potentials, Fe-S clusters can sense environmental changes. For example, they can undergo molecular changes in the presence of oxygen, iron or reactive oxygen and nitrogen species.[89,90]

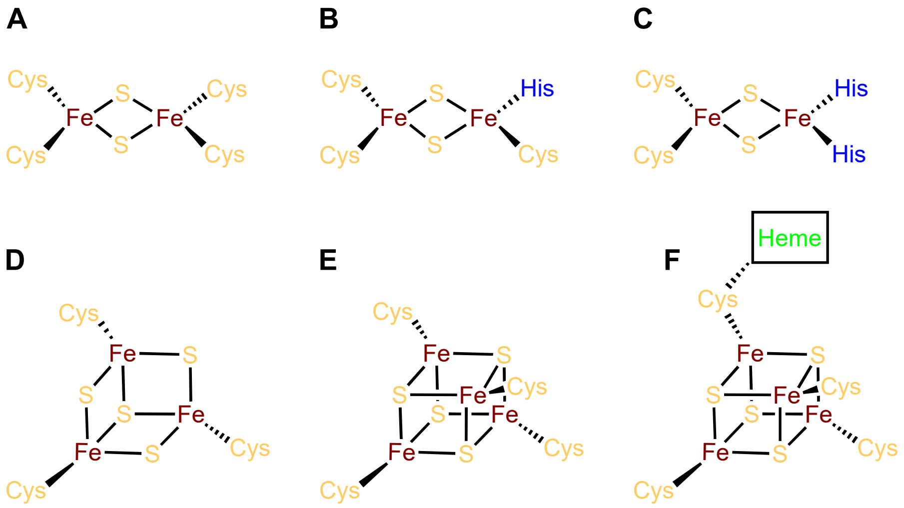


Figure 4: Overview of the different Fe-S types based on structure and nuclearity. A) Rhombic [2Fe-2S] cluster. B) NEET-type [2Fe-2S] cluster. C) Rieske-type [2Fe-2S] cluster. D) [3Fe-4S] cluster. E) Cubane [4Fe-4S]. F) Cubane [4Fe-4S] + thiolate group of heme.[adapted after 91]

Fe-S clusters occur in different forms and structures. The most common types are the rhombic [2Fe-2S] clusters and cubane [4Fe-4S] clusters. All protein-bound Fe-S clusters consist of iron and inorganic sulfur, except for rubredoxins which bind one Fe atom per protein monomer.[92] The iron atoms are bound to the protein via thiol groups of cysteine residues. The cubane Fe-S clusters are some of the most common in nature since they exhibit a tendency for self-assembly. One of the main processes mediated by these cofactors is electron transfer. This is facilitated by the uniform delocalization of valence electrons that leads to a low reorganization energy. In addition, [4Fe-4S] clusters can be classified in ferredoxin-type and high potential iron-sulfur proteins (HiPIP).[93] The resting state is similar for both families however, they differentiate in the transition to their active state. Ferredoxin-type clusters achieve their active state by reduction, whereas HiPIP need to be oxidized to enter their active state. The crown-shaped [3Fe-4S] clusters are closely related to the cubane centers, except for one iron ion is absent in this structure. Due to the missing metal ion the delocalization is not as pronounced as for cubane Fe-S clusters. Furthermore, a [4Fe-4S] cluster can bind larger functional groups such as heme groups by forming thiolate bonds with one of the coordinating cysteine residues (Figure 4).[91]

Rhombic [2Fe-2S] clusters are among the most common cluster structures next to the cubane Fe-S clusters. In comparison to the cubane clusters, rhombic clusters do not exhibit any delocalization of electrons, but they are still active as cofactors in redox processes. Usually, cysteines are the amino acids in a protein that bind the cluster near the backbone, however histidines can also contribute by ligating both lower and higher nuclearity cluster types, thus giving rise to new coordination patterns and cluster subclasses.[94,95] A [2Fe-2S] can be classified as NEET-type as soon as one of the coordinative cysteines is replaced by a histidine residue (Figure 4).[96] The substitution of one more cysteine residue leads to the transition to a Rieske-type [2Fe-2S] cluster.[97] These Fe-S clusters have a wider range of reduction potential due to the difference in the net charge caused by the coordination of two histidine residues. Rubredoxins are small proteins which belong to the class of Fe-S proteins and contain one Fe atom coordinated by four cysteinyl residues in their active site.[92] The tetrahedral coordination environment of the metal ion in rubredoxins is similar to the one of [2Fe-2S] proteins.[92,98] Despite their complex and very different structures, Fe-S clusters are synthesized in the same manner in vivo.[99]

1.3.2 Biogenesis of iron-sulfur clusters

The biosynthesis, trafficking and target-specific insertion of Fe-S clusters require a large number of proteins and are strictly regulated processes. There are three key machineries that facilitate the assembly of Fe-S proteins in the cell. The iron-sulfur cluster (ISC) machinery located in the mitochondria of eukaryotes and bacteria is a functional operon comprising thirteen genes coding for proteins fulfilling full maturation of most iron-sulfur proteins. The sulfur utilization factor (SUF) is another machinery that partakes in the assembly of Fe-S proteins and the one responsible for maturating housekeeping Fe-S proteins under oxidative-stress conditions (Figure 5). In the case of cytosolic and nuclear Fe-S proteins their maturation is executed by the cytosolic iron-sulfur cluster assembly (CIA) machinery.[100,101]

The first step of the ISC-promoted cluster formation is the desulfurization reaction of cysteine to alanine leading to persulfate release which is promoted by a 5'-phosphate-dependent cysteine desulfurase IscS.[102,103] In the next step, the sulfur ions are transferred to IscU which acts as a scaffold protein and undergoes a transient complex with IscS for the assembly of [2Fe-2S] and possibly [4Fe-4S] clusters in presence of electrons, sulfur anions and iron cations in the surrounding milieu.[104–106] In addition to sulfur, the de novo formation of Fe-S clusters depends on a supply of iron. To date, no dedicated iron donor – whether a low-molecular-weight complex or an iron-specific chaperone – has been identified in any of the established biogenesis pathways. As a result, it is presumed that iron is taken directly from the labile iron pool, which in yeast and mitochondria contains approx. 150 µM Fe^{2+}.[107] Still, the ongoing requirement for iron in Fe-S cluster assembly raises the possibility that a specialized iron delivery system has yet to be discovered.

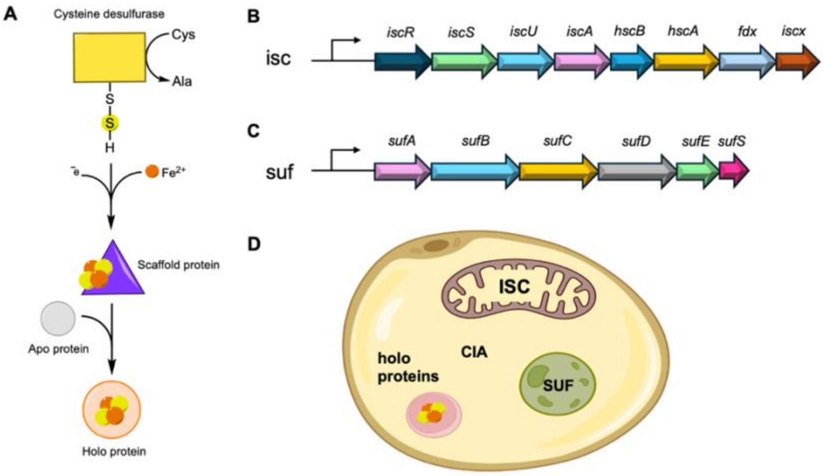


Figure 5: Overview of gene clusters and biochemical pathways involved in the synthesis of Fe-S proteins. A) overview of iron-cluster formation and assembly in eukaryotes, B) *isc* operon, C) *suf* operon, D) localization of the different machineries involved in the cofactor maturation process. Following cellular components were depicted for a simplified view of the localization of different assembly machineries: red = nucleus, green = plastid, brown = mitochondria, yellow = eukaryotic cell. Proteins that have similar functions among the two gene clusters are shown in the same colour.

The first partial reaction is similar to the one catalyzed by the ISC machinery and comprises the desulfurization of cysteine to alanine via SufS and the release of persulfate as intermediate with the only difference being the assistance of SufE which stabilizes the persulfate its further transfer takes place.[108–112] The cluster assembly is collectively executed by SufU, SufA similar to IscU and IscA, respectively, and SufB which exists in complex with SufC and SufD (Figure 5).[113–115] In addition, SufC is an ATPase which might promote the release of the newly synthesized Fe-S cluster to an apo protein thus completing the cycle of maturation during oxidative-stress conditions.[108,114]

The first reaction consists in the NTPases-assisted *de novo* assembly of an Fe-S cluster on the complex formed by Cfd1 and Nbp35.[116,117] The synthesized cluster is then transferred to apo proteins, a process aided by Nar1 and Cia1. At the same time, Nar1 is dependent on the cluster synthesis activity of Cfd1 and Nbp35 as these two are responsible to fully maturate the said transfer protein. Even though the CIA pathway appears straightforward, this sequential assembly is still dependent on the ISC machinery for glutathione (GSH) export.[118] The elucidation of the CIA machinery mechanism indicated that the biogenesis and transfer of [4Fe-4S] clusters is fulfilled solely by the CIA machinery, whereas the assembly and maturation of cytosolic [2Fe-2S] proteins depend on both mitochondrial and cytosolic synthesis systems.

1.3.3 Maturation strategies of Fe-S clusters

The recombinant production of Fe-S clusters faces several challenges due to their O_2-sensitivity. However, one bottleneck in this approach represents the fact that no machinery can fully accommodate unnaturally high amounts of the over-expressed gene, this resulting in not fully maturated Fe-S proteins. On the other hand, this can be remediated by either chemical or semi-enzymatic reconstitution with inorganic iron and sulfur sources or with recombinantly produced IscS and L-cysteine, respectively. The chemical reconstitution approach can prove slightly problematic as it has been shown to lead to impurities visible with electronic absorption spectroscopy (EAS), incomplete cluster formation or formation of Fe-S artifacts. If the geometry of the investigated Fe-S cluster is yet not established, chemical reconstitution could lead to erroneous cluster identification as this artificial state might differ from the protein's native cluster. Here the semi-enzymatic reconstitution presents a more elegant alternative which mimics the Fe-S biosynthetic pathway and can facilitate *de novo* assembly in native conformation. The handling and study of Fe-S proteins in cofactor-bound state can be a complex process depending on the number of clusters, their structure and their role. The full maturation of an Fe-S cluster can however be facilitated by using a suitable reconstitution approach.

1.4 Asp1 as an Fe-S protein

The occurrence of an Fe-S cluster as cofactor in pyrophosphatase was observed for the first time in *S. pombe* Asp1.[5] The signature motifs of the ATP-grasp domain and the histidine acid phosphatases are not known to be involved in binding of cofactors, so the existence of one in the pyrophosphatase domain was surprising. Such a versatile cofactor could fulfill a variety of biological relevant functions. It could be responsible for oligomerization events since a dimeric state has been observed for Asp1.[1] Alternatively it could serve as structural element in the protein and facilitate a certain conformation. The Fe-S cluster could even be involved in the regulation of the pyrophosphatase activity. Although Asp1 has been known to the research community for more than a decade it was only recently discovered that Asp1 is an Fe-S protein. This raises the question of how does the Asp1 Fe-S cluster exist in vivo?

Recent electronic absorption spectra of the recombinantly produced pyrophosphatase domain of Asp1 ($\text{Asp1}^{371\text{-}920}$) indicated the presence of an Fe-S cluster bound to the protein, a result which was further supported by resonance Raman and EPR spectroscopy.[5] The existence of Fe-S cluster and its stoichiometry can be quantified by collecting an electronic absorption spectrum in the visible light range that displays the specific spectral features caused by metal ions bound in highly-ordered geometrical patterns. The cofactor identified as a $[2Fe\text{-}2S]^{2+}$ cluster indicates the presence of only 0.8 Fe/monomer protein which deems cofactor maturation during gene overexpression as incomplete. It was possible to increase the stoichiometry to 2 Fe/monomer using chemical reconstitution with iron and sulfur sources under aerobic conditions. In addition, it was determined that $\text{Asp1}^{371\text{-}920}$ easily undergoes DTH-promoted degradation which makes the involvement of the pyrophosphatase domain in complex with the cofactor in redox events very unlikely. The collected EAS spectrum paired with biological data and site-directed mutagenic experiments indicated that the presence of a $[2Fe\text{-}2S]^{2+}$ cluster in the $\text{Asp1}^{371\text{-}920}$ inhibits the phosphatase activity significantly and increases the overall kinase activity.

Our group was able to isolate the Asp1 pyrophosphatase domain ($\text{Asp1}^{365\text{-}920}$) in its native state and characterize the domain focusing on the biological relevance of the Fe-S cluster and its coordination environment.[3] The expression of the *asp1*$^{+}$ gene was conducted in specialized cell strains that yielded higher amounts of cluster loaded protein based on the increase of the extinction coefficient. Furthermore, the recombinantly produced Asp1 pyrophosphatase domain was isolated for the first time anaerobically with native and fully maturated iron-sulfur cluster, as the provided inductive-coupled plasma mass spectrometry (ICP-MS) data indicated 2 Fe/monomer protein metal content in agreement with a [2Fe-2S] cluster.[3] Chemical reconstitution was shown to increase the cofactor stoichiometry and even lead to impurities visible in electronic absorption spectra. Mössbauer and X-ray absorption spectroscopic data paired with site-directed mutagenic experiments demonstrated that the presumed cystine residues are involved in ligating the Fe-S to the pyrophosphatase domain and that is very unlikely that other amino acid residues play a role in this aspect.

In addition, the biological data indicates that the Asp1 Fe-S cluster does not inhibit the pyrophosphatase domain and thus the biological function of the cofactor remains unclear.

The tertiary structure of Asp1 remains partially elusive the structure of the kinase domain in a both free and ligand-bound state is reported.[38] The presented crystallographic data indicates that in the presence of phosphorylated inositol substrates IP_6 and $5\text{-}IP_7$ the kinase domain adopts a closed conformation with an increase of ordered secondary structure elements.[38,119] To this day, the biological function of Asp1 Fe-S cluster and the regulation of the pyrophosphatase activity are not fully understood.

2 Aim of this Thesis

Structure predictions using AlphaFold were not in line with previous biochemical data and opened up new questions about the structure and function of the Fe-S cluster in Asp1. Therefore, the aim of this thesis is to gain insights into the structure, location and function of the Fe-S cluster.

The existing studies on Asp1 leave gaps in the collective knowledge about this protein as the Fe-S cluster obtained by chemical reconstitution may be different to the native cluster. Furthermore, EAS was shown to not be a suitable method for identifying the cluster form. Despite of novel coordination pattern being discovered and the presence of a high amount of cysteine residues in the pyrophosphatase domain of Asp1 it remains challenging to identify possible coordination sites based on sequence analysis and to fully characterize them with ubiquitous methods used in the research niche studying metalloproteins.[120,121] Additionally, the results of influence of the Fe-S activity are controversial and the lack structural information about this protein remains considerable even shortly after the publication of the structure of the Asp1 kinase domain (Asp1$^{1\text{-}364}$).

One first goal of this thesis was to generate optimized Asp1 constructs to produce sufficient samples of homogeneous protein for structural and spectroscopic studies. For this, an optimization of both expression and purification conditions and protocols was foreseen. The next goal of this work consisted in the characterization and narrowing of the cluster coordination environment. A differentiation of the many cysteines in the Asp1 pyrophosphatase domain between involved and not involved in cluster binding was needed. A third goal consisted in the characterization of the Asp1 Fe-S cluster using spectroscopic and biophysical methods. Here should also be clarified, if the reagents used in the experiments influence a certain cluster form or oligomeric state. A further goal of this study was to investigate if the cluster has any impact on the protein structure using circular dichroism spectroscopy. Here a differentiation among samples is needed, between tagged and untagged samples and different metal contents. At last, it was determined if the cluster plays a role in protein oligomerization using size exclusion chromatography which might later on impact the assembly of Asp1 molecules in solution.

3 Material and Methods

3.1 Chemicals and buffers

All chemicals were of analytical grade or better. All buffers were prepared using Milli-Q water.

Table 1: Overview of the used buffers and media.

Media / Buffer	Composition	pH	Use
LB medium (AppliChem)	20 g LB premixed powder, 1 L deionized water	-	Cultivation media
TB medium (AppliChem)	36 g TB premixed powder, 1 L deionized water	-	Cultivation media
2x YT medium (AppliChem)	31 g of 2x YT premixed powder, 1 L deionized water	-	Cultivation media
Equilibration buffer (Ni-NTA)	100 mM sodium phosphate, 10% glycerol, 1 L Milli-Q water	8.0	Chromatography
Wash buffer (Ni-NTA)	100 mM sodium phosphate, 10% glycerol, 30 mM imidazole, 1 L Milli-Q water	8.0	Chromatography
Elution buffer (Ni-NTA)	100 mM sodium phosphate, 10% glycerol, 300 mM imidazole, 1 L Milli-Q water	8.0	Chromatography
Equilibration buffer (GSTrap)	100 mM sodium phosphate, 1 mM DTT, 10% glycerol, 1 L Milli-Q water	7.4	Chromatography
Wash buffer (GSTrap)	100 mM sodium phosphate, 10% glycerol, 1 mM DTT, 1 L Milli-Q water	7.4	Chromatography
Elution buffer (GSTrap)	100 mM sodium phosphate, 1 mM DTT, 10% glycerol, 10 mM red. GSH, 1 L Milli-Q water	7.4	Chromatography
Equilibration buffer (TEV)	50 mM sodium phosphate, 200 mM sodium chloride, 10% glycerol, 2 mM TCEP, 1 L Milli-Q water	8.0	Chromatography
Wash buffer 1 (TEV)	50 mM sodium phosphate, 200 mM sodium chloride, 10% glycerol, 10 mM imidazole, 2 mM TCEP, 1 L Milli-Q water	8.0	Chromatography
Wash buffer 2 (TEV)	50 mM sodium phosphate, 200 mM sodium chloride, 10% glycerol, 30 mM imidazole, 2 mM TCEP, 1 L Milli-Q water	8.0	Chromatography
Elution buffer (TEV)	50 mM sodium phosphate, 200 mM sodium chloride, 10% glycerol, 300 mM imidazole, 2 mM TCEP, 1 L Milli-Q water	8.0	Chromatography
Equilibration buffer (PreScission)	1xPBS, 0.5% Triton-X-100, 5 mM DTT, 1 L Milli-Q water	7.3	Chromatography
Wash buffer (PreScission)	1xPBS, 0.5% Triton-X-100, 5 mM DTT, 1 L Milli-Q water	7.3	Chromatography
Elution buffer (PreScission)	50 mM Tris-HCl, 10 mM red. GSH, 1 L Milli-Q water	8.0	Chromatography

3.2 Strains and plasmids

All used strains are listed in Supplementary table 1. The four different variants of the Asp1$^{365-920}$ domain were generated via In-Fusion HD Cloning Kit from Takara Bio. Inc. (Shiga, Japan) and are codon-optimized for heterologous gene expression in *E. coli*. Mutant variants of the Asp1 pyrophosphatase domain GST-Asp1$^{365-920\ \mathrm{H397A}}$, GST-Asp1$^{365-920\ \mathrm{C607S}}$, GST-Asp1$^{365-920\ \mathrm{C663S}}$ and GST-Asp1$^{365-920\ \mathrm{C607S,\ C663S,\ C864S,\ C879S}}$ (GST-Asp1$^{365-920\ \mathrm{QM}}$) were cloned as part of a previous study and are not codon-optimized for *E. coli*.[3] Mutant variants GST-Asp1$^{383-920\ \mathrm{C643S}}$ and GST-Asp1$^{383-920\ \mathrm{C643S\ C644S}}$ were cloned as part of this work. The mutation of cysteine reisudes to serines is common in order to investigate the roles of cysteines in the protein structure. The sequence coding for serine with the highest usage frequency in *E. coli* (AGC) and the least number of base pairs to be exchanged was chosen as target sequence for site-directed mutagenesis. This way, these two constructs were codon optimized for heterologous expression in *E. coli*. These were originally derived from an expression vector pKM36. Site-directed mutagenesis and cloning into pKM36 via homologous recombination were conducted to generate the above-mentioned glutathione-S-transferase-tagged proteins.

3.3 Molecular biology

3.3.1 Cloning and site-directed mutagenesis

Four different constructs of the Asp1$^{365-920}$ pyrophosphatase domain were generated using In-Fusion cloning. For this procedure the In-Fusion HD Cloning Kit from Takara Bio. Inc. (Shiga, Japan) was used. In a first step the entry vector was linearized using suitable primers. A set of primers with 15 bp extensions complementary to vector ends were designed specifically to amplify each gene fragment. Next the assembly of the linearized vector and amplified gene of interest was set up in a 1:1 ratio in a total volume of 2 µL. Extraction and purification of DNA from agarose gel slices was performed using the NucleoSpin Gel and PCR Clean-up Kit from Macherey-Nagel (Düren, Germany). The procedure was carried out according to the protocol included in the kit. A volume of 0.5 µL of the In-Fusion HD Cloning Kit from Takara Bio. Inc. (Shiga, Japan) was added to the reaction mixture. The reaction was incubated for 15 min at 50 °C. Next the cells were transformed with the reaction mixture as described in section 3.3.2. Plasmid isolation from bacterial cell cultures (10 mL 2-fold yeast extract tryptone (2x YT) cultures, incubated overnight) was performed using the NucleoSpin Plasmid Mini Kit from Macherey-Nagel (Düren, Germany). The procedure was carried out according to the protocol included in the kit. Concentrations of isolated DNA samples were determined via absorption at 260 nm with the NanoDrop 2000c of Thermo Fisher Scientific (Waltham, USA).

3.3.2 Heat-shock transformation of chemically competent cells

Stocks of chemically competent cells of different strains were stored at -80 °C until needed. A 50 µL cell suspension of competent cells was incubated with 2.5 µL (10-100 ng) plasmid solution on ice for

30 min. Next the cells underwent a heat-shock for 60 s at 42 °C and were further chilled on ice for 2 min. Cell recovery was promoted by adding 950 µL 2x YT media, without antibiotic, under sterile conditions and the cell suspension was incubated for 1 h at 37 °C and 300 rpm. The pellet and the supernatant were separated by centrifugation at 6,000 x *g* for 5 min. A volume of 950 µL of the supernatant was discarded, the pellet was resuspended in the residual 50 µL of 2x YT medium and spread over agar plates containing ampicillin (100 µg/mL).

3.3.3 Serial dilution patch testing

The constructs pJR2-3XL-*asp1* and pJR2-3XL-*asp1*$^{\text{I808D}}$ for expression in *S. pombe* were generated in a previous study.[47] PCR fragments were generated by site-directed mutagenesis using a QuikChange II site-directed mutagenesis kit from Agilent Technologies and cloned into pJR2-3XL via homologous recombination in *S. cerevisiae*.[122,123] The quadruple mutant pJR2-3XL-*asp1*$^{\text{365-920, C607S, C663S, C864S, C879S}}$ or pJR2-3XL-*asp1*$^{\text{QM}}$ was obtained by digestion of pJR2-3XL with the restriction enzymes NotI and PstI from New England Biolabs (Ipswich, USA), followed by leucine selection.[122,123]

3.4 Electrophoretic methods

3.4.1 Agarose gel electrophoresis

A solution of 1% agarose containing 0.12‰ Gel Red of the firm Biotum Inc (Hayward, USA) as a UV-fluorescent nucleic acid dye was poured in a horizontal gel tray and a comb added for the formation of gel chambers. The electrophoresis was carried out in a chamber of Peqlab Biotechnology GmbH (Nürnberg, Germany) with a power supply of VWR International (Darmstadt, Germany) at 180 V (cont.), 120 V for 55 min. The GeneRuler 1 kb ladder (10; 8; 6; 5; 4; 3.5; 3; 2.5; 2; 1.5; 1; 0.75; 0.5; 0.25 kb) of Fermentas (St. Leon-Rot, Germany) was applied on gel as well and used as a DNA marker for the identification of fragment size. The gel documentation was conducted using the Gel iX20 Imager system of Intas Science Imaging Instruments (Göttingen, Germany).

3.4.2 Denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Protein purification and expression was analyzed via Denaturing Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE). The gels were cast using the Bio-Rad Mini Protean Tetra System from Bio-Rad Laboratories, Inc. (Berkeley, USA). First the separating gel was polymerized because of its higher concentration and was overlaid with isopropanol to prevent desiccation. Next the stacking gel was cast over the separating gel and a sample comb added. The PageRuler Plus Prestained Protein Ladder was used as standard (250, 130, 100, 70, 55, 35, 25, 15, 10 kDa) and a volume of 3 µL was loaded onto the PAGE. Separating gel (15%, 4 gels): 6.7 mL 30% (v/v) acryl amide/bisacryl amide (37.5:1), 6.7 mL 3x gel buffer, 2.0 g glycerol, 4.7 mL H_2O, 100 µL 10% (w/v) ammonium persulfate (APS) solution, 100 µL TEMED. Stacking gel (4%, 4 gels): 1.6 mL

30% (v/v) acryl amide/bisacryl amide (37.5:1), 3.0 mL 3x gel buffer, 7.4 mL H_2O, 100 µL 10% (w/v) APS solution, 100 µL TEMED.

All samples were diluted to a near identical concentration with a volume of 45 µL. After the addition of 15 µL 6x SDS-loading buffer, the samples were denatured by heating at 99 °C for 10 min. The gels were loaded in electrophoretic chambers which were filled up with 1x TGS buffer. The PAGE was run with 35 mA per gel for 90 min. The gels were stained with *Coomassie G-250* solution for 30 min and destained overnight in Milli-Q H_2O.

For expression controls the pellet samples were prepared by transferring 45 µL of the lysed cell suspension to a 1.5 mL Eppendorf tube. Cell pellet and supernatant were separated by centrifugation at 11,000 x *g* for 5 min. The supernatant was transferred to a different 1.5 mL tube and mixed with 15 µL of 6x SDS-loading buffer. The pellet was resuspended in 45 µL of Milli-Q H_2O and 15 µL of the 6x SDS-loading buffer were added to the homogenous sample as well.

3.5 Heterologous protein production

3.5.1 Expression tests

Small scale expression of all newly cloned constructs was conducted while varying a multitude of parameters in order to identify optimal expression conditions. The expression tests were performed in 250 mL of culture media in Erlenmeyer flasks with baffles. The starting culture was prepared in a total volume of 50 mL in 2x YT medium. A volume of the starting culture was added in a 1:100 ratio and the small culture was incubated at 37 °C and 120 rpm until the optical density measured at 600 nm (OD_{600}) reached values between 1.0 and 1.6. Afterwards 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce gene overexpression. Expression tests were conducted at 20 °C, 25 °C and 30 °C for 3 h and overnight after induction. A 1 mL culture sample was collected for SDS-PAGE analysis after 3 h and the next morning. Different culture media that were tested comprised lysogeny broth (LB), terrific broth (TB), and 2x YT media. Furthermore, different bacterial strains cell strains (*E. coli* BL21(DE3) *ΔiscR*, *E. coli* NiCo21 (DE3) and *E. coli* BL21(DE3)), were tested to identify a suitable one leading to high protein yields and purity. The cultivation medium that was used for starting cultures was either LB or 2x YT medium. Control samples before and after induction were collected for analysis by SDS-PAGE and were normalized to an $OD_{600} = 1$ during the sample preparation for them to be comparable.

3.5.2 Aerobic expression and protein isolation of Asp1 variants

Starting cultures of containing the plasmid were prepared in 10 mL 2YT-medium supplemented with ampicillin (100 µg/mL) using colonies from a fresh agar plate and were incubated at 160 rpm and 37 °C overnight. Large cultures in 2YT-medium were inoculated from starting cultures in a 1:100 ratio. The medium was supplement with ampicillin (100 µg/mL), and ferric ammonium citrate (2 mM final concentration). Cells were cultivated aerobically at 120 rpm and 37 °C until the OD_{600} reached values between 1.2 and 1.6. Gene expression was induced by adding 1 mM IPTG. To facilitate Fe-S metabolism and a maximal assembly of cluster in vivo 2 mM L-cysteine and 25 mM sodium fumarate were added just after induction was started. The cultures were cultivated for 3 h after IPTG induction. Cell harvest was conducted for 10 min at 5,000 x *g* and 4 °C. Cell pellets were then stored at -20 °C or lysed directly.

For cell lysis, cells were resuspended in 100 mM sodium phosphate buffer, 10% glycerol, pH 7.4. etylenediaminetetraacetic acid (EDTA) free cOmplete protease inhibitor cocktail tablets from Roche (Basel, Switzerland) and 10 ng DNase from Roche (Basel, Switzerland) per g cell pellet were added to the cell suspension to prevent protease activity and to promote cleavage of any DNA traces activity during cell lysis, respectively. After vortexing at RT, the homogenous suspension underwent sonication using a device from Bandelin electronic (Berlin, Germany) for 20 min with an amplitude of 60%, a pulse frequency of 1 s every 3 s on a VS70/T sonotrode. The lysates were clarified by centrifugation at 40,000 x *g* for 30 min. During this procedure, a Protino-GSTrap HP column with a bed volume of 5 mL from Macherey-Nagel (Düren) was connected to an ÄKTAstart chromatography system from Cytiva (Malborough, USA), and equilibrated with 5 column volumes (CV) of 100 mM sodium phosphate buffer, 10% glycerol, pH 7.4 at a flowrate of 1 mL/min. Next, the lysate was loaded onto the column at the same flowrate to enable adequate time for the affinity binding to occur. The interaction between GST-tagged proteins and the column as stationary phase containing immobilized glutathione is slow and higher flowrates than 1 mL/min were shown to inhibit full-capacity binding. After the flow-through fraction was collected the column was washed with 10 CV of 100 mM sodium phosphate buffer, 10% glycerol, pH 7.4 at 2 mL/min to remove improbable unspecific binding proteins. In the last step the GST-tagged protein was eluted using 5 CV of 100 mM sodium phosphate buffer, 10% glycerol, 10 mM reduced glutathione (GSH), pH 7.4.

The cell lysis and protein purification of His-tagged constructs of Asp1$^{365\text{-}920}$ were conducted in the same manner as mentioned above with the only exception being the chromatography column which was used. A Protino-Ni-NTA column with a bed volume of 5 mL from Macherey-Nagel (Düren) was used for the immobilized metal affinity chromatography (IMAC) to isolate His-Asp1$^{365\text{-}920}$ and His-Asp1$^{383\text{-}918}$. In this type of column material, a Ni^{2+} ion with six coordination sites undergoes a chelate complex with nitrilotriacetic acid (NTA) which is immobilized on a matrix surface. The Ni^{2+} ion is coordinated by three negatively charged oxygen atoms and one partially negatively charged N atom of NTA leaving two unoccupied coordination sites exposed to various molecules in solution. If any other molecules with higher affinity are present the two unoccupied coordination sites of a Ni^{2+} ion are occupied by two H_2O molecules. Histidine residues from a His-tags or from histidine-rich proteins are the probable candidates at occupying the last two coordination sites of one Ni^{2+} ion. In comparison to GST affinity chromatography is the interaction between histidine residues and Ni^{2+} ions much faster than the one between the GST-tag and reduced glutathione. This allows the loading of the lysate and washing of the column at higher flowrates. In case of His-Asp1$^{365\text{-}920}$ and His-Asp1$^{383\text{-}918}$ the lysate was loaded at

1.5 mL/min and the wash step was conducted at 2 mL/min with 100 mM sodium phosphate buffer, 10% glycerol, 30 mM imidazole, pH 8.0. The elution was performed with 5 CV of 100 mM sodium phosphate buffer, 10% glycerol, 300 mM imidazole, pH 8.0. All fractions from the different chromatographic steps were collected via fractionation. SDS-PAGE samples were collected from all fractions for later analysis. At last buffer exchange was conducted using centrifugation tubes to remove glutathione or imidazole from the solution. Protein solutions were then either aliquoted and stored at -80 °C or further analyzed using spectroscopic methods.

3.5.3 Production of TEV protease

A starting culture with 100 mL of LB medium supplemented with ampicillin (100 µg/mL) was incubated at 37 °C and 160 rpm overnight. The next day a main culture of 4 L LB medium was supplemented with ampicillin (100 mg/mL) and inoculated with 10 mL of the starting culture. The main culture was incubated at 37 °C and 120 rpm until the OD_{600} reached 0.8 and then induced with 1 mM IPTG. The culture was further incubated at 16 °C overnight post induction. Cell harvest was conducted at 5,000 x *g* and 4 °C for 10 min. The cell pellets were stored at -20 °C until needed.

The cell pellet was thawed and resuspended in 120 mL of 50 mM sodium phosphate buffer, 200 mM sodium chloride, 10% glycerol, 2 mM TCEP, pH 8.0. EDTA-free cOmplete protease inhibitor cocktail tablets from Roche (Basel, Switzerland) and 10 ng DNase from Roche (Basel, Switzerland) per g cell pellet were added to the cell suspension to prevent protease activity and to promote cleavage of any DNA traces activity during cell lysis, respectively. After vortexing at RT, the homogenous suspension underwent sonication using a device from Bandelin electronic (Berlin, Germany) for 8 min with an amplitude of 30%, a pulse frequency of 1 s every 3 s on a MS72 sonotrode. The lysates were clarified by centrifugation at 40,000 x *g* for 30 min. During centrifugation, a Protino-Ni-NTA column with a bed volume of 5 mL from Macherey-Nagel (Düren) was connected to an ÄKTAstart chromatographic system from Cytiva (Malborough, USA), and equilibrated with 4 CV of 50 mM sodium phosphate buffer, 200 mM sodium chloride, 10% glycerol, 2 mM TCEP, pH 8.0 at a flowrate of 1.5 mL/min. Next, the supernatant was applied to the column at the same flowrate after which a manual gradient with imidazole was applied to remove unspecific binding proteins from the column. The buffer system remained the steps, but the final imidazole concentration was increased to 10 mM and 30 mM imidazole for 10 and 4 CV, respectively. Afterwards, the His-tagged Tobacco-Etch-Virus (TEV) protease (HHHHHHHHHH) was eluted with 50 mM sodium phosphate buffer, 200 mM sodium chloride, 10% glycerol, 2 mM TCEP, pH 8.0. All fractions from the different chromatographic steps were collected via fractionation. SDS-PAGE samples were collected from all fractions for later analysis. At last buffer exchange was conducted using centrifugation tubes to remove the imidazole from the solution. Protein solutions were aliquoted and stored at -80 °C.

3.5.4 Production of PreScission protease

A starting culture with 100 mL of LB medium supplemented with ampicillin (100 µg/mL) and 1% glucose was incubated at 37 °C and 160 rpm overnight. Next main cultures prepared with LB medium, and the supplements mentioned above were inoculated with the starting culture to obtain a final OD_{600} of 0.1 per L medium. The main culture was incubated at 37 °C and 120 rpm until the OD_{600} reached values between 0.6 and 0.8 and then induced with 1 mM IPTG. The culture was further incubated at 37 °C for 5 h. Cell harvest was conducted at 7,000 x *g* and 4 °C for 15 min and the pellet was subsequently washed with 30 mL of 1x PBS. The cell suspension was again centrifuged at 7,000 x *g* for 10 min and pellets were stored at -20 °C until needed.

The cell pellet was resuspended at 4 °C in lysis buffer with 4 mL of 1x PBS, 0.5% Triton X-100, 5 mM 1,4-dithiothreitol (DTT), pH 7.3 per g pellet. EDTA-free cOmplete protease inhibitor cocktail tablets from Roche (Basel, Switzerland) and 10 ng DNase from Roche (Basel, Switzerland) per g cell pellet were added to the cell suspension to prevent protease activity and to promote cleavage of any DNA traces activity during cell lysis, respectively. Furthermore, lysozyme was added to the lysis mixture in a 20 µg/mL dilution ratio. After vortexing at RT, the homogenous suspension underwent sonication using a device from Bandelin electronic (Berlin, Germany) for 8 min with an amplitude of 60%, a pulse frequency of 20 s on a VS70/T sonotrode. The cell suspension was then incubated on ice and under shaking for 30 min until the sonication procedure was repeated once more. The lysates were clarified by centrifugation at 40,000 x *g* for 45 min and the supernatant was decanted into a falcon tube. During centrifugation, a Protino-GSTrap column with a bed volume of 5 mL from Macherey-Nagel (Düren) was connected to an ÄKTAstart chromatographic system from Cytiva (Malborough, USA), washed with 6 CV of Milli-Q H_2O and equilibrated with 12 CV of 1x PBS, 0.5% Triton X-100, 5 mM DTT, pH 7.3 at a flowrate of 1.0 mL/min. Next, the supernatant was applied to the column at the same flowrate after which the column was washed with 40 CV of 1x PBS, 0.5% Triton X-100, 5 mM DTT, pH 7.3. The GST-tagged HRV-3C-Protease (PreScission protease) was eluted with 5 CV of 50 mM Tris-HCl buffer, 10 mM reduced glutathione, pH 8.0. All fractions from the different chromatographic steps were collected via fractionation. SDS-PAGE samples were collected from all fractions for later analysis. After SDS-PAGE analysis the elution fractions containing the protease were collected and the volume was reduced to approx. 8 mL using concentration tubes. Next, the protein solution was dialyzed twice against 2 L of 50 mM Tris-HCl buffer, 150 mM sodium chloride, 20% glycerol, 10 mM EDTA, 1 mM DTT, pH 8.0 for 3 h using a dialysis membrane with a MWCO of 1.000 kDa. Protein solutions were aliquoted and stored at -80 °C.

3.5.5 Anaerobic cell lysis and protein isolation of Asp1 variants

Most of the Asp1 purification were carried out under strictly anaerobic conditions. The solubilization of the cell pellet and additives for the cell lysis was conducted in the same manner as for the aerobic purification but inside an anaerobic glove box from Coy Laboratory Products (Grass Lake, USA) filled with 98% N_2 and 2% H_2. The falcon tube containing the bacterial suspension was then closed and transferred out of the glovebox and incubated under argon flow for several minutes to guarantee anaerobic conditions throughout the sonication procedure. Afterwards the cell suspension was transferred inside the glovebox, and to fresh centrifugation tubes already incubated under anaerobic conditions to remove any possible oxygen sources. The subsequent cell harvest was performed as mentioned in the previous chapter and the clarified cell lysates were transferred one more time to the anaerobic chamber where the affinity chromatography followed.

3.6 Size exclusion chromatography

Size-exclusion chromatography (SEC) was used to determine the oligomerization state of different Asp1$^{365\text{-}920}$ variants (analytical) or as a second purification step after affinity chromatography or chemical reconstitution (preparative). This method consists in the separation of biomolecules by hydrodynamic radius and can be used in an analytical or preparative manner or for desalting as the sample composition does not affect the resolution. The chromatography resin, consisting of a porous matrix of spherical particles of different sizes, is defined as the stationary phase. The buffer system used during the chromatography run is the liquid phase.

Two different columns were used for analytical SEC. The HiLoad 16/600 Superdex 75 pg with a bed volume of 120 mL from Cytiva (Malborough, USA) was used for His-Asp1$^{365\text{-}920}$ and His-Asp1$^{383\text{-}918}$ (66.0 kDa) and the HiLoad 16/600 Superdex 200 pg with a bed volume of 120 mL from Cytiva (Malborough, USA) was used for GST-Asp1$^{365\text{-}920}$,GST-Asp1$^{383\text{-}918}$ (90.9 kDa) and all related states due to the considerable size of the GST fusion protein. The columns were equilibrated with 2 CV of the running buffer at 1 mL/min and max. 0.5 MPa. The course of the chromatography was observed by monitoring the absorbance of 280 nm. SDS-PAGE samples were prepared from the collected fractions for later analysis. Afterwards the fractions in which the target protein was present were combined. A calibration using the low molecular weight kit (LMW Kit) from Cytiva (Malborough, USA) was used to determine the retention volumes of proteins of known size. Based on these values a regression curve was plotted, and a calibration equation could be established, which enabled the determination of exact experimental hydrodynamic radius of the biomolecules present in solution.

Preparative SEC was conducted using the HiLoad 16/600 Superdex 200 pg with a bed volume of 120 mL from Cytiva (Malborough, USA) after affinity chromatography as a further purification step or after chemical reconstitution in order to remove Fe-S aggregates from solution.

3.7 Proteolytic digest

An affinity tag (GST-, His-tag) is needed in order to be able to isolate proteins via affinity chromatography. However, this can be removed to facilitate different experiments subsequent to protein purification. Cleavage sites for the TEV and PreScission protease are encoded on the expression vectors used in this study, pET16b-TEV and pGEX-6-P-2, between the affinity tag and the sequence of the Asp1 pyrophosphatase domain, respectively. For His-tagged constructs the concentration of the isolated TEV protease was determined using UV/Vis spectroscopy and then added in a 1:20 ratio (TEV protease : protein). The solution was further supplemented with 1 mM DTT and 0.5 mM EDTA, pH 8.0. The digested sample was incubated overnight at 4 °C under shaking conditions. For GST-tagged protein variants the samples were incubated with PreScission protease in a 1:10 ratio (PreScission protease : protein) overnight at 4 °C under shaking conditions. The cleavage buffer for the PreScission protease consisted of 50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.0. The previously digested sample was loaded on an equilibrated Protino Ni-NTA or GSTtrap HP column depending on the type of affinity tag at 1 mL/min and the flow-through was collected via fractionation. Elution of the column-bound affinity tag and protease was conducted at 1 mL/min with 300 mM imidazole or 10 mM red. GSH in buffer, the eluate being collected via fractionation as well. SDS-PAGE samples were collected from all fractions for latter SDS-PAGE analysis.

3.8 Spectroscopic methods

3.8.1 Electronic absorption spectroscopy

EAS was used to measure the iron-sulfur content of His-Asp1$^{365-920}$, GST-Asp1$^{365-920}$, His-Asp1$^{383-918}$, GST-Asp1$^{383-9218}$, GST-Asp1$^{365-920\ C643S}$, GST-Asp1$^{365-920\ C607S}$, GST-Asp1$^{365-920\ C663S}$, GST-Asp1$^{365-920\ C864S}$, GST-Asp1$^{365-920\ H397A}$, GST-Asp1$^{383-918\ C868S}$, GST-Asp1$^{365-920\ C643S\ C644S}$, GST-Asp1$^{365-920\ QM}$. Electronic absorption spectra were recorded using a Cary-60 spectrophotometer from *Agilent Technologies* (Ratingen, Germany) with 1 nm bandwidth, a scanning speed pf 120 nm/min and a 1-cm-path-length quartz cuvette at RT. For Fe-S oxygen sensitivity assays a spectrum was recorded every 30 min and the cuvette was purged with oxygen between measurements.

3.8.2 Mössbauer spectroscopy

Mössbauer spectra were recorded on a conventional spectrometer with alternating constant acceleration cooled with an Oxford Instruments Variox cryostat, using a ^{57}Co/Rh (1.8 GBq) γ-source as part of a collaboration. Samples were measured at 80 K and with no applied magnetic field. Isomer shifts are relative to iron metal at 300 K. Spectra were simulated and fitted using Lorentzian quadrupole doublets using in house software.

3.9 Chemical reconstitution

Aerobically or anaerobically isolated protein in a 100-500 µM concentration range were incubated on ice with 5 mM DTT under anaerobic conditions for 30 min and then supplemented with 5x molar excess of ferric chloride ($FeCl_3$) and 5x molar excess sodium sulfide (Na_2S). First the total required volume of the Fe source was added bitwise in 5 steps. The protein sample was further incubated for 30 min on ice before the addition of the sulfur source proceeded in an identical manner. Chemically reconstituted samples were incubated overnight at 4 °C at 300 rpm to facilitate maximal Fe-S assembly inside the protein. The next day desalting was conducted to remove unspecifically bound Fe-S aggregates using a HiLoad 16/600 Superdex 200 pg column from Cytiva (Malborough, USA).

For Mössbauer spectroscopy, aerobically purified GST-Asp1$^{365\text{-}920}$ and Asp1$^{365\text{-}920}$ were reconstituted following the described protocol but with ^{57}Fe instead of ferric chloride. The desalting of the Mössbauer samples was conducted using a HiLoad 26/600 Superdex 200 pg column from Cytiva (Malborough, USA).

For cluster transformation assays the chemical reconstitution was performed according to the described protocol but with 10x molar excess of ferric chloride ($FeCl_3$) and sodium sulfide (Na_2S).

For chemical reconstitution comparative assays, the procedure was followed as described above, however ferric ammonium citrate and reduced glutathione were used as an additional Fe source and reductant, respectively.

3.10 Determination of protein concentration via EAS

3.10.1 Determination of protein concentration via EAS

Protein concentration of different samples could be routinely determined using electronic absorption spectroscopy due to the absorption features of aromatic amino acids at 280 nm. A spectrum of a sample was recorded at a suitable dilution to generate absorption values under 1 and thus guarantee linearity according to the Lambert-Beer law. The same equation was then used for the calculation of protein concentration under consideration of the dilution factor. For all comparisons of EAS data, a spectrum of an undiluted sample was additionally recorded to properly observe Fe-S signals at the respective wavelengths.

For comparison purposes the spectra of undiluted samples were normalized on the extinction coefficient by dividing all absorption values of the single spectra by the value calculated protein concentration.

3.10.2 Determination of protein concentration via Bradford-assay

Measurements of protein concentration for inductive coupled plasma mass spectrometry (ICP-MS) samples were conducted using the Bradford-assay. This colorimetric procedure is based on the absorption shift of Coomassie G 250, a triphenyl methane dye, in acidic settings from 465 nm to 595 nm. A volume of 800 µL of the samples were mixed with 200 µL of Bio-Rad Protein assay dye reagent concentrate and incubated for 5 min. The absorption of each sample was measured at 595 nm and the protein concentration determined using a calibration curve (measurement of various bovine serum albumin (BSA) solution with concentrations between 0-150 µg/mL).

3.11 Determination of metal content

Protein samples were precipitated with 3% trace-metal grade nitric acid before analysis. The samples were centrifuged for 10 min at 11,000 x *g* to remove sedimented protein. The supernatant was transferred to a metal-free centrifugation tube from *VWR* (Radnor, PA, USA). The Fe content of the protein was determined by ICP-MS using an Agilent 7500ce ICP-MS instrument from Agilent Technologies (Ratingen, Germany) in the Universität zu Köln. Samples were measured in biological triplicates and if possible, always at 10 µM protein concentration.

Due to volume limitations of the reductive chelation assay and concentration limitations of reconstituted samples, ICP-MS samples resulting from these experiments were measured at lower concentrations than 10 µM, this being considered for the set-up of the experimental measurements.

3.12 Biophysical methods

3.12.1 Reductive chelation of Fe-S clusters

Chelation of ferrous and ferric iron forms from GST-Asp1$^{365\text{-}920}$ variants was conducted by adding a 10-, 100- or 1000-fold excess of chelator (EDTA or bipyridyl, BIPY) to the stock protein solution. Next a few crystals of sodium dithionite (NaDT) were added to ensure a complete reduction of Fe($^{3+}$) to Fe($^{2+}$). Protein samples were then incubated at 4 °C for 4 h and halfway through another sodium dithionite crystals were added. Desalting was conducted using a HiLoad 16/600 Superdex 200 pg column in 100 mM sodium phosphate buffer, 10% glycerol, pH 7.4. Sample for EAS and ICP-MS were prepared to verify the success of the reductive chelation experiments.

3.12.2 CD Spectroscopy

Information regarding the secondary structure of GST-Asp1$^{365\text{-}920}$ and His-Asp1$^{365\text{-}920}$ in different states was collected via circular dichroism spectroscopy, a method that rests upon the difference between left and right circular polarized light when hitting an optically active molecule. Distinct secondary structure elements contained by proteins leads to characteristic circular dichroism (CD) spectra in far-UV range. The analysis of this data was conducted using the webserver (https://bestsel.elte.hu) BeSTSel, which delivered an estimation of the distribution of the secondary structure elements as α-helices, antiparallel and parallel β-sheets, turns and random coils.[124–126]

The measurements for the determination of the secondary structure elements were conducted using a JASCO J-815 CD spectropolarimeter (Minneapolis, USA). Samples were prepared in a concentration range of 1.5 to 8 µM with a total volume of 350 µL. The recorded parameters during the measurements are the ellipticity in mdeg, the absorbance in mAU and voltage in V in a wavelength range from 190 to 260 nm. The blank measurement was conducted using 10 mM sodium phosphate buffer, pH 7.4, this being the sample preparation buffer. All parameters are summarized in table

The measurements for the determination of the melting temperature were conducted using a JASCO J-815 CD spectropolarimeter and a temperature interval. The experimental data is depicted as ellipticity in mdeg and for it to be comparable it must be normalized on the most common and repetitive structural feature found in proteins. Thus, the data is normalized on the mean residue weight (MRW). This parameter is calculated as depicted in Equation 1 with *M* being the molecular weight of the protein (Da) and *N* the number of amino acids. The mean residue ellipticity at a certain wavelength can be calculated as shown (Equation 1), where θ is the observed ellipticity (degree) at wavelength λ, d is the thickness (cm) and c the concentration (g/mL).

Equation 1 $$MRW = \frac{M}{N-1}$$

Equation 2 $$[\theta]_{mrw,\lambda} = \frac{MRW \cdot \theta_\lambda}{10 \cdot d \cdot c}$$

The data was processed by subtracting the background of the buffer first and then the experimental values were converted to mean molar residue ellipticity. The minimum of the CD signal at 20 °C was identified and all values at that wavelength were normalized. Last, the melting temperature was determined by plotting the normalized data as a saturation curve and reading out the temperature value at the graphical turning point (0.5).

Table 2: Overview of the parameters used in the CD measurements for the determination of secondary structure.

Parameter	**Settings**
Horizontal axis	Wavelength [nm]
Vertical axis 1	CD [mdeg]
Vertical axis 2	HAT [V]
Vertical axis 3	Absorbance [mAU]
Start wavelength	260 nm
End wavelength	190 nm
Data interval	
Temperature	20 °C
D.I.T	1 s
Bandwidth	5 nm
Scan rate	50 nm/min
Baseline correction	none
Accumulations	20
Mode temperature gradient	
CD detector	PMT

Table 3: Overview of the parameters used in the CD measurements for the determination of the melting temperature.

Parameter	**Settings**
Horizontal axis	Wavelength [nm]
Vertical axis 1	CD [mdeg]
Vertical axis 2	HAT [V]
Vertical axis 3	Absorbance [mAU]

Start wavelength	260 nm
End wavelength	190 nm
Data interval	
Start temperature	20 °C
End temperature	90 °C
Temperature interval	0.5 °C
Heating rate	1 °C/min
D.I.T	1 s
Bandwidth	5 nm
Scan rate	100 nm/min
Baseline correction	none
Accumulations	10
Mode temperature gradient	
CD detector	PMT

3.12.3 Reduction and oxidation experiments

For the reduction and oxidation assays protein samples were prepared with a concentration of 133 µM, which results in an optimal signal to noise ratio during the collection of EAS data. Three of the samples were treated with the reducing agents NaDT, DTT and GSH. Equivalents of the reagents were added stepwise, after which the samples were incubated for 30 min on ice. Next a sample for CD spectroscopic measurements was collected and the sample was measured using EAS. The fourth sample was incubated under aerobic conditions for the ligated metal ions to undergo oxidation by atmospheric oxygen. One equivalent equaled an incubation of the reaction tube with open lid for 30 min, after which air bubbles were pipetted actively in the sample. Prior to the collection of EAS data, a sample for CD spectroscopic experiments was acquired. A maximum of 16 equivalents were added to the four samples.

3.12.4 Analytical ultracentrifugation

Analytical ultracentrifugation measurements were conducted using an analytical ultracentrifuge Proteome Lab XL-A (Beckman-Coulter). Experiments included 10 µM of chemically reconstituted protein with 10-fold excess of iron and sulfur sources. The measurements were performed using a standard double stor cell (titanium) with an optical path length of 12 mm using an An-60Ti rotor. The temperature was set to 20 °C and the speed adjusted to 60,000 rpm which correspond to 262,000 x *g*. Data analysis was conducted using a continuous distribution Lamm equation model, c(s), implemented in the Software Sedfit (version 16p25; https://sedfitsedphat.github.io).[127]

4 Results

4.1 Structure prediction and analysis using AlphaFold2

The release of AlphaFold2 enabled accurate structure prediction based on protein sequence. A breakthrough consisted in the recent launch of AlphaFold2, a platform based on artificial intelligence and able to predict three-dimensional structures of proteins with atomic-level accuracy from amino acid sequences. The release of more than 200 million protein structure predictions aided researchers tremendously in elucidating the tertiary structures of proteins with both common and unconventional architectures, among which some metalloproteins could be found.[128,129] Although very advantageous, this software has a major flaw regarding proteins binding cofactors or ligands. Often the existence of such compounds in close proximity to proteins can lead to structural changes at atomic levels or can occur if at all upon substrate binding. The calculations behind the AlphaFold2 structure prediction software do not consider the existence of such ligands and cofactors which most probably can be misleading if a strategy for structure determination is solely based on such a prediction. Nevertheless, the use of such prediction model can prove advantageous.

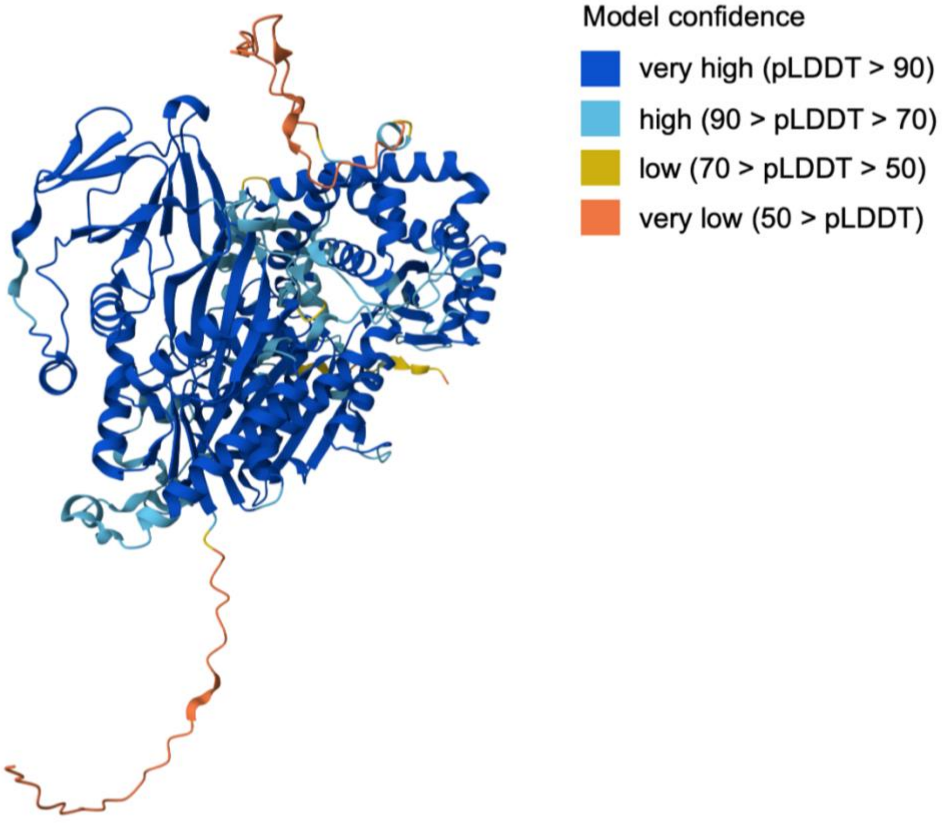


Figure 6: Predicted structural model of *S. pombe* Asp1 by AlphaFold2.[130,131] The predicted tertiary structure is depicted in four colors according to the model confidence legend. AlphaFold2 produces a per residue model confidence score (pLDDT) between 0 and 100. Regions below 50 pLDDT may be unstructured in isolation. Regions in the structure of the protein are colour coded corresponding to the model confidence. Dark and light blue correlate with very confident and confident predictions regarding the structural elements. Regions depicted in yellow and orange are not certain to occur in the presented structure and are attributed the lowest confidence scores.

The codon-optimized Asp1$^{365-920}$ and Asp1$^{383-918}$ constructs, harbouring only the pyrophosphatase domain, were validated using structure prediction via AlphaFold2. A structure was predicted based on the sequence of full-length *S. pombe* Asp1. The coloring in dark blue shows that most of the predicted structure for Asp1 was calculated with confidence, since the corresponding confidence model values lies over 90 (Figure 6). This hints towards the calculations on which the model is based to very accurate and close to the real Asp1 structure. Certain areas between secondary structure elements (mostly helices) with high confidence are depicted in light blue, which indicates a lower confidence value. Only isolated elements are illustrated in yellow or orange, in which case the model confidence values lie under 50. Interestingly, these are unstructured regions, such as loops or turns, that might not be relevant to the protein structure. The attribution of such low confidence values to possibly unstructured regions is not common, since these appear to be flexible based on the calculations which AlphaFold2 predictions are based on. However, this value was deemed too low to be accepted as part of the structure analysis of Asp1. Only the areas over 70 were taken into consideration. In summary, the predicted Asp1 structure is overall calculated with high confidence and can be relied upon based on the confidence model.

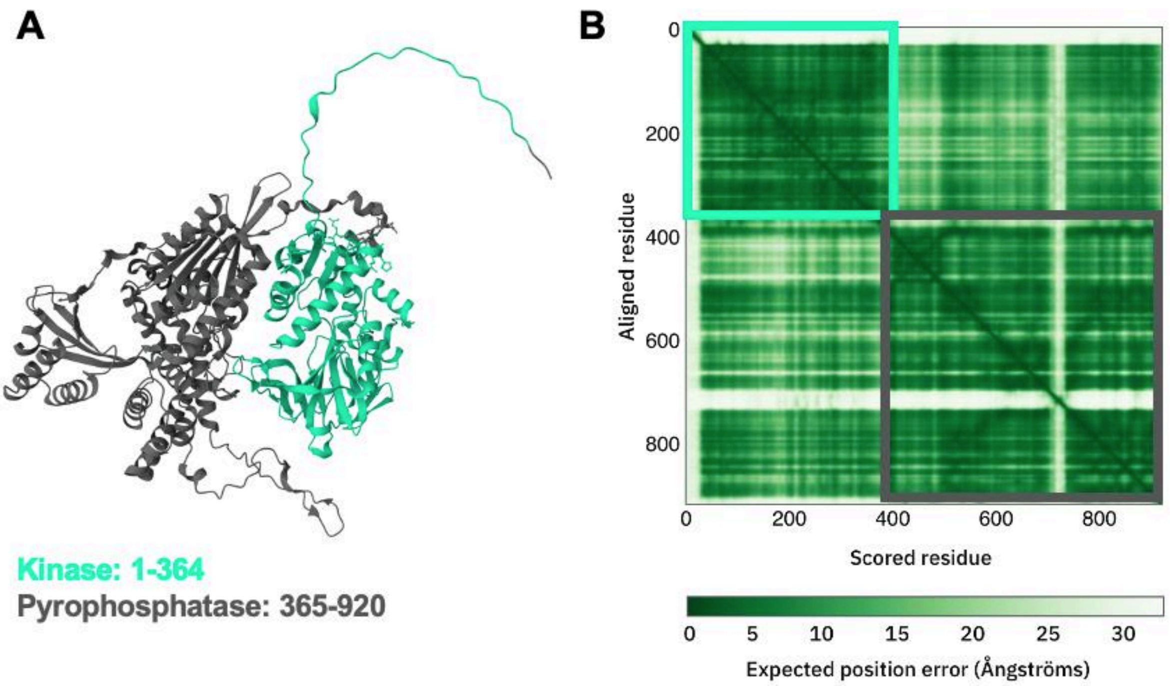


Figure 7: Separation of *S. pombe Asp1* in its two domains.[130,131] A) Predicted structure of full-length Asp1 with its two domains according to the predicted aligned error (PAE) diagram. B) The PAE diagram for *S. pombe* Asp1. Instead, the shade of green indicates the expected distance error in Å, ranging from 0 Å to an arbitrary cut-off of 31 Å. The colour at (x, y) corresponds to the expected distance error in the residue x's position when the predicted and the true structures are aligned on residue y. A dark green tile corresponds to a good prediction (low error), whereas a light green tile indicates poor prediction (high error).

The model was accompanied by a predicted aligned error (PAE) diagram, which measures the confidence in the relative position of two residues within the predicted structure, indicating the reliability of relative positions and orientations of different domains (Figure 7).[130,131] In this diagram, two of the darkest tiles indicate the two domains of Asp1 with a low expected position error. This hints towards the predicted structure and the implied division of the domains to be accurate. The only white areas in the PAE diagram correspond to loops and other areas of the Asp1 structure which were deemed as structurally irrelevant by other secondary structure prediction tools (Figure 9). Here the expected position error is much higher, this being typical for potentially unstructured elements. Thus, the separation of the Asp1 domains at residue 364 was correct and the constructs of the Asp1 pyrophosphatase domain (Asp1$^{365\text{-}920}$ and Asp1$^{383\text{-}918}$) were validated for their further use in spectroscopic and chromatographic experiments. The Asp1 constructs were validated by checking if the Asp1 structure prediction model aligns with the domain separation proposed in the protein constructs. The PAE diagram confirmed that the pyrophosphatase domain comprises amino acids 365 to 920 with high confidence.

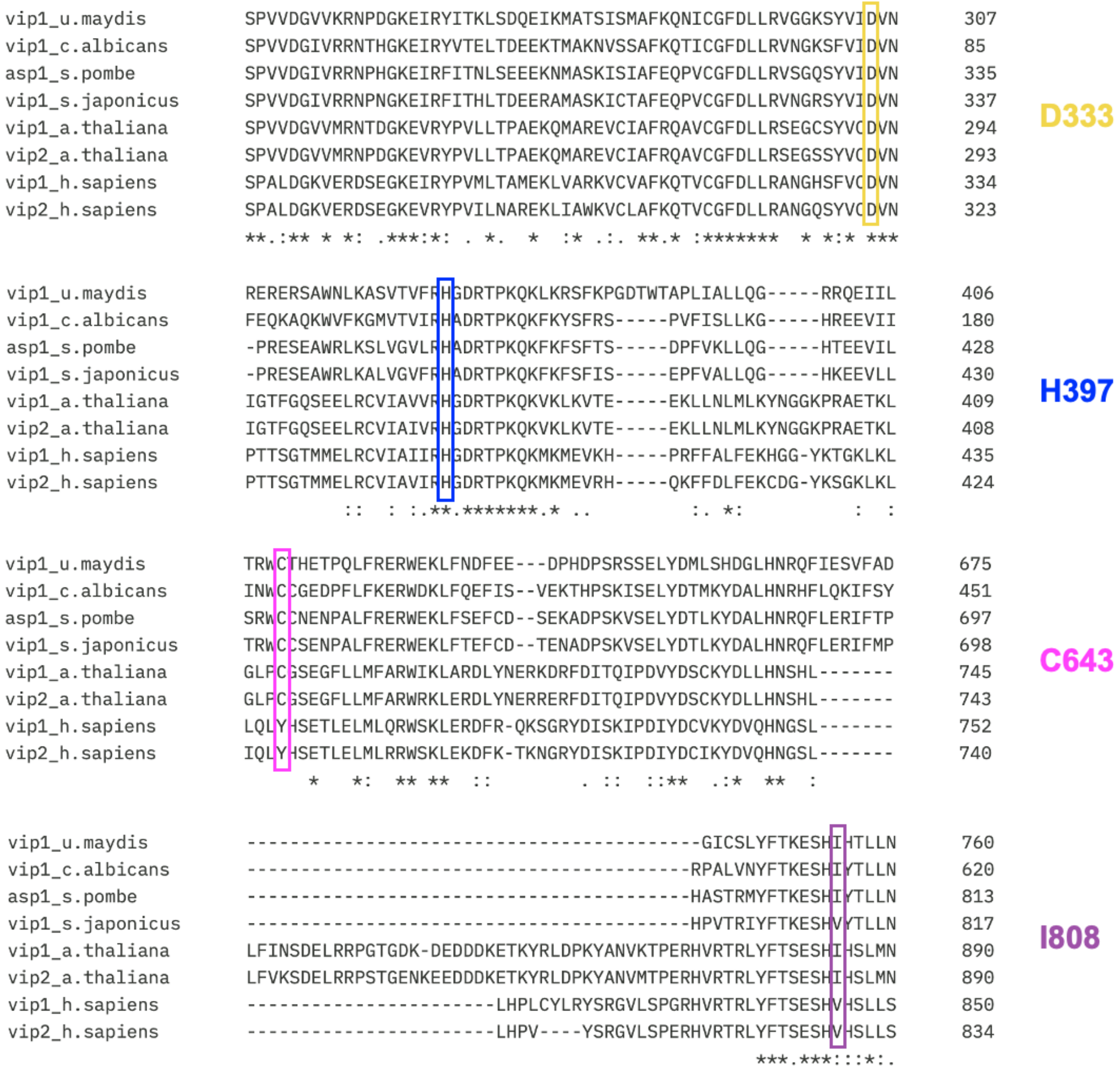
Table 4: Relevant amino acids of *S. pombe* Asp1 and their pLDDT score based on the AlphaFold2 structure prediction model.[130,131]

Amino acid	**pLDDT Score**
D333	95.70
H397	97.41
C607	89.61
C643	77.64
C644	77.00
C663	94.53
I808	96.64
C864	92.03
C868	80.03
C879	81.82

Initially C607 and C663 were believed to be located in the active site of the Asp1 pyrophosphatase domain and coordinate a [2Fe-2S] cluster together with other two cysteines (C864, C879).[3] However, the release of AlphaFold2 and accurate sequence-based structure prediction revealed that these two residues are located on the surface of the Asp1 pyrophosphatase domain, too far away to be able to contribute to the coordination of the same Fe-S cluster. Previous Asp1 studies further revealed that C663 is involved in cluster binding, and that the mutation of this residue impacts the metal content per monomer protein.[3] Thus, the question arose on how these four cysteines could be involved in the coordination of the same cluster or if there are more clusters bound to the Asp1 pyrophosphatase domain.

4.2 Sequence alignments of Asp1 orthologs

Previous studies on *S. pombe* Asp1 highlighted some conserved amino acids residues in both domains, that could play a pivotal role in the regulation of the function and the cofactor binding. Parts of a sequence alignment comprising Asp1 orthologs from different organism classes show that certain amino acids appear highly conserved among the species and the role of some has already been investigated and determined. Furthermore, the cysteine C643 is the only cysteine conserved among all species except human PPIP5K, in some cases this being accompanied by a second cysteine in immediate proximity (*Candida albicans, S. pombe and Schizosaccharomyces japonicus,* Figure 8). This cysteine pair becomes ever more interesting since cysteines and histidine residues which are involved in cluster coordination are usually conserved among isoforms and orthologs.



```
vip1_u.maydis        SPVVDGVVKRNPDGKEIRYITKLSDQEIKMATSISMAFKQNICGFDLLRVGGKSYVIDVN    307
vip1_c.albicans      SPVVDGIVRRNTHGKEIRYVTELTDEEKTMAKNVSSAFKQTICGFDLLRVNGKSFVIDVN    85
asp1_s.pombe         SPVVDGIVRRNPHGKEIRFITNLSEEEKNMASKISIAFEQPVCGFDLLRVSGQSYVIDVN    335
vip1_s.japonicus     SPVVDGIVRRNPNGKEIRFITHLTDEERAMASKICTAFEQPVCGFDLLRVNGRSYVIDVN    337
vip1_a.thaliana      SPVVDGVVMRNTDGKEVRYPVLLTPAEKQMAREVCIAFRQAVCGFDLLRSEGCSYVCDVN    294
vip2_a.thaliana      SPVVDGVVMRNPDGKEVRYPVLLTPAEKQMAREVCIAFRQAVCGFDLLRSEGSSYVCDVN    293
vip1_h.sapiens       SPALDGKVERDSEGKEIRYPVMLTAMEKLVARKVCVAFKQTVCGFDLLRANGHSFVCDVN    334
vip2_h.sapiens       SPALDGKVERDSEGKEVRYPVILNAREKLIAWKVCLAFKQTVCGFDLLRANGQSYVCDVN    323
                     **.:** * *: .***:*: . *.  *  :* .:. **.* :*******  * *:* ***

vip1_u.maydis        RERERSAWNLKASVTVFRHGDRTPKQKLKRSFKPGDTWTAPLIALLQG-----RRQEIIL    406
vip1_c.albicans      FEQKAQKWVFKGMVTVIRHADRTPKQKFKYSFRS-----PVFISLLKG-----HREEVII    180
asp1_s.pombe         -PRESEAWRLKSLVGVLRHADRTPKQKFKFSFTS-----DPFVKLLQG-----HTEEVIL    428
vip1_s.japonicus     -PRESEAWRLKALVGVFRHADRTPKQKFKFSFIS-----EPFVALLQG-----HKEEVLL    430
vip1_a.thaliana      IGTFGQSEELRCVIAVVRHGDRTPKQKVKLKVTE-----EKLLNLMLKYNGGKPRAETKL    409
vip2_a.thaliana      IGTFGQSEELRCVIAIVRHGDRTPKQKVKLKVTE-----EKLLNLMLKYNGGKPRAETKL    408
vip1_h.sapiens       PTTSGTMMELRCVIAIIRHGDRTPKQKMKMEVKH-----PRFFALFEKHGG-YKTGKLKL    435
vip2_h.sapiens       PTTSGTMMELRCVIAVIRHGDRTPKQKMKMEVRH-----QKFFDLFEKCDG-YKSGKLKL    424
                              ::  : :.**.*******.* ..         :. *:          :  :

vip1_u.maydis        TRWCTHETPQLFRERWEKLFNDFEE---DPHDPSRSSELYDMLSHDGLHNRQFIESVFAD    675
vip1_c.albicans      INWCCGEDPFLFKERWDKLFQEFIS--VEKTHPSKISELYDTMKYDALHNRHFLQKIFSY    451
asp1_s.pombe         SRWCCNENPALFRERWEKLFSEFCD--SEKADPSKVSELYDTLKYDALHNRQFLERIFTP    697
vip1_s.japonicus     TRWCCSENPALFRERWEKLFTEFCD--TENADPSKVSELYDTLKYDALHNRQFLERIFMP    698
vip1_a.thaliana      GLFCGSEGFLLMFARWIKLARDLYNERKDRFDITQIPDVYDSCKYDLLHNSHL-------    745
vip2_a.thaliana      GLFCGSEGFLLMFARWRKLERDLYNERRERFDITQIPDVYDSCKYDLLHNSHL-------    743
vip1_h.sapiens       LQLYHSETLELMLQRWSKLERDFR-QKSGRYDISKIPDIYDCVKYDVQHNGSL-------    752
vip2_h.sapiens       IQLYHSETLELMLRRWSKLEKDFK-TKNGRYDISKIPDIYDCIKYDVQHNGSL-------    740
                           *   *:  ** **  ::        . ::  ::**  .:*  **  :

vip1_u.maydis        ------------------------------------------GICSLYFTKESHIHTLLN    760
vip1_c.albicans      -----------------------------------------RPALVNYFTKESHIYTLLN    620
asp1_s.pombe         -----------------------------------------HASTRMYFTKESHIYTLLN    813
vip1_s.japonicus     -----------------------------------------HPVTRIYFTKESHVYTLLN    817
vip1_a.thaliana      LFINSDELRRPGTGDK-DEDDDKETKYRLDPKYANVKTPERHVRTRLYFTSESHIHSLMN    890
vip2_a.thaliana      LFVKSDELRRPSTGENKEEDDDKETKYRLDPKYANVMTPERHVRTRLYFTSESHIHSLMN    890
vip1_h.sapiens       -----------------------LHPLCYLRYSRGVLSPGRHVRTRLYFTSESHVHSLLS    850
vip2_h.sapiens       -----------------------LHPV----YSRGVLSPERHVRTRLYFTSESHVHSLLS    834
                                                                ***.***:::*:.
```

Figure 8: Sequence alignment of Asp1 orthologs under consideration of catalytic sites in the protein. The sequence alignment was conducted using the online tool Clustel Omega. Only relevant areas of the alignment are shown.[132]

The amino acid sequence of the Asp1 pyrophosphatase domain contains eleven cysteines, out of which only one is conserved. It seems almost contradictory that a biologically relevant cofactor would be coordinated by different cysteines and possibly in different areas of the protein, when from an evolutionary point of view these are regions that exhibit little change in order to preserve the function of the cofactor. However, this could also indicate the possibility that many of the present cysteine residues play a role in the maintenance of the domain structure, especially since substrates of the pyrophosphatase domain are large inositol molecules with single or double phosphate groups at each carbon atom. Asp1 is not the only ortholog of the PPIP5K/Vip1 family that exhibits changes in the key motif of the phosphatase-like domain. A similar deviation at position 808 from the classical histidine acid phosphatase signature motif has been observed for the C-terminal domain of *S. japonicus* Asp1 (Figure 8).[133] This opened a new perspective where the Asp1 pyrophosphatase domain of *S. pombe* Asp1 was exchanged for the one from the *S. japonicus* Asp1 ortholog. These two orthologs exhibit high sequence identity and share the same isoleucine 808 in the pyrophosphatase domain. The newly exchanged pyrophosphatase domain exhibit the same function in *S. pombe*, showing that the organism does not differentiate between its own domain or one from its orthologs as long as the key signature motifs remain unchanged.

When taking a look at other residues besides cysteines it can be observed that amino acids H397 and D333 are conserved among all investigated species. The aspartate at position 333 was reported to be a key residue in the maintenance of the kinase activity, as its loss-of-function mutation can lead to a 'kinase-dead' protein. Moreover, residues H397 and I808 have been shown to reduce the pyrophosphatase activity and thus lead to increased formation of 1-IP$_7$ and 1,5-IP$_8$ metabolites in vivo much over wildtype levels (*S. pombe, C. albicans*, Figure 8).[47] It appears that Vip1 from *S. japonicus* and PPIP5K1 and PPIP5K2 from *H. sapiens* exhibit one more peculiarity, as the overall highly relevant I808 residue is naturally substituted with a valine at this position. This feature indicates that the pyrophosphatase activity in these species might underlie a different regulation system as the one proposed for *S. pombe* Asp1 and orthologs presenting the same conserved residues (Figure 8).

4.3 Expression and purification of the Asp1$^{365-920}$ pyrophosphatase domain

New gene constructs of the Asp1 pyrophosphatase domain, based on a codon-optimized sequence, were designed to overcome the limitation imposed by the use of rare codons. Structural and spectroscopic studies require high amounts of pure and homogenous protein. Obtaining the protein required for these experiments has been a challenge in previous studies. To increase the protein yields the original *asp1*$^{+}$ gene sequence was designed in a codon-optimized version for heterologous gene overexpression in *E. coli* using synonymous codons. This step led to a major improvement of the expression protocol, since a notably higher protein yield (2x-3x more protein) was obtained from the same amount of cell pellet. The original *asp1*$^{+}$ gene sequence, consisting of *S. pombe* typical codons, complicated the heterologous expression since these codons were not the naturally occurring in *E. coli*. Codon frequency is different in every organism and rare codons act rate-limiting in the expression of heterologous proteins. Due to this major enhancement, this gene was used for the cloning of all following DNA constructs used in this study.

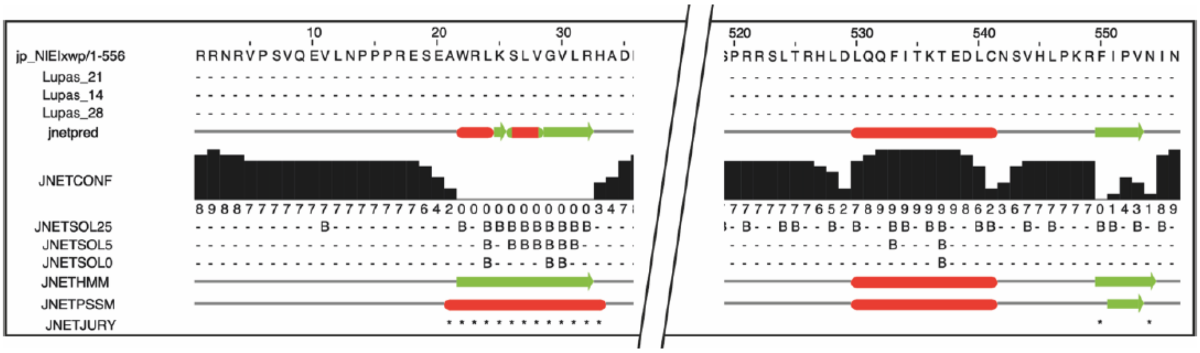


Figure 9: Secondary structure prediction of the Asp1 pyrophosphatase domain using JPred 4 Incorporating Jnet. Illustrated is the secondary structure prediction for the first and the last 35 amino acids of Asp1$^{365-920}$. α-Helices are depicted as red tubes whereas β-sheets are shown as green arrows for JNETPSSM, JNETHMM and jnetpred. JNETCONF decides upon the reliability of the prediction based on numbers ranging from 0 to 9, with large numbers indicating a high reliability.[134]

A shorter variant (Asp1$^{383-918}$) of wildtype Asp1$^{365-920}$ was designed without some potentially unstructured regions to facilitate crystallization. The preparation of screens for crystallization conditions has proved difficult in previous studies and no crystals resulted from those experiments. The optimization of DNA constructs for crystallization experiments was attempted using an online tool for secondary structure prediction called JPred 4 Incorporating Jnet. This was used to identify possible unstructured areas in the predicted secondary protein structure elements like loops or turns that might impede the crystallization of the Asp1 pyrophosphatase domain. According to the prediction model (Figure 9), the first 18 and last two amino acids of the pyrophosphatase domain are not structurally relevant. This discovery was further validated using AlphaFold2 in order to determine if deleting the 20 amino acids would be adequate or larger unstructured areas should be removed. The predicted AlphaFold2 model confirmed that the deletion of the 20 amino acids in question would suffice to create a pyrophosphatase construct that contains less unstructured regions. Thus, a shorter construct of the pyrophosphatase domain was designed showcasing only amino acids 383 to 918 so that the protein variant exhibits less flexibility and mobility in the terminal regions during crystallization.

The construct Asp1$^{383-918}$ does not exhibit increased stability through the truncation of some terminally unstructured regions based on the structure prediction. The shorter version of the Asp1 pyrophosphatase domain was designed only for screening for crystallization. The removal of some unstructured regions in the predicted structure model might promote the crystallization process and imply higher stability of the protein. Experiments showed that the Asp1$^{383-918}$ did not lead to the formation of any crystals as presumed. Despite thorough structure analysis and prediction using AlphaFold2 and JPred 4 Incorporating Jnet the designed construct did neither crystallize nor did the changes to the sequence provide any stability increase. In conclusion, no crystallization conditions could be found for the Asp1$^{383-918}$ pyrophosphatase domain and thus no protein crystals were obtained.

Asp1 was reported to exist as a dimer under physiological conditions, an aspect which is very interesting when considering the presence of an Fe-S cluster in the pyrophosphatase domain.[1] Such a cofactor can play a role in oligomerization events depending on its position in the protein structure. To investigate this, the Asp1$^{365-920}$ construct was cloned with a His- and a GST-tag. The GST-tag is known to dimerize easily due to formation of disulfide bonds between two GST molecules.[135,136] This resulted in obtaining two constructs, His-Asp1$^{365-920}$ and GST-Asp1$^{365-920}$.

Expression tests show that expression for 3 h results in the highest expression levels. Expression tests conducted with GST-Asp1$^{365-920}$, His-Asp1$^{365-920}$ and with truncated pyrophosphatase variants GST-Asp1$^{383-918}$ and His-Asp1$^{383-918}$ indicate that induction occurs successfully based on a more prominent band at approx. 90 kDa for GST-tagged variants and 65 kDa for His-tagged constructs on the SDS-PAGE in both types of induction control samples (Figure 10, Figure 11, Figure 12). However, the band from the samples harvested after 24 h is not as intense as the one after only 3 h, indicating that the intracellular levels of GST- and His-Asp1$^{365-920}$, as well as GST- and His-Asp1$^{383-918}$ decrease upon extensive expression time. A reason for this might consist in the fact that the culture is saturated after a long incubation time and *E. coli* cells enter the death phase. Also, oxygen limitation due to high culture density and saturation could play a role in the decrease of the overall expression yield. At the same time, it is possible that the nutrient resources in the culture medium are depleted after long incubation times. This leads to the cells dying and further exposing the expressed target protein to protease activity. Protein degradation in bacteria is further promoted by stress conditions, such as nutrient depletion. Summarizing, a short expression time leads to significantly higher yields in the cell pellet for the different Asp1$^{365-920}$ and Asp1$^{383-918}$ variants.

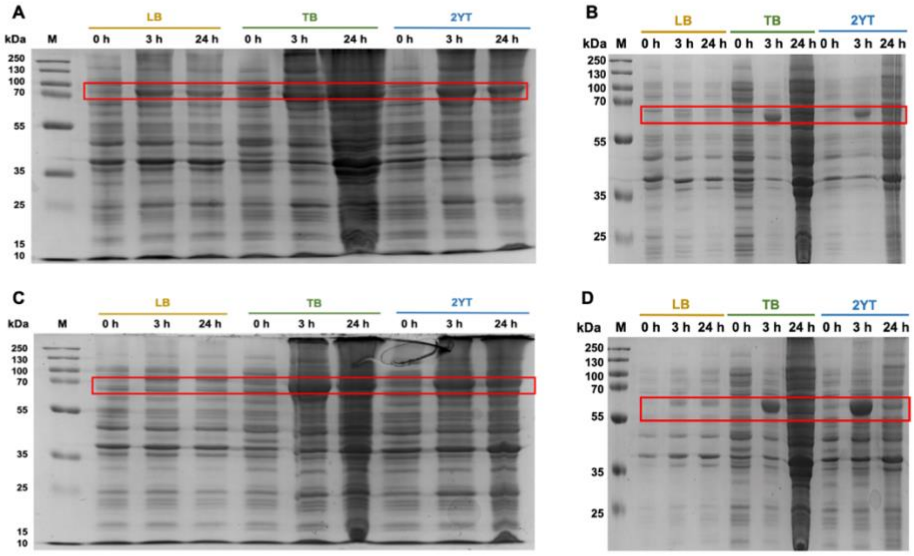


Figure 10: Expression tests in *E. coli* BL21(DE3) *ΔiscR* cells using different culture media. A) GST-Asp1$^{365\text{-}920}$, B) His-Asp1$^{365\text{-}920}$, C) GST-Asp1$^{383\text{-}918}$, D) His-Asp1$^{383\text{-}918}$.

A rich medium such as 2x YT is the optimal cultivation medium for the expression of Asp1 variants. The choice of cultivation medium has a great impact on the expression levels. Cultivation in LB medium does not appear adequate since the lowest protein levels were recorded in the expression controls from this batch (Figure 10, Figure 11, Figure 12). Expression in TB and 2YT media leads to bands stronger in intensity for all the different tested Asp1 variants, indicating a probably better cell growth and that a high protein amount could be harvested after 3 h of gene overexpression after cultivation in this medium. Thus, the expression of Asp1$^{365\text{-}920}$ and Asp1$^{383\text{-}918}$ variants is best conducted in a nutrient-rich medium like 2x YT.

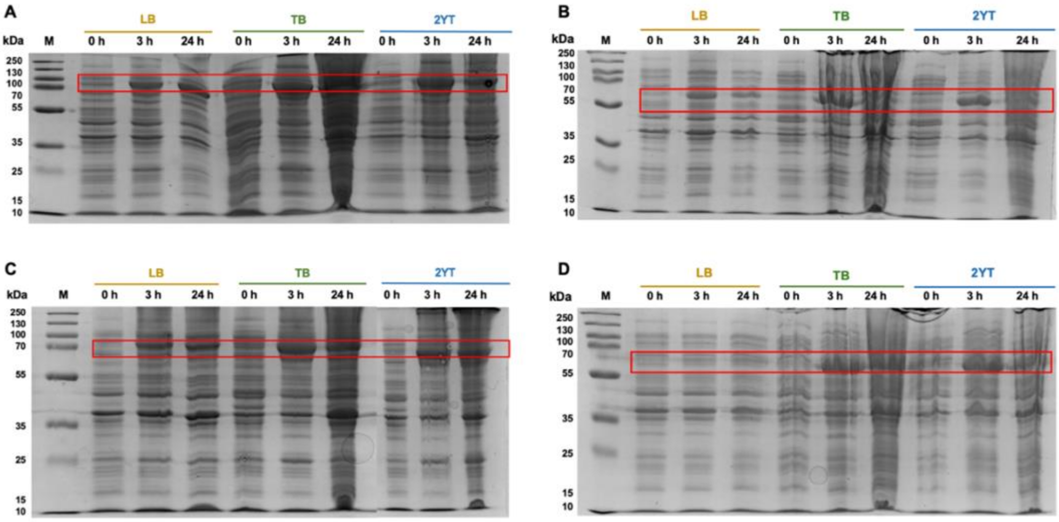


Figure 11: Expression tests in *E. coli* NiCo21 (DE3) cells using different culture media. A) GST-Asp1$^{365\text{-}920}$, B) His-Asp1$^{365\text{-}920}$, C) GST-Asp1$^{383\text{-}918}$, D): His-Asp1$^{383\text{-}918}$.

The same expression tests were conducted in two alternate bacterial cell strains to investigate if this has any influence of the expression of Asp1$^{365\text{-}920}$ variants. The expression results from the last set of expression tests using the classic *E. coli* BL21(DE3) strain do not differ significantly from tests where

E. coli NiCo21 (DE3) cells were used. The bacterial *E. coli* NiCo21 (DE3) strain was chosen for the expression tests since it is deficient for two proteases and can improve purity of target proteins isolated by IMAC (Figure 10, Figure 11, Figure 12). The protein expression strain NiCo21(DE3) has been engineered to minimize *E. coli* protein contamination of IMAC fractions: GlmS is mutated to eliminate binding to IMAC resins and three other proteins (SlyD, ArnA and Can) are tagged to enable rapid removal by chitin affinity chromatography.[137] Based on literature and previous work the expression in *E. coli* BL21(DE3) *ΔiscR* proved to be beneficial for enhancing maturation and synthesis of Fe-S proteins, since the deletion of the *isc* repressor (IscR) enhances transcription of several genes from the *isc* operon which are involved in cluster assembly.[138,139] Since there were no significant differences in the expression levels using three different bacterial strains, the *E. coli* BL21(DE3) *ΔiscR* strain was chosen for future large-scale expression all Asp1 variants. The expression in this strain appears to be optimal among the three on the knowledge from previous studies and little differences in expression levels among the conducted expression tests.

A short expression period of 3 h at 30 °C is optimal to produce 3-6 g cell pellet, from which a protein amount between 75 mg sand 360 mg can be isolated. The expression tests presented in this work were conducted at 30 °C after gene overexpression was induced with IPTG. Prolonged expression at 20 °C did not prove beneficial for Asp1 variants at all. This might happen since the *E. coli* metabolism is focused on maintaining a correct growth pattern and surviving under not optimal growing conditions. The level of expression at 20 °C was lower than at 30 °C and thus the expression was conducted only at this beneficial higher temperature.

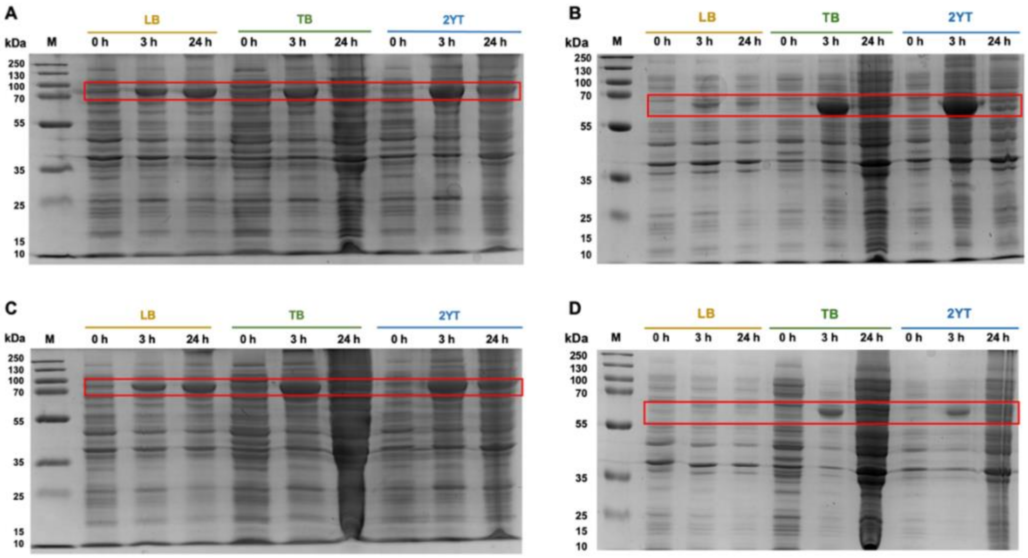


Figure 12: Expression tests in *E. coli* BL21(DE3) cells using different culture media. A) GST-Asp1$^{365-920}$, B) His-Asp1$^{365-920}$, C) GST-Asp1$^{383-918}$, D) His-Asp1$^{383-918}$.

GST- and His-tagged Asp1$^{365-920}$ and Asp1$^{383-918}$ variants cloned in *E. coli* BL21(DE3) *ΔiscR* can be isolated in the soluble fraction using the sodium phosphate buffer system supplemented with 10% glycerol. Protein purification attempts were conducted with glycerol as an additional buffer component to determine if it has any impact on the purification success. The experiments showed that the GST- and His-tagged Asp1$^{365-920}$ variants were expressed in a soluble fraction only after the addition of glycerol to all buffers. Glycerol is known to be a stabilizer due to its ability to accumulate around the hydration shell of proteins in solution.[140,141] Buffer supplementation with 10% glycerol made a tremendous difference in the solubility of all variants of the pyrophosphatase domain which was noticeable by the thick and intense band in the lysate fraction.

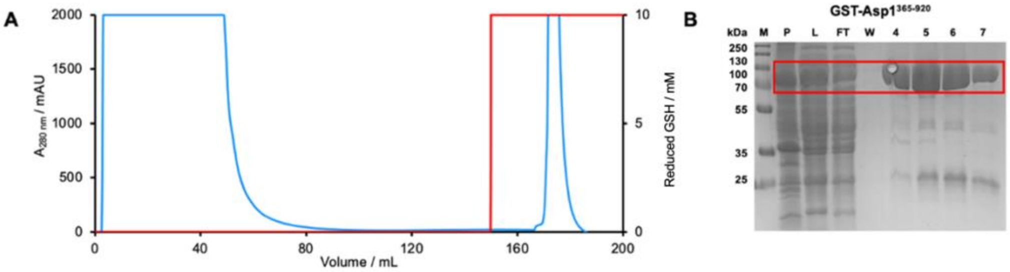


Figure 13: Purification of wildtype GST-Asp1$^{365-920}$ via GSTrap affinity chromatography. A) Chromatogram showing the absorbance at 280 nm in mAU (blue) and the concentration of reduced GSH in mM (red). B) SDS-PAGE showing the results of the anaerobic isolation of GST-Asp1$^{365-920}$. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, FT = flow-through, W = wash. The numbers indicate the different elution fractions.

The addition of glycerol consisted in the last optimization step for the purification protocol. The elution of the target protein during affinity chromatography (reduced GSH, Figure 13) led to a high peak, indicating a very high amount of protein. SDS-PAGE analysis showed the eluted protein to be reasonably pure and that there is still a high amount of remaining target protein in the flow-through (Figure 13.B). So, a second run was attempted using the flow-through fraction and yield similar to the one from the first run could be obtained. The chromatographic data show that the optimized purification protocol facilitates the extraction of large amounts of reasonably pure protein. Thus, it was possible to isolate and obtain soluble and pure fractions of GST- and His-tagged Asp1$^{365-920}$ and Asp1$^{383-918}$ variants. An untagged variant of Asp1$^{365-920}$ and Asp1$^{383-918}$ can be generated using the PreScission protease to cleave the GST-tag. Data of the untagged Asp1$^{365-920}$ domain was necessary to determine if the affinity tags have any influence on the characteristics of the domain. To obtain such protein variant the HRV-3C-Protease (PreScission) protease was used to remove the GST-tag at the corresponding cleavage site (LEVLFQGP) encoded on the vector.[142,143] The digest of GST-Asp1 was conducted using cleavage buffer (section 3.7) to maximize the activity of the protease. The separation of the protein from the GST-tag was conducted using a chromatographic run which resulted in two baseline separated peaks that contain the untagged pyrophosphatase domain and the GST-tag (Figure 14). This experiment enabled the generation of an untagged pyrophosphatase domain to be used for different spectroscopic and chromatographic studies presented in this work.

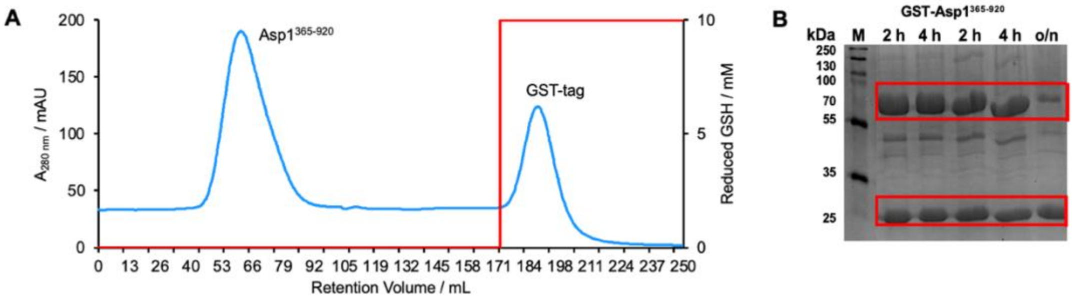


Figure 14: Separation of protein and affinity tag (A) after proteolytic digest of wildtype GST-Asp1$^{365\text{-}920}$ using PreScission protease (B). A) Chromatogram showing the absorbance at 280 nm in mAU (blue) and the concentration of red. GSH in mM (red). B) SDS-PAGE showing the results at different points in time of the prior proteolytic digest of GST-Asp1$^{365\text{-}920}$. The proteolytic digest was conducted as mentioned in section 3.7.

The isolation and purification of His-tagged protein variants of Asp1$^{365\text{-}920}$ was conducted similarly the only difference being the affinity chromatography type. The fractions collected during Ni-NTA affinity chromatography and loaded on the SDS-PAGE indicate that the His-tagged construct can be isolated pure and in high amounts (Figure 15). As in the case of the GST-tagged variant, the high amount of protein remaining in the flow-through fraction enable a second chromatography run. The addition of the glycerol in the lysis and running buffers increases the stability of His-Asp1$^{365\text{-}920}$ and indirectly the solubility so that the corresponding band in the pellet and the lysate fraction are almost identical. The established purification protocol results in high yields of pure and soluble His-Asp1$^{365\text{-}920}$.

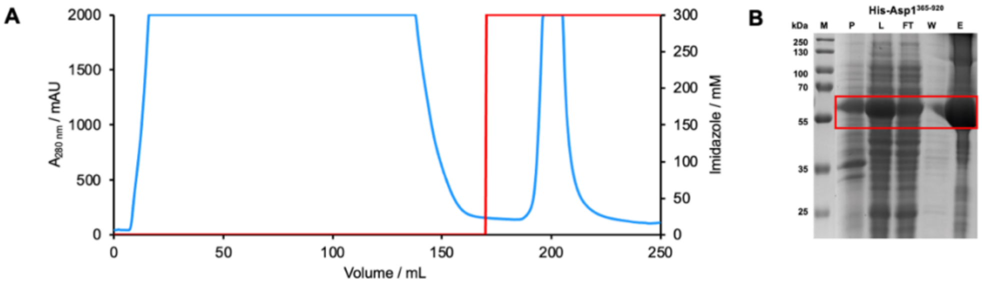


Figure 15: Purification results of wildtype His-Asp1$^{365\text{-}920}$ via IMAC. A) Chromatogram showing the absorbance at 280 nm in mAU (blue) and the imidazole concentration in mM (red). B) SDS-PAGE showing the results of the anaerobic isolation of His-Asp1$^{365\text{-}920}$. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, FT = flow-through, W = wash. The numbers indicate the different elution fractions.

An untagged variant of Asp1$^{365\text{-}920}$ and Asp1$^{383\text{-}918}$ can be generated using the TEV protease (HHHHHHHHHH) to remove the His-tag. The digest of His-tagged protein variants was conducted to establish an alternate way of obtaining untagged protein. The His-tag was removed using the corresponding TEV protease and under suitable experimental conditions to maximize protease activity. This kind of proteolytic digest appears to be as successful as the one where the PreScission protease was used. According to the chromatogram (Figure 16.A) there is no residual undigested protein. However, the results of the proteolytic digest cannot be properly monitored via SDS-PAGE as the absence of the His-tagged consists in a very small size difference (3 kDa) between the tagged and the digested versions of the Asp1 pyrophosphatase domain (Figure 16.B). This established protocol allows an alternative way to generate of sufficient amounts of untagged protein for the subsequent spectroscopic and chromatographic studies.

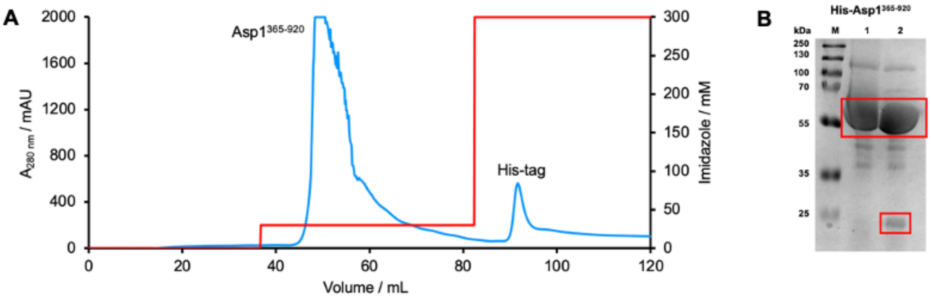


Figure 16: Separation of protein and affinity tag (A) after proteolytic digest of wildtype His-Asp1$^{365\text{-}920}$ using TEV protease (B). A) Chromatogram showing the absorbance at 280 nm in mAU (blue) and the imidazole concentration in mM (red). B) SDS-PAGE showing the results of the priorly-conducted proteolytic digest of His-Asp1$^{365\text{-}920}$ and subsequent separation of tag and protein. The proteolytic digest was conducted as mentioned in section 3.7.

4.4 Characterization of the Asp1$^{365\text{-}920}$ Fe-S cluster

EAS and ICP-MS were used to investigate if the content of the [2Fe-2S] cluster can be elevated in the sample in vitro. For this experiment, two samples of as-isolated GST-Asp1$^{365\text{-}920}$ from the same purification batch were chemically reconstituted using different amounts of iron and sulfur. Chemical reconstitution of [2Fe-2S] proteins is routinely conducted using a 5-fold excess of iron and sulfur sources, whereas for the reconstitution of higher nuclear cluster, such as a [4Fe-4S] cluster, 10 equivalents of the chemicals are needed. Besides the increased concentration of the two iron and sulfur compounds, the parameters of the reconstitution protocol remained unchanged. Size exclusion chromatography was performed to remove Fe-S aggregates. The electronic absorption spectrum of the GST-Asp1$^{365\text{-}920}$ apo variant exhibits none of the previously described spectroscopic features. The protein sample was prepared using reductive chelation with 10-fold excess of bipyridyl and were desalted via SEC. In addition, a negative control was prepared from the same purification batch by chelating as-isolated GST-Asp1$^{365\text{-}920}$ in order to have a sample without cluster for the comparison.

The GST-tag was chosen as affinity tag for different Asp1$^{365\text{-}920}$ constructs since the protease needed to cleave off the GST-tag does not require any EDTA in the buffer. The His-tag was not the best option for an affinity tag since the TEV-protease, which is needed to cleave off the His-tag, is inhibited by metal ions. The inhibition of the TEV-protease can only be eluded if EDTA is present in the cleavage buffer. The presence of EDTA could however impact [2Fe-2S] cluster coordinated by Asp1$^{365\text{-}920}$, as demonstrated in the following experiments.

The Asp1 pyrophosphatase domain (Asp1$^{365\text{-}920}$) appears to bind a [4Fe-4S] cluster which does not coincide with the native form. The negative control is depicted in grey and indicates the electronic spectrum of GST-Asp1$^{365\text{-}920}$ in the absence of the cofactor. The electronic spectrum of the as-isolated Asp1 pyrophosphatase domain is shown in yellow and exhibits the prominent broad peak at 410 nm and the shoulder at 320 nm with molar extinction coefficient values of 10 and 7 $mM^{-1}*cm^{-1}$ (Figure 17.A). The spectrum of the wildtype pyrophosphatase domain reconstituted with 5-fold excess of iron and sulfur sources (orange curve) shows similar bands but a higher overall intensity as the spectrum of the previously mentioned state with overall 1 Fe/monomer occupancy. The red curve corresponding to as-isolated GST-Asp1$^{365\text{-}920}$ chemically reconstituted with 10-fold excess of iron and sulfur shows the same two well-defined absorption bands with molar extinction coefficient values of 16 (410 nm) and 22 (320 nm) $mM^{-1}*cm^{-1}$ which are in good agreement with literature on [4Fe-4S] proteins (Figure 17.A).[121,144] The electronic absorption spectra show that the three samples of GST-Asp1$^{365\text{-}920}$ (yellow, orange and red curves) bind a [4Fe-4S] cluster based on the characteristic absorption band at 410 nm. The different intensity indicates that for each sample a different amount of molecules bind the [4Fe-4S] cluster. Also, the two samples chemically reconstituted with different amounts of iron and sulfur sources both coordinate a [4Fe-4S] cluster based on the course of the electronic spectrum (Figure 17). Overall, the ICP-MS data shows that the number of cluster-binding molecules gradually increases upon chemical reconstitution with 5- and 10-fold excess of iron and sulfur sources and achieves 2 Fe/monomer (orange curve) and 3.6 Fe/monomer (red curve), respectively (Figure 17). The only difference among the three cluster-containing samples (yellow, orange and red curve) is the number of molecules that bind the [4Fe-4S] cluster. Based on the ICP-MS data, the chemically reconstituted GST-Asp1$^{365\text{-}920}$ sample corresponding to the orange curve binds 2 Fe/monomer. This data paired with the observations from the electronic absorption spectrum indicate that GST-Asp1$^{365\text{-}920}$ reconstituted with 5-fold excess of iron and sulfur (orange curve) binds a [4Fe-4S] cluster and only half of the molecules in the sample bind the cluster altogether, hence leading to an overall value of 2 Fe/monomer. Similarly, it can be observed that the as-isolated GST-Asp1$^{365\text{-}920}$ sample exhibits a [4Fe-4S] cluster as well, but less than half of the molecules bind the cofactor when compared to the sample reconstituted with 10-fold excess of iron and sulfur. Zinc ions can bind free binding sites inside an Fe-S cluster, and these are mostly present in the as-isolated GST-Asp1$^{365\text{-}920}$ sample.[145–147] The zinc levels in the four samples were analyzed additionally to the iron levels to obtain a better understanding of the cluster occupancy in the case that not enough iron is present in the sample or expression medium. In Figure 17.C it can be observed that the apo and reconstituted wildtype protein achieve residual zinc levels of only 0.3 Zn/monomer. The zinc content of the as-isolated sample amounts to approx. 0.6 Zn/monomer, which shows that zinc ions that are usually present in expression media can bind to unoccupied binding sites inside the Asp1 Fe-S cluster. This data shows that possibly not enough iron is present during expression to guarantee full occupancy of the Asp1 Fe-S cluster after aerobic expression and anaerobic isolation. The EAS paired with the ICP-MS data indicate that chemical reconstitution leads to the formation of a [4Fe-4S] cluster that is biologically not relevant and does not coincide with the reported data on the as-isolated Asp1 pyrophosphatase domain.

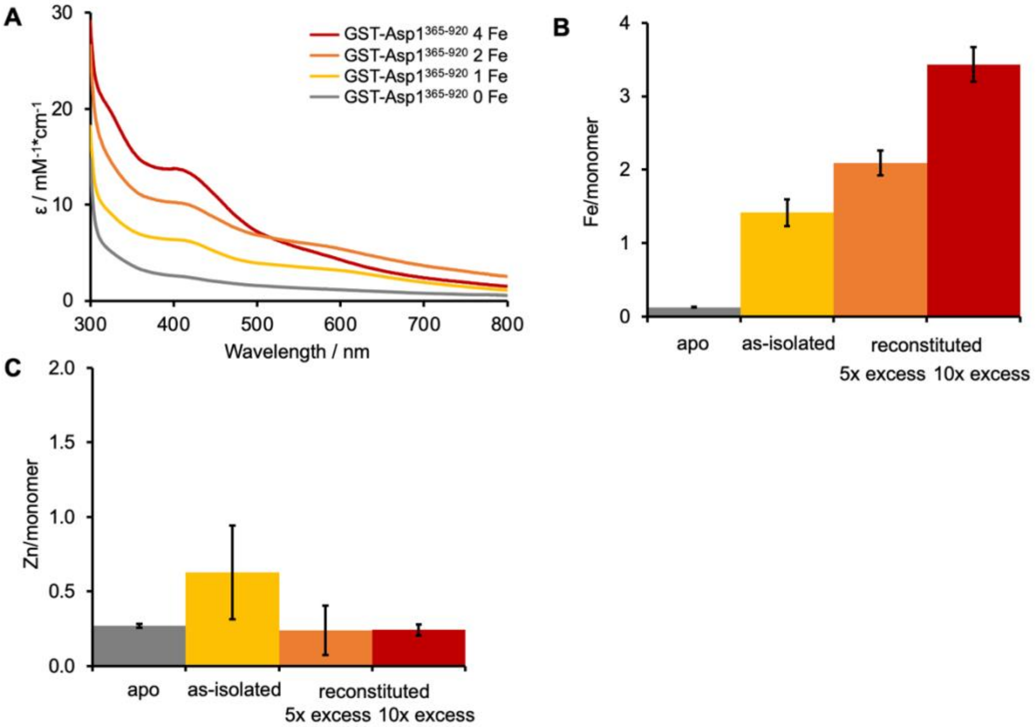


Figure 17: Overview of EA spectra and iron content of wildtype GST-Asp1$^{365\text{-}920}$ in different cofactor-bound states. A) The molar extinction coefficient ε in $mM^{-1}*cm^{-1}$ (y-axis) was plotted as a

function of the wavelength in nm (x-axis). B) The measured iron levels of the different samples are shown in corresponding colours. C) The measured zinc levels of the different samples are shown in corresponding colours. The measurements consisting of biological triplicates have additional error bars depicting the standard deviation.

Relevant studies regarding Asp1 indicate ferric chloride as the routinely-used iron source during reconstitution experiments. Some questions that became apparent during this study were if the chemical reconstitution can be conducted with the same success using other reductants and iron sources and if these changes affect in any way the spectral features of the Asp1 pyrophosphatase domain. With the goal of answering those questions, chemical reconstitution experiments with DTT and GSH as reductants and ferric chloride and ferric citrate as iron sources were performed according to section 3.9. It was chosen to reconstitute with 10-fold excess, to investigate if the same [4Fe-4S] cluster mentioned in the experiment above could be reconstituted under different conditions. The samples were desalted using SEC on a 16/600 HiLoad 200 pg column before the analysis via EAS and ICP-MS commenced.

The use of DTT and GSH as reductants during chemical reconstitution experiments does not lead to any significant differences in the EAS and ICP-MS data for GST-Asp1$^{365-920}$. In Figure 18.A the resulting electronic absorption spectra of the 4 combinations of reagents can be observed. Overall, all spectra indicate the broad absorption band at 420 nm which is characteristic for a [4Fe-4S] cluster. The spectra of the two samples where DTT was used as reductant show the same bands, with a slight difference in the values of the molar extinction coefficient (16 and 20 $mM^{-1}*cm^{-1}$at 420 nm, Figure 18.A). The same can be said overall about the other two samples were GSH was used as reductant. Interestingly, in both latter samples the shoulder signal at 320 nm appears more prominent than the first-mentioned two samples. GSH is not only the natural binding partner to the GST-tag, but literature shows that this small molecule consisting of three amino acids can occupy coordination sites of Fe-S clusters where organic thiol groups from cysteine residues are usually involved.[148–150] This could occur during the reconstitution experiments, leading to a different electronic environment around the *S. pombe* Asp1 pyrophosphatase domain that ultimately results in more distinct spectral properties.

An iron content of 4 Fe/monomer is achieved for GST-Asp1$^{365-920}$ in all samples despite using DTT and GSH as reductants and two different iron sources during chemical reconstitution. As in the previous experiment, ICP-MS data was collected to obtain a full overview over the iron contents resulting from the four different experimental set-ups. Figure 18.B shows the ICP-MS data and that three out of four samples bind almost exactly 4 Fe/monomer protein, as expected when using the 10-fold reconstitution protocol. In case of the sample treated with GSH and ferric citrate 3.5 Fe/monomer could be measured however the results are in good agreement with the known standard deviation for this type of measurement. The standard deviation in measured ICP-MS data can occur due to the Bradford determination of protein concentration that is part of the sample preparation for ICP-MS. In addition, the low values below 0.3 Zn/monomer protein verify the fact that iron is the only metal occupying the binding site of the [4Fe-4S] cluster in the 10-fold reconstituted samples of GST-Asp1$^{365-920}$ (Figure 18.B). Due to the high excess of iron in the sample, there are no free binding sites that could be occupied by zinc ions. This data indicates overall that the use of different reductants and iron sources does not influence the reconstitution process or the quality significantly and that the [4Fe-4S] cluster that results from the chemical reconstitution does not coincide with the native cluster of the Asp1 pyrophosphatase domain.

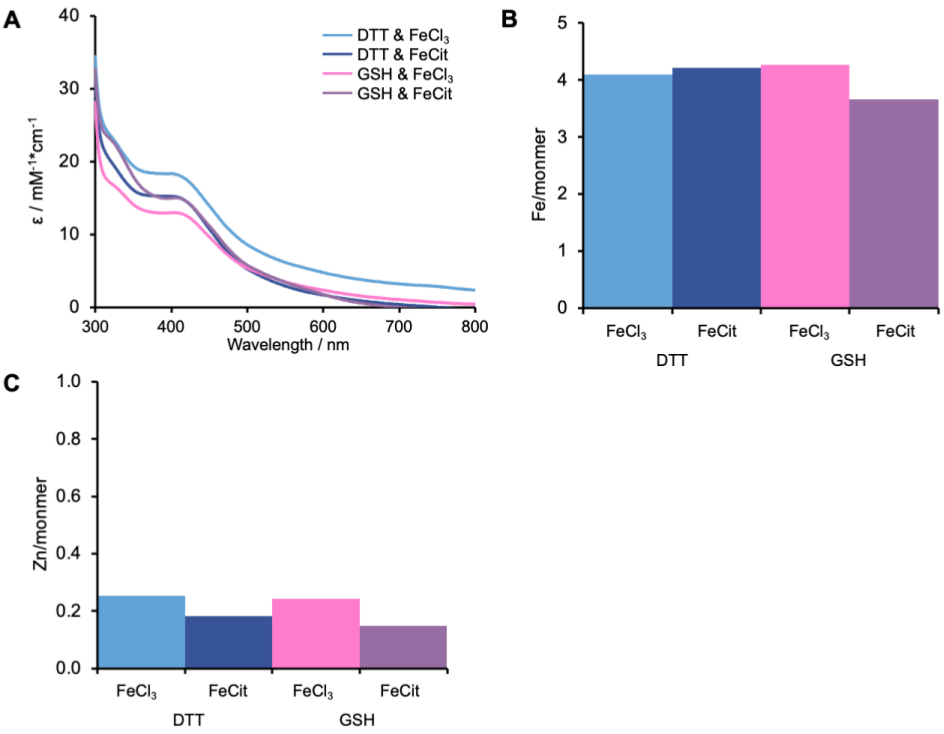


Figure 18: Influence of reductive agent and iron source on the EA spectra and iron contents in GST-Asp1$^{365-920}$ after chemical reconstitution with 10-fold excess of iron and sulfur sources. A) The molar extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). B) The measured iron levels of the different samples are shown in corresponding colours. C) The measured zinc levels of the different samples are shown in corresponding colours.

4.5 Characterization of the oligomeric state of Asp1$^{365\text{-}920}$

The oligomeric state of Asp1 variants was characterized via SEC due to the reveal of cysteines C607 and C663 by AlphaFold2 to be located at the surface of the protein. During the analysis of the model generated by AlphaFold2 the question arose if the cluster binding has any influence of the oligomerization state of the protein and a second Asp1 protein could contribute to the formation of an intermolecular [2Fe-2S] cluster. The following experiments serve to obtain a better understanding of these events and determine if the Asp1 Fe-S cluster has any influence on the oligomerization state or if certain oligomerization states lead to changes in the Fe-S cluster architecture.

The Asp1 pyrophosphatase domain is oligomeric dispersed with values above 600 kDa in the absence of the Fe-S cluster (apo form). In this experiment two samples of the apo pyrophosphatase domain were compared with regard to the oligomeric state. The apo state was achieved with two different approaches. The first sample analyzed in this experiment is the GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$, which was characterized in previous studies and deemed to not contain the cofactor.[3] The second sample consists of the wildtype sample of the shorter Asp1 pyrophosphatase domain that underwent reductive chelation to remove the cofactor (Figure 19). Both SEC chromatograms show only one well-resolved peak at 45.0 mL for GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$ and 46.7 mL for GST-Asp1$^{383\text{-}918}$. The experimentally determined hydrodynamic radius could be calculated using a calibration equation and identified the two peaks as high-order oligomers with molecular weights between 800-600 kDa. These values correspond to approx. 8 and 7 molecules of GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$ and GST-Asp1$^{383\text{-}918}$, respectively. These results show that absence of 4 Fe and 4 S per protein monomer leads to an oligomerization, where only one size of a large oligomer is present in the sample.

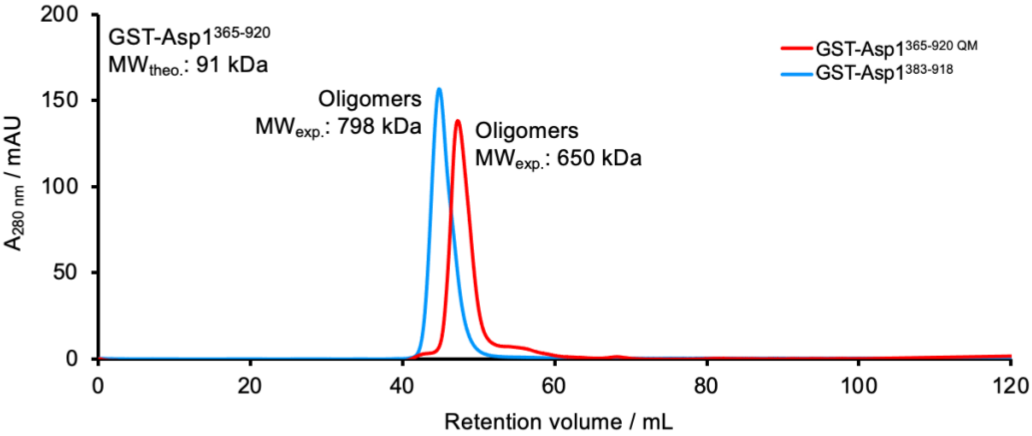


Figure 19: Overlayed SEC chromatograms showing the oligomeric state of apo GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$ and GST-Asp1$^{383\text{-}918}$. The SEC was conducted using a HiLoad 16/600 Superdex 200 pg column. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the retention volume on the primary y-axis. The sample concentration amounted to 200 µM each. The running buffer consisted of 100 mM sodium phosphate, 10% glycerol, pH 7.4. The theoretical molecular weight is abbreviated as $\mathrm{MW_{theo.}}$ and experimentally determined molecular weight as $\mathrm{MW_{exp.}}$.

The effect of glycerol on the oligomeric state of wildtype Asp1 (GST-Asp1$^{365\text{-}920}$) was investigated using SEC. Two chromatography runs were conducted under the same conditions with the only difference being the presence of glycerol in the running buffer for one sample. Figure 20 shows a comparison of the two SEC chromatograms, in which four peaks are visible. The same oligomeric states are noticeable under both conditions, however their ratios to each other are not identical. It can be observed that the amounts of monomer and dimer are higher in the sample treated without glycerol. Moreover, the trimeric oligomeric state is the most prominent one among all with a retention volume of 58.5 mL. Summarizing glycerol plays a role in the overall solubility and stability of the protein but does not significantly influence a certain low-order oligomeric state.

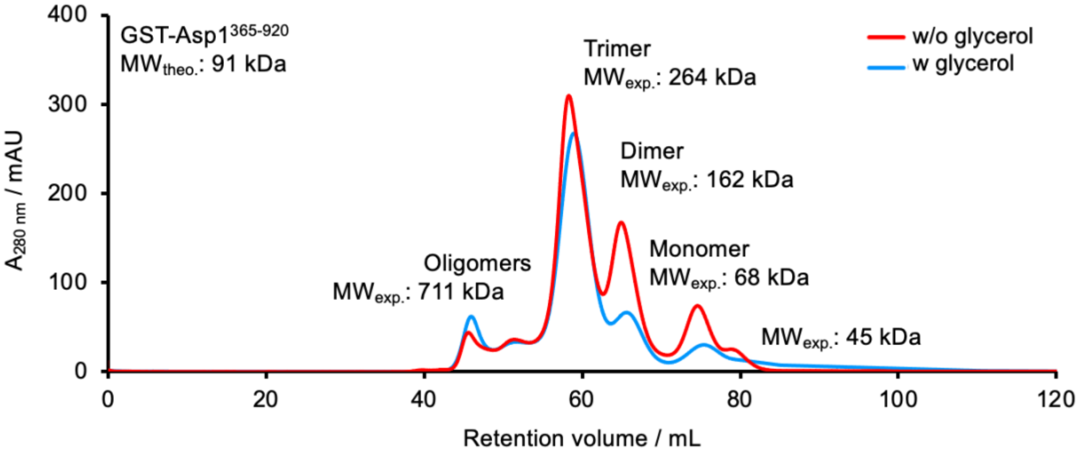


Figure 20: Overlayed SEC chromatograms showing the influence of glycerol on the oligomerization state of as-isolated GST-Asp1$^{365-920}$. The SEC was conducted using a HiLoad 16/600 Superdex 200 pg column. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the retention volume on the primary y-axis. The sample for the SEC run with glycerol was prepared with a concentration of 15.5 mg/mL and the one for the run without glycerol had a concentration of 18 mg/mL. The running buffer consisted of 100 mM sodium phosphate, pH 7.4. The theoretical molecular weight is abbreviated as MW$_{theo.}$ and experimentally determined molecular weight as MW$_{exp.}$.

The sample investigated in this experiment was chemically reconstituted with 5-fold molar excess of iron and sulfur. Figure 21 shows the SEC chromatogram of GST-Asp1$^{365-920}$ reconstituted with 5-fold excess of iron and sulfur, thus coordinating 2 Fe/monomer. The metal content of this sampled was confirmed by ICP-MS and this sample is one of the measured biological triplicates presented in the ICP-MS data above (figure 21). Only two peaks are visible, one attributed to high-grade oligomers at 48 mL and the second at 63 mL representing the dimer state. Oligomers are often found in the resulting chromatograms despite performing a centrifugation step prior to injecting the sample on the column. This further indicates that oligomerization occurs in solution and the peak at 48 mL is not due to precipitated protein in the sample (Figure 21). Overall, it can be observed that GST-Asp1$^{365-920}$ exists as dimer after reconstitution with 5-fold iron and sulfur sources and that a considerable part of the protein is in an oligomeric state.

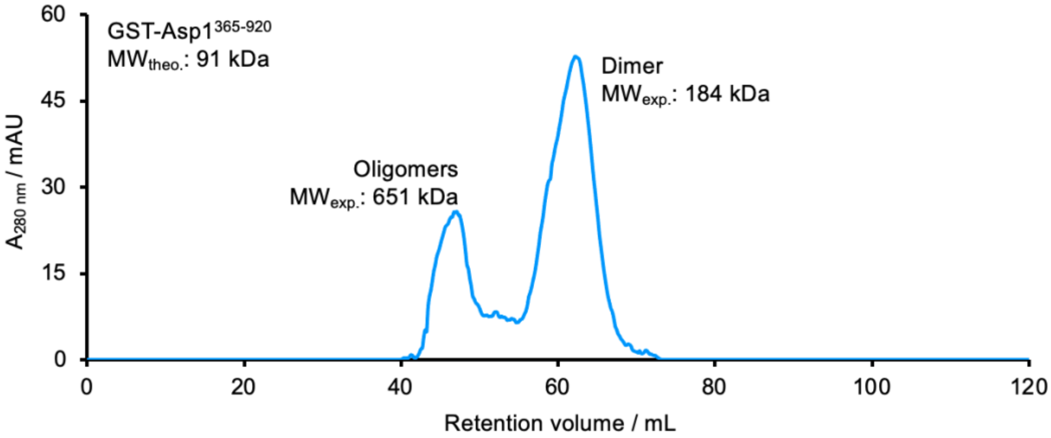


Figure 21: Chromatogram of GST-Asp1$^{365\text{-}920}$ after chemical reconstitution with 5-fold excess of iron and sulfur. The SEC was conducted using a HiLoad 16/600 Superdex 200 pg column. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the retention volume. The running buffer consisted of 100 mM sodium phosphate, 10% glycerol, pH 7.4. The theoretical molecular weight is abbreviated as MW$_{\text{theo.}}$ and experimentally determined molecular weight as MW$_{\text{exp.}}$.

The use of different reductants during chemical reconstitution does not influence the oligomeric state of GST-Asp1$^{365\text{-}920}$ significantly. Since reduction agents are commonly used for reconstituting the Fe-S clusters, it was of interest to analyze if the used reagents lead to any changes in the distribution of the oligomeric species of the Asp1 pyrophosphatase domain. This was investigated by conducting chemical reconstitution of GST-Asp1$^{365\text{-}920}$ with 5-fold excess of iron and sulfur sources and using DTT, GSH and TCEP as reducing agents. Based on the chromatograms showed in Figure 22, no significant difference in the distribution of the oligomeric state of the Asp1 pyrophosphatase domain can be observed. The samples are prominently in a dimeric state and less sample in a monomeric state. Some high-order oligomers are present as well in the volume area of 40-58 mL but it the used reducing agent does not make a difference in the oligomerization and that none of them appear able to influence the dispersion in any negative or positive way (Figure 22). Summarizing, the three tested reductants do not influence the oligomeric state in any way and Asp1 appears to occur mostly as dimer.

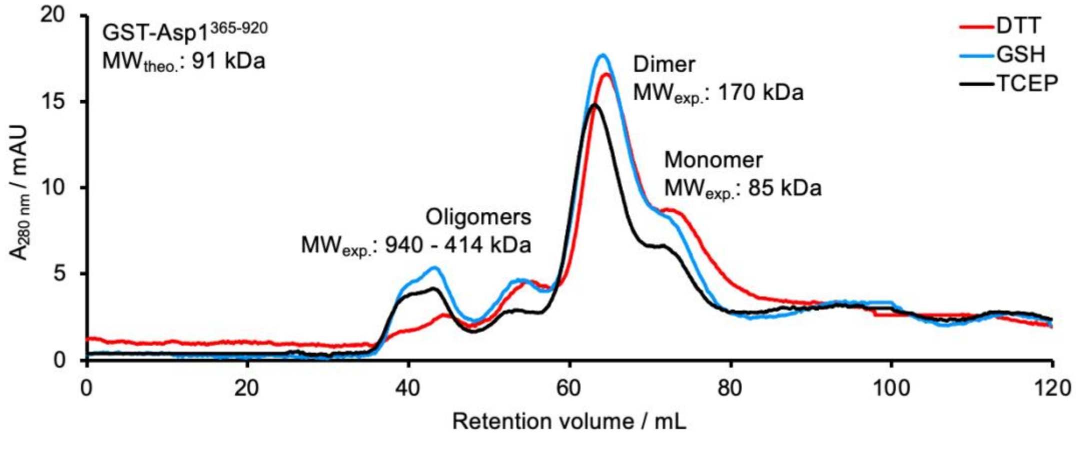


Figure 22: SEC chromatograms showing the influence of the reducing agent on the oligomerization state of GST-Asp1$^{365\text{-}920}$ during chemical reconstitution with 5-fold molar excess of iron and sulfur. The SEC was conducted using a HiLoad 16/600 Superdex 200 pg column. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the retention volume on the

primary y-axis. The running buffer consisted of 100 mM sodium phosphate, 10% glycerol, pH 7.4. The theoretical molecular weight is abbreviated as $MW_{theo.}$ and experimentally determined molecular weight as $MW_{exp.}$.

GST-Asp1$^{365-920}$ is monomeric dispersed after reconstitution with 10-fold iron and sulfur sources and implied binding of 4 Fe/monomer by the pyrophosphatase domain. In the previous experiments it was shown that as-isolated GST-Asp1$^{365-920}$ exists mostly trimeric state. So far, all of four observed metal contents for the Asp1 pyrophosphatase domain have been characterized by SEC beside the one where four iron ions are bound in one or two clusters per monomer protein. Therefore GST-Asp1$^{365-920}$ samples were prepared by chemical reconstitution with 10-fold molar excess of iron and sulfur sources. Since the expected outcome is the ligation of four iron ions by one monomer but the exact mechanism behind it is still unknown GSH was used as an alternative reductant to DTT, as it might contribute to metal coordination. At the same time iron (Fe^{3+}) citrate was used as a second source of iron alongside iron (Fe^{3+}) chloride in this reconstitution experiment since bioavailability among these chemicals can differ. All four reconstitution experiments were conducted under the same conditions with incubation for cluster assembly overnight at 4 °C and SEC runs in 100 mM sodium phosphate, 10% glycerol, pH 7.4. Remarkably, the coordination of four iron ions per monomer confirmed via ICP-MS (Figure 18) significantly shifts the oligomerization equilibrium on the side of the monomer (Figure 23). GST-Asp1$^{365-920}$ appears predominantly monomeric dispersed with a retention volume of 69.2 mL with only a small fraction of the sample assuming a high-order oligomerization state. The variation of reductant and iron source does not seem to affect the oligomerization state in any way and all conditions lead to a fully occupied [4Fe-4S] cluster per monomer GST-Asp1$^{365-920}$.

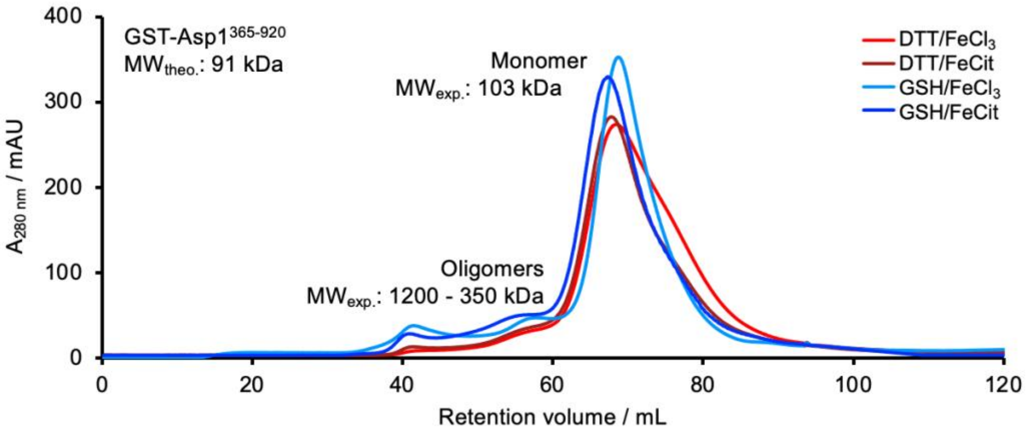


Figure 23: SEC chromatograms showing the differences in the oligomerization state of GST-Asp1$^{365-920}$ depending on the used reducing agent or iron source during chemical reconstitution (10-fold excess). The SEC was conducted using a HiLoad 16/600 Superdex 200 pg column. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the retention volume on the primary y-axis. The running buffer consisted of 100 mM sodium phosphate, 10% glycerol, pH 7.4. The theoretical molecular weight is abbreviated as $MW_{theo.}$ and experimentally determined molecular weight as $MW_{exp.}$.

The monomeric dispersed state of the [4Fe-4S]-binding GST-Asp1$^{365-920}$ was further confirmed via analytical ultracentrifugation, a technique that can be used to determine the oligomeric state in solution. The sedimentation profile shown in Figure 24.A indicates the relative fast sedimentation of the oligomers present in the sample with a complete measurement time of approx. 6 h. The concentration decreases over time as more particles sediment to the bottom of the measuring cell. At the same time the sedimentation front is shifted ever more towards the bottom. The *s*-value refers to the sedimentation coefficient, a measure of how quickly a molecule or particle settles or sediments in a centrifugal field and the *c*(*s*) distribution represents the distribution of sedimentation coefficients (*s*) within a sample, essentially showing the relative abundance of different sized particles based on their sedimentation behavior. The plateau was reached at approx. 0.2 AU. Looking at the *c*(*s*) distributions in Figure 24.B to Figure 24.D it is evident that there are two major classes of oligomeric states, one being the monomer with an *s*-value of 6.1 and molecular weight of 100 kDa and the second one being an initially unidentifiable species of approx. 51 kDa with a *s*-value of 3.9. The high-order oligomers detectable from Figure 23 are visible in Figure 24.D where many peaks attributed to different large oligomeric states are found over a broad region spanning 40 S. The higher the s-value the faster is the sedimentation of the observed molecule and thus the larger its molecular weight. This approach enabled the confirmation of the monomeric dispersed state of GST-Asp1$^{365\text{-}920}$ that was achieved after chemical reconstitution.

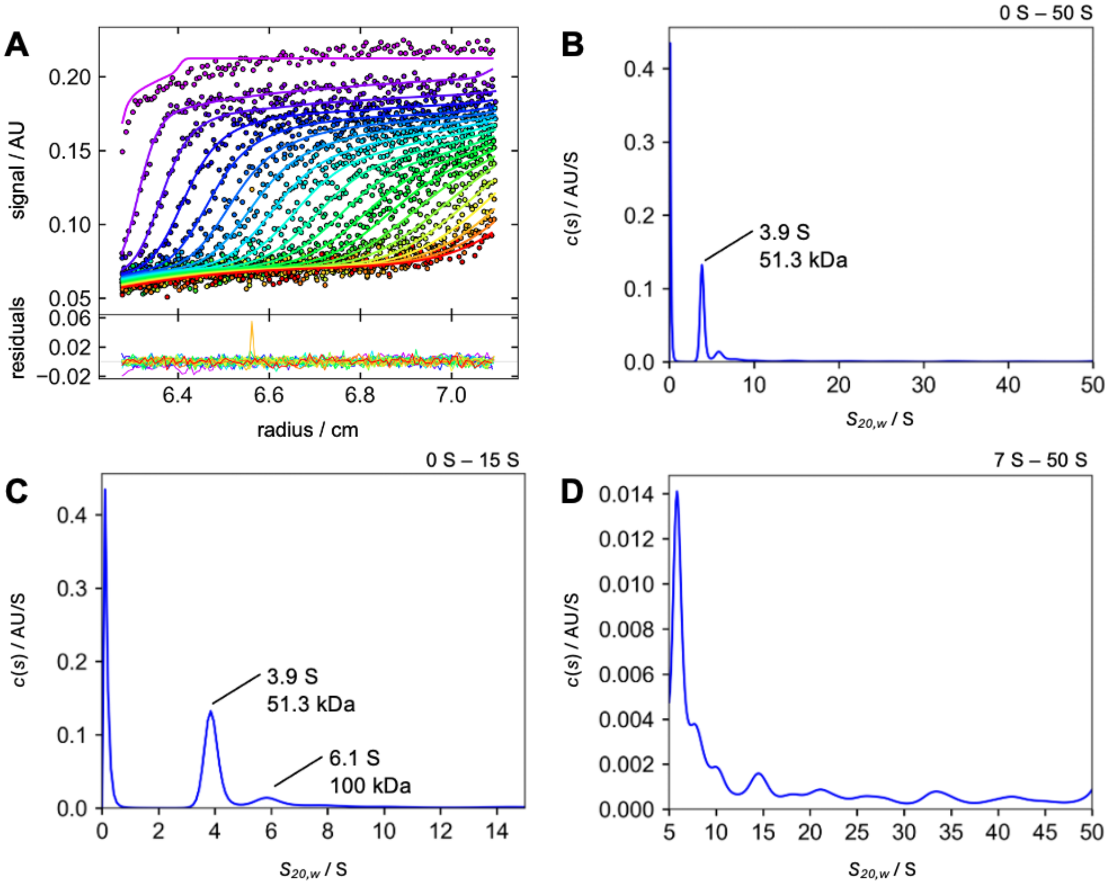


Figure 24: Sedimentation profile (A) and single *c*(*s*) distributions (B-D) of GST-Asp1$^{365\text{-}920}$ reconstituted with 10-fold excess of iron and sulfur measured with AUC. The sample was measured in 10 mM sodium phosphate, 1% glycerol, pH 7.4, at 20 °C and 60,000 *x g*. The *c(s)* distributions were normalized to one, according to the area under the curve for the complete *s*-value range from 0 to maximum *s*-value. A) temporal concentration change as signal in AU plotted as function of the radius of the measuring cell in cm. The coloring of the profile is a time scale, where blue signifies the beginning and red the end of the measurement. B) data fitted on the *c*(*s*) model in the *s*-value range of 0-50 S. The *s*-values are standardized on a $s_{20,w}$ value, with the absorbance being plotted against the $s_{20,w}$ value. C) data fitted on

the $c(s)$ model in the s-value range of 0-15 S. The s-values are standardized on a $s_{20,w}$ value, with the absorbance being plotted against the $s_{20,w}$ value. D) data fitted on the $c(s)$ model in the s-value range of 7-50 S. The s-values are standardized on a $s_{20,w}$ value, with the absorbance being plotted against the $s_{20,w}$ value.

4.6 Reductive chelation of the Asp1 Fe-S cluster

The experiments presented earlier throughout this study focused on inducing states of the Asp1 pyrophosphatase domain where different amounts of iron ions are ligated and characterizing them using a variety of methods. During the reductive chelation assays an attempt was made at gathering information about the localization of the binding site of the cluster by removing all cofactor from the protein sample.

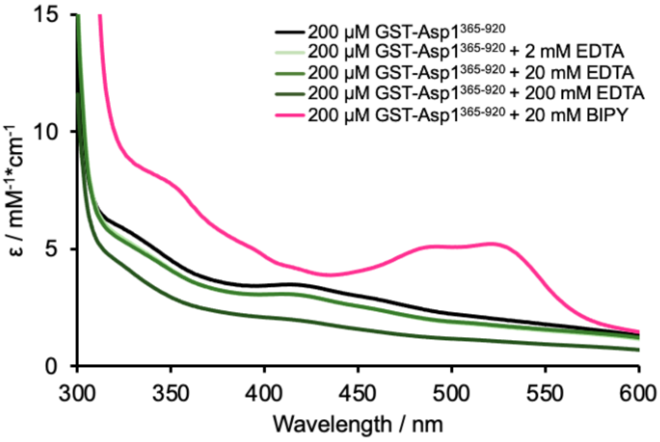


Figure 25: Electronic spectra of GST-Asp1$^{365-920}$ after conducting the reductive chelation assay with diferrent reagents. The molar extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). In this assay 2,2-bipyridine (BIPY) and ethylenediaminetetraacetic acid (EDTA) were tested as chelators due to their different chemical properties. EAS data was collected just after adding the chelating agents in different molar excess to the protein sample. In black the EAS spectrum of the used GST-Asp1$^{365-920}$ sample without any chelator and three spectra of GST-Asp1$^{365-920}$ under the chelating influence of EDTA in green shades.

EDTA is not a suitable chelator to remove all iron from the Asp1 pyrophosphatase domain under reducing conditions. Samples of GST-Asp1$^{365-920}$ were prepared according to section 3.12.1 for reductive chelation. It is noticeable that the molar extinction coefficient decreases slightly upon addition of 10- and 100-fold molar excess of EDTA to protein concentration but achieves significantly lower values when a 1000-fold excess is used (2.04 mM^{-1}*cm^{-1}at 410 nm, Figure 25). The typical absorption bands of the Asp1 pyrophosphatase domain are still perceivable for all EDTA-containing samples but the one with the highest amount of the chelator.

The EAS spectrum of GST-Asp1$^{365-920}$ with 100-fold molar excess of BIPY is very different in comparison to the EDTA spectra with two peaks at 485 nm and 522 nm (5.05 mM^{-1}*cm^{-1}at 485 nm, 5.20 mM^{-1}*cm^{-1}at 522 nm, Figure 25) and a more prominent shoulder at 320 nm.[151] The binding kinetics of BIPY are very fast as a spectroscopic red shift occurs upon addition of the chelator which is met with a colour development to carmine red, noticeable with the unaided eye.[151] Excess reagents were removed using SEC and a 16/600 HiLoad Superdex 200 pg column (Figure *29*). EA spectra were again recorded after desalting was completed to check the chelation status.

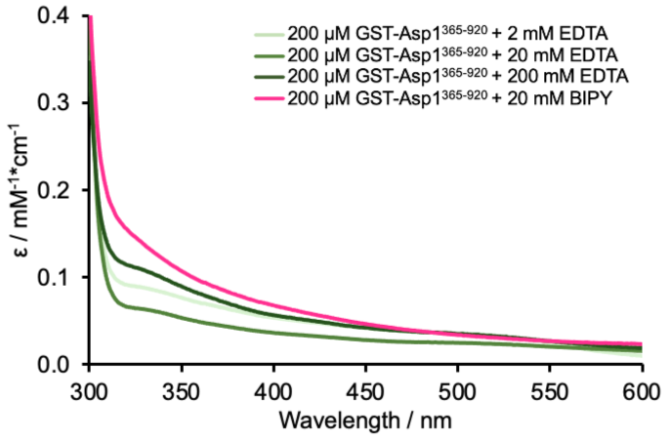


Figure 26: Electronic spectra of GST-Asp1$^{365\text{-}920}$ after conducting the reductive chelation with different reagents and SEC for desalting. The molar extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). Every illustrated curve is the average of at least three spectra recorded from the fractions that were collected for each sample.

Despite the overall higher background signal, the addition of 100-fold molar excess of BIPY is most effective for the complete chelation of the Asp1 Fe-S cluster, so that no residual absorption bands are visible in the EAS spectrum. Opposite to the spectrum of GST-Asp1$^{365\text{-}920}$ supplemented with 20 mM BIPY the EDTA-containing samples still exhibit the typical shoulder at 320 nm even if the reached molar extinction coefficient is below 0.15 mM^{-1}*cm^{-1}at the specified wavelength. Despite the same incubation time and three different molar excess ratios it was not possible to chelate all iron ions from the Asp1 pyrophosphatase domain using EDTA (Figure 26).[152]

EAS alone is not enough to quantify the remaining iron amount per monomer protein, so this technique was once again combined with ICP-MS for a better understanding of the Fe-S occupancy after reductive chelation. ICP-MS samples were collected from all fractions belonging to differently treated samples, the results being depicted in Figure 27. It is noticeable that the Fe/monomer levels are lowest in the sample treated with 100-fold molar excess of BIPY. Slightly higher levels of Fe/monomer can be observed for the samples containing EDTA, whereas the sample of GST-Asp1$^{365\text{-}920}$ treated with 200 mM EDTA unexpectedly reaches Fe/monomer values almost twice as high (0.25 Fe/monomer, Figure 27.A) as for the sample incubated with BIPY. Interestingly, all samples contain much more zinc than iron per monomer protein despite the excessive amounts of chelator. The samples treated with 2 mM and 20 mM EDTA reach values of approx. 0.35 Zn/monomer whereas the zinc contents of the 200 mM EDTA and 20 mM BIPY are double as high (Figure 27.A). Studies regarding bipyridine-bound supracomplexes show that a 10-fold molar excess of the chelator suffices in fully complexing all iron (Fe^{2+}) ions from a solution and that three BIPY molecules are needed to chelate an iron ion. The amount of sodium dithionite (NaDT) added to the sample during reductive chelation fully enables the reduction of ligated iron (Fe^{3+}) to iron (Fe^{2+}) which is then easily sequestered by BIPY. The most probable explanation for the elevated zinc levels is the provenience of the GST-Asp1$^{365\text{-}920}$ sample which as-isolated was shown to bind approx. 1 Zn/monomer. A rationale why the excessive amount of EDTA is not sufficient to complex zinc ions present in the sample lies in the higher affinity of EDTA towards trivalent ions. [152] Ferric iron has a higher charge density and interacts better with the negative groups of EDTA.

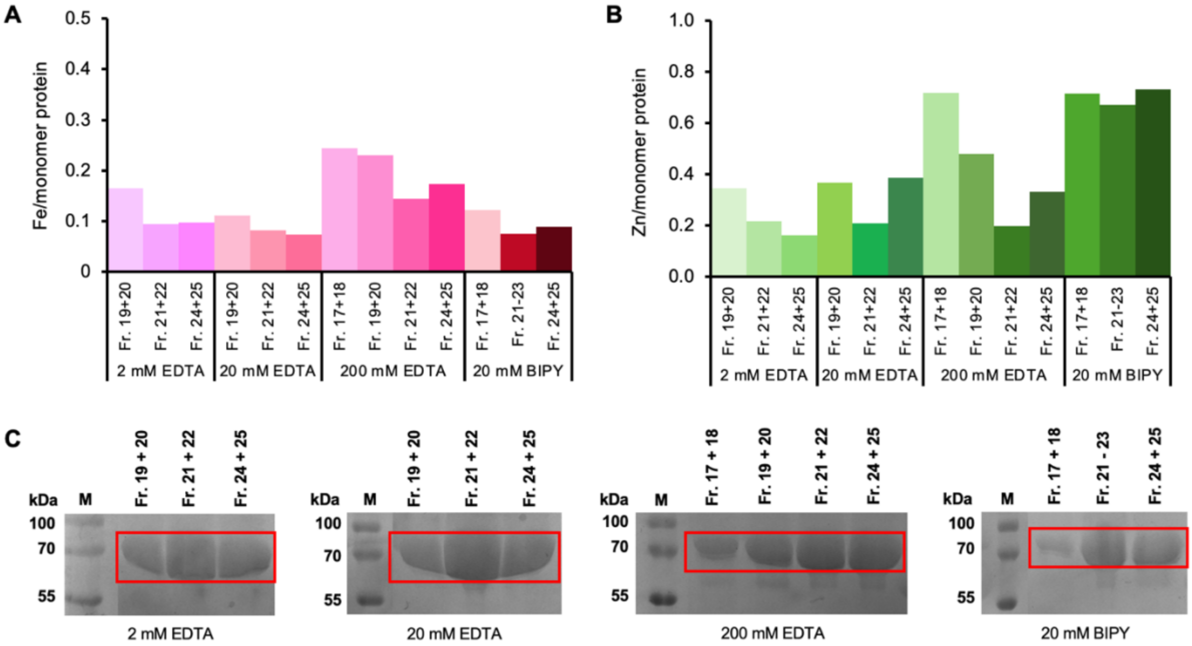


Figure 27: Overview of metal content and purity of the samples used in the reductive chelation assay. A) Iron content of different fractions during SEC after reductive chelation. B) Zinc content of different fractions during SEC after reductive chelation. C) Overview of denaturing SDS-PAGES of the collected fractions during SEC.

The overall observations led to a more detailed analysis of the chemical structure and properties of the two used reagents (Figure 28). EDTA is a polar compound often used for chelation and known best for forming strong complexes with di- and trivalent metal ions such as Fe^{2+}, Mg^{2+}, Zn^{2+}, Ca^{2+}, Ni^{2+}.[152] Due to its size and polarity is can easily sequester metal ions bound to the surface of proteins or in accessible areas. Bipyridine, on the other hand, is a highly hydrophobic compound which is sterically unhindered compared to EDTA and has better access to inaccessible areas of the protein. It was reported that BIPY is more selective towards ferrous (Fe^{2+}) than ferric iron (Fe^{3+}). The octahedral geometries that result from the orbital overlap and electronic interactions lead to a favourable and stable complex. On the other hand, ferric (Fe^{3+}) iron is more positively charged and exhibits increased ligand field stabilization conditions that cannot be met by BIPY. This results in a less stable complex which is not favoured. Overall, BIPY is a better iron chelator than EDTA under reducing conditions. This could potentially indicate the iron ions are more accessible to a small and hydrophobic chelator such as BIPY under reducing conditions. This could implicate that a considerable amount of iron bound by the Asp1 pyrophosphatase domain occurs as ferric (Fe^{3+}) iron under physiological conditions but is sequestered by the protein making it less accessible for large molecules such as EDTA, which is the better chelator for trivalent ions. A possible explanation for the data might be that the ferric ($Fe^{3+)}$ iron which would be the predominant species under physiological conditions is slightly shielded by the pyrophosphatase domain. This rationale would also be in good agreement with the data from previous studies, that characterized the Asp1 Fe-S cluster as a [2Fe-2S] cluster, where one iron ion occurs in the Fe^{3+} oxidation state and the second transitions between the Fe^{2+} and Fe^{3+} oxidations states.[153]

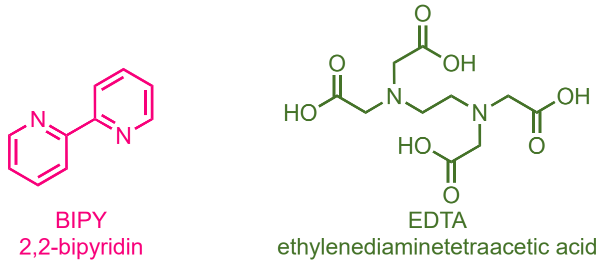


Figure 28: Chemical structure of 2,2-bipyridin (BIPY) and ethylenediaminetetraacetic acid (EDTA) which were used as chelating agents.

The SEC runs conducted at the end of the reductive chelation experiment did not serve only to remove the reagents but also to estimate the oligomerization state in the absence of the cofactor. Figure 29 shows that all four samples are very diversly dispersed from monomer to trimers and high order oligomers. In Figure 29.A is noticeable that a 10-fold elevation in the molar excess of the chelator leads to a different kind of oligomeric state than in the samples incubated with only 2 and 20 mM EDTA. The same can be observed for BIPY where approx. a third of the sample is oligomeric dispersed to the same grade as the sample treated with maximal excess of EDTA (Figure 29.B, Figure 29.C). The described oligomers are so large that they range between 988 and 1300 kDa and are comprised in the void volume of the SEC column.

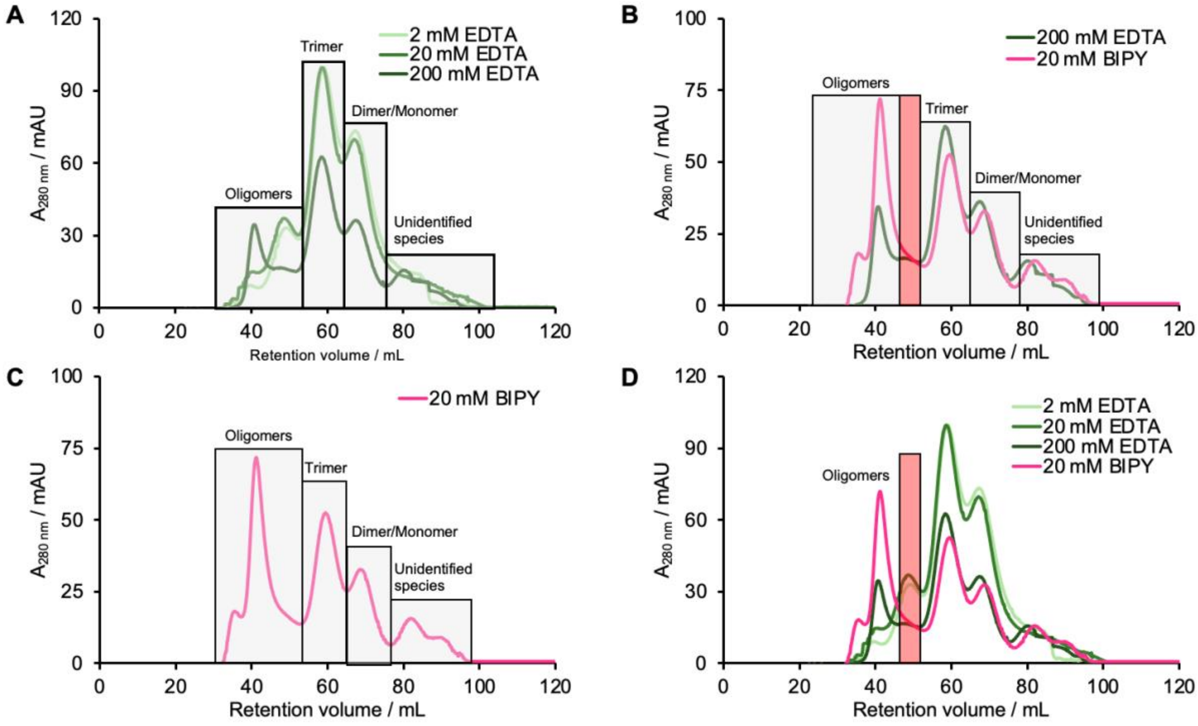


Figure 29: SEC data from the reductive chelation assays. The SEC was conducted using a HiLoad 16/600 Superdex 200 pg column. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis. A) SEC chromatograms of the samples treated with EDTA. B) SEC chromatograms of the samples treated with 200 mM EDTA and 20 mM BIPY. C) SEC chromatograms of the BIPY-containing sample. D) SEC chromatograms of

all samples from the reductive chelation experiment. A missing oligomeric fraction was marked in red square as a visual aid.

Overall, it can be observed that the oligomerizations states are most divers when the Asp1 Fe-S cluster is not occupied at all. A certain amount of oligomerization can certainly be attributed to the presence of the GST-tag that makes such fusion proteins prone to dimerization.[135,136,154–158] Moreover, it has been reported that Asp1 is mostly dimeric dispersed without the influence of any affinity tag.[1] In addition to that, the use of BIPY appears to double the high-grade oligomers even though it is the better chelator for GST-Asp1$^{365\text{-}920}$.

4.7 Reduction and oxidation of the Asp1$^{365\text{-}920}$ Fe-S cluster

The aim of this experiment was to assess the redox sensitivity of the Fe-S cluster in GST-Asp1$^{365–920}$. Both reducing and oxidizing conditions were applied, and changes in the cluster were monitored using EAS and CD spectroscopy. Previous studies on Asp1 revealed that the Fe-S cluster is EPR silent in the as-isolated form, suggesting that it is fully oxidized and exists in the $[2Fe\text{-}2S]^{2+}$ state. [3] Attempts to reduce the cluster with one or two equivalents of sodium dithionite (NaDT) failed to yield an EPR-active species, indicating that these conditions were insufficient to reduce the cluster. Since EPR spectroscopy is a powerful method for characterizing Fe-S cofactors, determining whether the cluster can be reduced under suitable conditions is of great interest. Successfully reducing the Fe-S cluster could reveal valuable insights in regard to electron transfer. For the reduction experiments a stepwise titration of GST-Asp1$^{365\text{-}920}$ with different reductants was performed to explore whether the Fe-S cluster can be chemically reduced, and to determine the conditions necessary to achieve this. The three reagents used for the reductive experiments were NaDT, DTT and reduced GSH. The reducing agents were added stepwise to GST-Asp1$^{365\text{-}920}$.

On the other hand, the motivation for using oxidative conditions on the Fe-S was to assess its potential sensitivity to oxygen, a factor that may be relevant to its physiological function as an oxygen sensor in vivo. In this study, the oxidation of the metal ions coordinated in the Fe-S cluster was conducted under aerobic conditions (exposure to atmospheric oxygen). An identical sample of the GST-Asp1$^{365\text{-}920}$ was prepared for the oxidation experiment. The sample was incubated under aerobic conditions on ice and every 30 min air was actively pipetted in the sample. One such step was described as one equivalent in the oxidation experiment. There are no equivalents specified as in the case of the reduction experiments, but the legend shows the total amount of time the sample has been incubated for. A step in this experiment consisted of incubation of GST-Asp1$^{365\text{-}920}$ for 30 min under aerobic atmosphere after active pipetting of air bubbles in the sample (section 3.12.3). Here the oxygen from the atmosphere serves as oxidation reagent for the Fe-S cluster.

The spectra are depicted in different colours to highlight the usage of another reduction or oxidation agent. All EAS spectra are depicted as a colour gradient to emphasize the spectral changes caused by the incubation with O_2 or the repetitive addition of NaDT, DTT, GSH. The curve in black represents the EA spectrum of the GST-Asp1$^{365\text{-}920}$ starting solution at the beginning of the experiment. The spectrum of the untreated sample exhibits two peaks at 410 nm and 460 nm with an extinction coefficient of 5-6 $mM^{-1}*cm^{-1}$ (Figure 30) and a shoulder peak at 320 nm with an extinction coefficient of 9 $mM^{-1}*cm^{-1}$ (Figure 30).

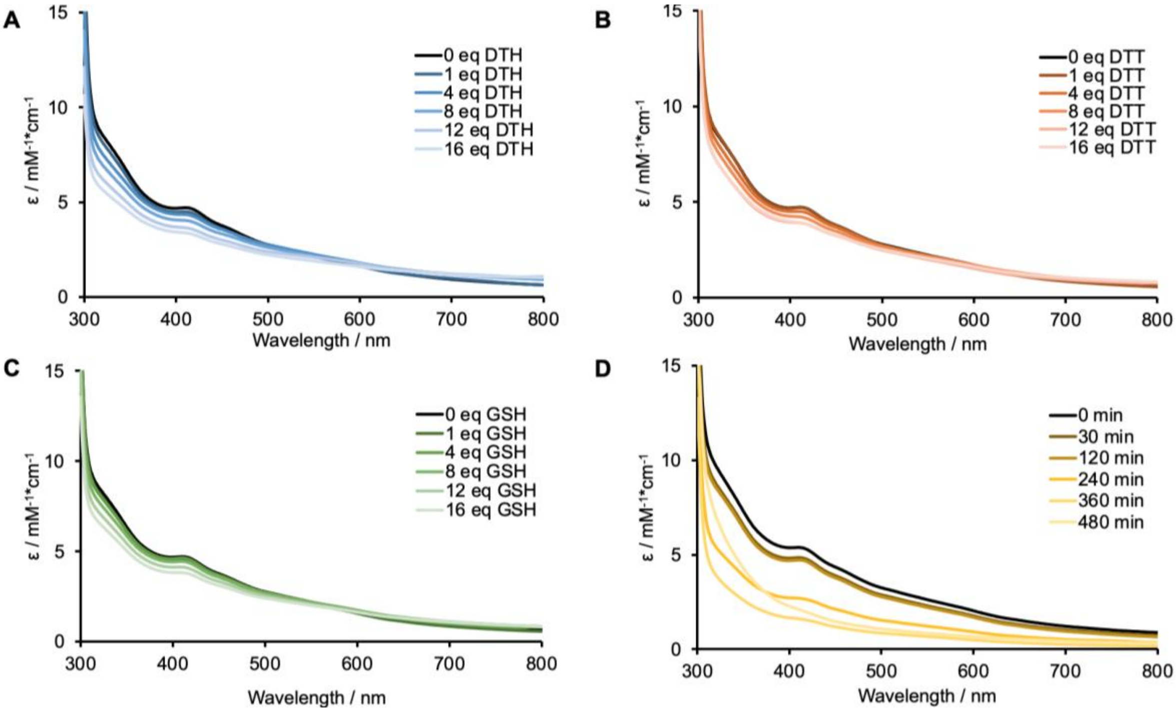


Figure 30: Electronic absorption spectra of GST-Asp1$^{365\text{-}920}$ during incubation with different reductive and oxidative agents. The molar extinction coefficient ε in mM^{-1}*cm^{-1} was plotted as a function of the wavelength in nm. Reductive agents were applied to the protein sample in 6 different amounts. Air was actively pipetted in the sample every 30 min until complete oxidation was reached. A) reduction with NaDT, B) reduction with DTT, C) reduction with GSH, D) oxidation with oxygen (O_2).

It can be observed that upon addition of the reagents all spectral features gradually decrease in intensity until the curves appear flattened. In case of NaDT the decrease in the signal occurs with the first equivalent already and continues in equal distances until the last EA spectrum at 16 eq does not exhibit any spectral features. When DTT and GSH were added the first indications of reduction of the Fe-S cluster were first observed after 4 and 8 eq of molar excess, respectively. Full reduction was reached after the addition of 16 eq of each reductant to the protein solution. The reduction of GST-Asp1$^{365\text{-}920}$ can be successfully conducted using NaDT as optimal reducing agent, however a higher number of equivalents is needed to achieve full reduction. Furthermore, the incubation of the same protein sample for 30 min under aerobic atmosphere shows a decrease in the signal intensity. The incubation for 240 min is significant as the molar extinction coefficient is twice as low as it was prior to oxygen incubation (3 mM^{-1}*cm^{-1} at 410 nm, Figure 30). The decrease in the signal indicates that the second iron ion of the cluster is being oxidized since the Asp1 Fe-S cluster was identified as oxidized subsequent to protein isolation. Full oxidation of the metal ions and possibly oxidative degradation is reached after 480 min of incubation under aerobic conditions.

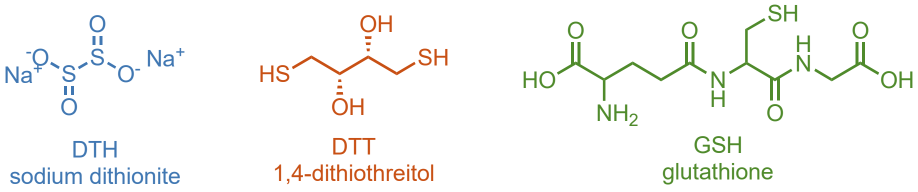


Figure 31: Chemical structures of the used reducing agents: NaDT, DTT and glutathione (GSH).

NaDT is the strongest reductant (-0.66 V) among the tested reagents, whereas DTT has a redox potential of -0.33V and is thus a milder reductant compared to NaDT (Figure 30.A, Figure 31).[159–161] The two thiol groups of DTT enable the efficient reduction of cysteine residues in proteins, however the redox potential of DTT is not enough to reduce Fe ions as effectively as NaDT (Figure 30.A, Figure 30.B). On the other hand, GSH which has lately proven indispensable in the mitochondrial iron-sulfur cluster biogenesis has been occasionally involved in the ligation of Fe-S in glutaredoxins.[149] This reducing agent has a similar redox potential as DTT which amounts at -0.264 V and is characterized as a mild reductant.[162] The peptide character of the compound and the thiol group are the reason while GSH is not only known for its redox properties but can contribute to coordination of Fe-S cluster (Figure 31). The reduction with GSH can be classified as mild as values no larger than -50 deg*cm^2*dmol^{-1} are recorded during the CD spectroscopic measurements as in the case of DTT (Figure 30.B, Figure 30.C).

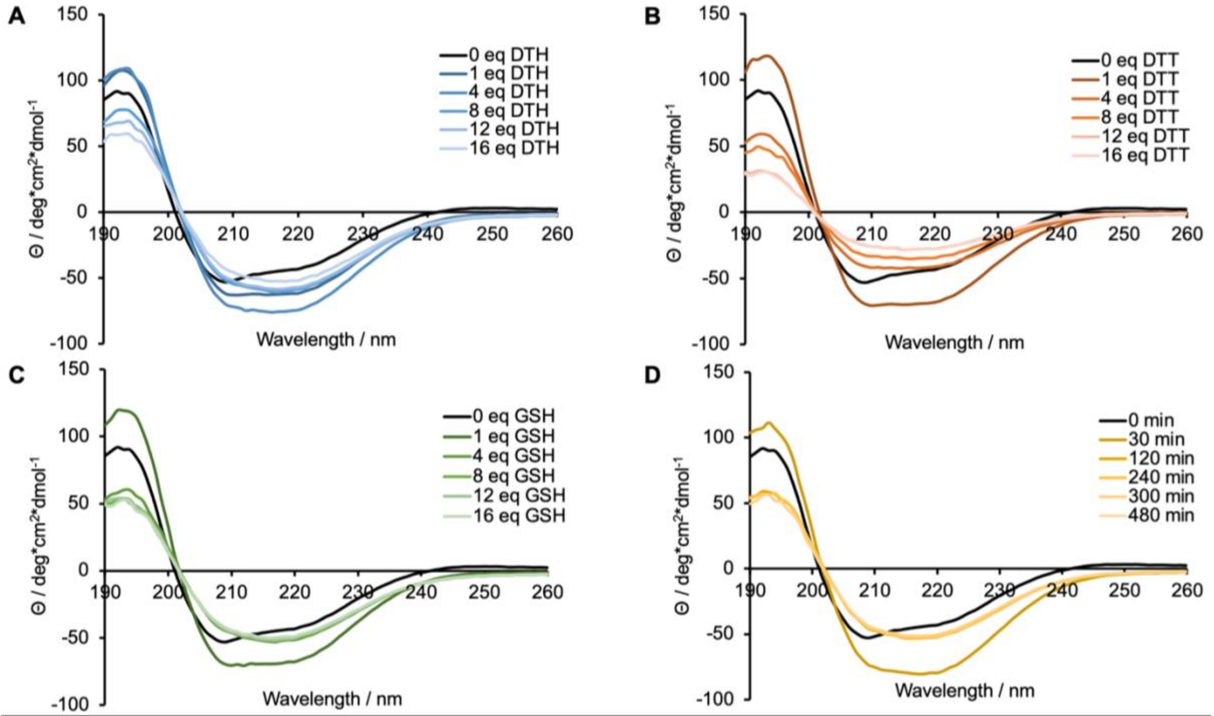


Figure 32: Single CD spectra of GST-Asp1$^{365\text{-}920}$ during incubation with different reductive and oxidative agents. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm. Reductive agents were applied to the protein sample in 6 different amounts. Air was actively pipetted in the sample every 30 min until complete oxidation was reached. A) reduction with NaDT, B) reduction with DTT, C) reduction with GSH, D) oxidation with O_2.

A sample for CD spectroscopy was collected out of every sample measured with EAS. All CD samples were prepared and normalized on the same chosen concentration that will yield a suitable signal to noise ratio. Interestingly, a significant difference can be observed between the GST-Asp1$^{365\text{-}920}$ sample without any reagent and the one after treatment with the first equivalent of reductant or oxidant (Figure 32). This might indicate that the protein assumes a different structural state upon the addition of the reagents. In all CD spectra where reductants were used it can be observed that the peak under 210 nm typical for helix is less prominent with each added equivalent of reducing agent. This indicates that the Asp1 Fe-S cluster influences the secondary structure of the pyrophosphatase domain. The cluster might stabilize helical structures in the oxidized cluster state which is found in the protein as-isolated state. Interestingly, the signal under 220 nm, typical for β-sheets, appears to be more prominent with each equivalent of reducing agent. All other curves from the O_2 oxidation experiment overlap and reach values of 50 deg*cm^2*dmol^{-1}.

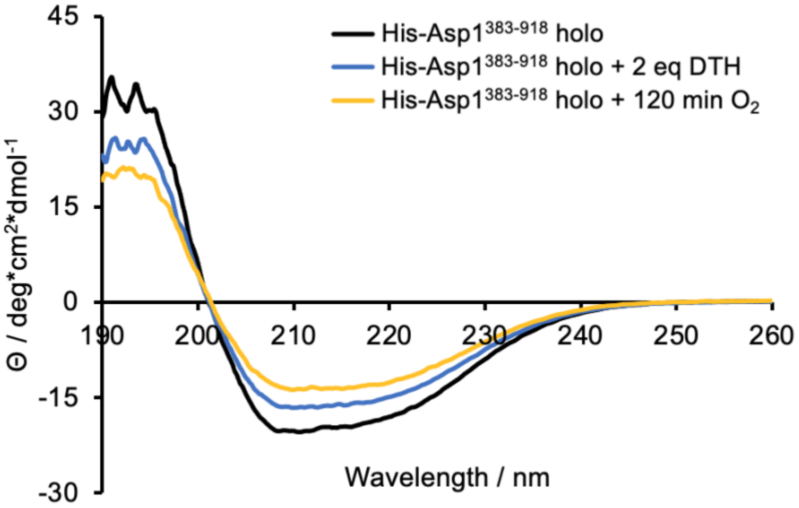


Figure 33: Influence of NaDT and O_2 on CD spectral properties of His-Asp1$^{383-918}$. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm.

The reduction assay was performed on the shorter version of the His-tagged pyrophosphatase domain to investigate if the previous findings are also applicable on a variant with a much smaller affinity tag and no sequence areas irrelevant to the secondary structure. Figure 33 shows that the incubation for 120 min under aerobic atmosphere is much more efficient in altering the coordinated iron species and thus the cluster than 2 equivalents of NaDT. However, a decrease of the minimum at 209 nm under the reference values (GST-Asp1$^{365-920}$, black curve, Figure 32) for the GST-tagged construct could be observed starting from 8 eq NaDT whereas this effect is noticeable already with 2 eq for the His-Asp1$^{383-918}$ (Figure 33). However, the usage of a His-tag was shown to have a negative impact on the stability of the protein variant upon longer storage times whereas the GST-tagged variant proved more stable. At the same time, the GST-tag is a fusion protein with a size of 26 kDa, which contains α-helices and β-sheets that contribute to the secondary structure. This is well reflected among all CD spectra that correspond to GST-tagged samples.

4.8 Analysis of secondary structure of the Asp1$^{365\text{-}920}$ pyrophosphatase domain

CD spectroscopy was used to obtain preliminary insights into the distribution of secondary structure elements of the pyrophosphatase domain of *S. pombe* Asp1. The GST-tagged wildtype Asp1$^{365\text{-}920}$ was analyzed in different cluster-bound states, 0 Fe/monomer indicating the apo state where no cofactor is present. Immediate after protein purification the pyrophosphatase domain was shown to bind only 1 Fe/monomer, whereas holo states with 2 and 4 Fe/monomer can be achieved upon chemical reconstitution with 5- or 10-fold equivalents of iron and sulfur sources, respectively. The quadruple mutant variant that was previously characterized was measured as part of this study as well as a negative control and comparison to the apo Asp1$^{365\text{-}920}$ construct (0 Fe/monomer). The interpretation of the CD data was complemented with an analysis based on the BeSTSel tool, which provided more detailed insights in the percentual distribution of secondary structure elements of the different states of GST-Asp1$^{365\text{-}920}$.

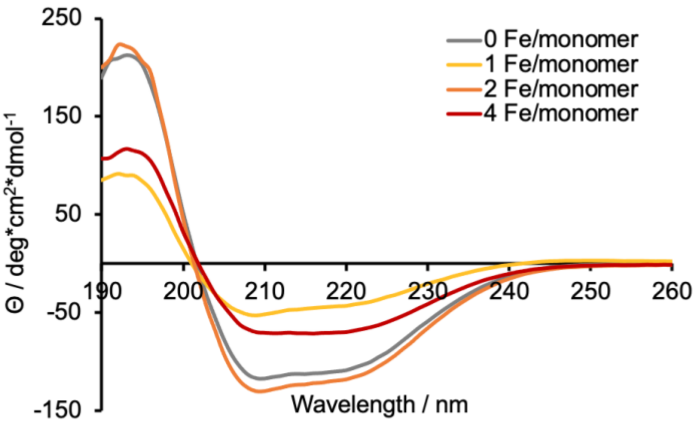


Figure 34: Overview of single CD spectra of GST-Asp1$^{365\text{-}920}$ with different amounts of iron. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm.

The graphically illustrated CD data of the apo GST-Asp1$^{365\text{-}920}$ sample shows two minima at 209 nm and 222 nm and one maximum at 195 nm (Figure 34). The peak at 209 nm is characteristic for helix structures, whereas the one at 222 nm for β-sheets. The data resulting from the apo GST-Asp1$^{365\text{-}920}$ describes a secondary structure consisting mainly of helix (Supplementary table 6, 28.8%) and random coil (Supplementary table 6, 40.7%). The remaining categories of secondary structure elements (antiparallel β-sheets, parallel β-sheet and turns) present similar values from 8.5% to 11.6% (Supplementary table 6). On the other hand, the coordination of 1 Fe/monomer causes a drastic decrease in the mean residues ellipticity (Figure 34). The curve appears to exhibit only one predominant minimum at 209 nm, whereas the second one almost disappears due to the curve being more flattened from 220 nm to 260 nm. At the same time, the calculated proportion of antiparallel β-sheet increases to 30.6% (Supplementary table 6), which is approximately three times higher than in the apo

GST-Asp1$^{365\text{-}920}$ sample. This shift is accompanied by a reduction in the estimated α-helical and β-turn content by about 20% and 8%, respectively. However, given the intrinsic limitations of CD-based secondary structure estimations and the relatively broad error margins of such deconvolutions, these differences should be interpreted cautiously. Upon addition of a second Fe atom per monomer (2 Fe/monomer), the overall shape of the CD spectrum becomes more similar to that of the apo form (Figure 38). The estimated α-helical content increases to 32.2%, the highest value observed among the analyzed samples, suggesting a potential trend toward helix stabilization. Nevertheless, due to the moderate resolution of CD spectroscopy, further methods would be required to confirm these structural transitions unambiguously.

CD spectroscopy is not only a reliable spectroscopic tool for estimation of average secondary structure of proteins, but it also serves to monitor folding and unfolding under various conditions and determine protein specific parameters. Protein denaturation data sets were recorded from the same set of samples which was analyzed in the previous single CD spectra. The manner of sample preparation was the same as the samples used for the denaturation curves coincide with the ones measured for the single CD spectra. Collection of such data sets was helpful to determine the melting temperature of GST-Asp1$^{365\text{-}920}$ in different states and thus estimate if the Fe-S cluster type has any influence on protein stability and folding. Based on Figure 35 the minima at 209 nm and 222 nm of the four samples which were discussed in Figure 34 and the differences among them are better distinguishable due to the compilation of many single spectra. Among all measurements, the 4 Fe/monomer data set exhibits the most different course as not only the amplitude is different but also the compression of the curves with increasing temperatures. The 0 and 2 Fe/monomer samples show great similarities in the graphical course although the amplitudes do not coincide. The data set for 1 Fe/monomer looks similar to the two mentioned before, however the ratio between the two minima at 209 nm and 222 nm is different making the first one more visible.

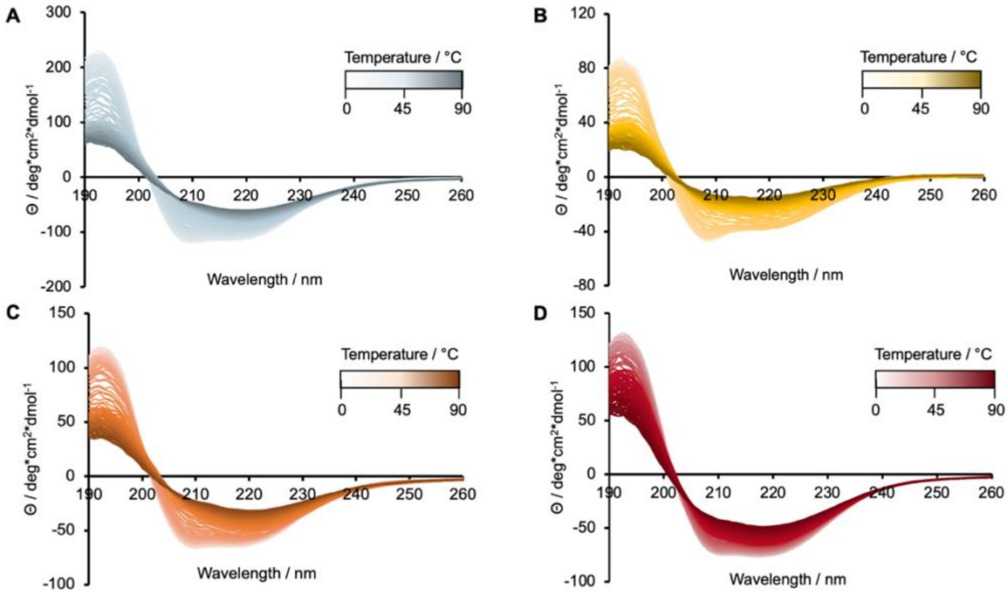


Figure 35: Denaturation curves of GST-Asp1$^{365\text{-}920}$ in four different states from apo to fully loaded holo form. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted

to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm. A) GST-Asp1$^{365\text{-}920}$ 0 Fe/monomer (apo). B) GST-Asp1$^{365\text{-}920}$ 1 Fe/monomer (as-isolated). C) GST-Asp1$^{365\text{-}920}$ 2 Fe/monomer. D) GST-Asp1$^{365\text{-}920}$ 4 Fe/monomer. The intensity of the colours correlate with increasing temperatures at which the data sets were recorded.

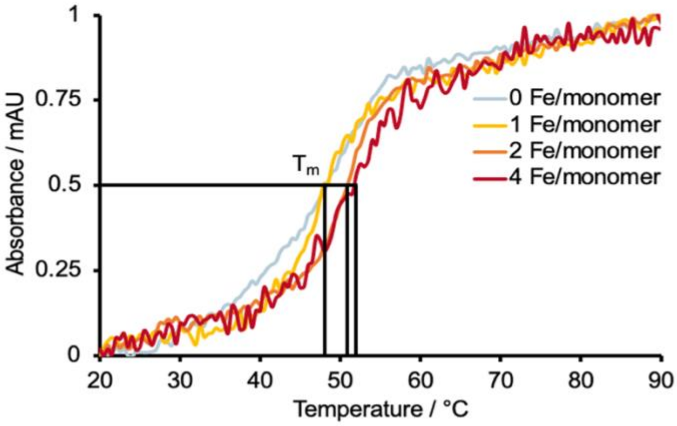


Figure 36: Normalized absorbance at 209 nm for the determination of the melting temperature of GST-Asp1$^{365\text{-}920}$ in different cluster-bound states. The normalized absorbance was plotted as a function of temperature resulting in saturation curves. The melting temperature T$_m$ is the corresponding value on the x-axis of the half-maximal absorbance (turning point).

Figure 36 indicates an increase in the melting temperature T$_m$ with higher amounts of iron per monomer protein. It appears that the coordination of only one iron ion per monomer does not make any difference in terms of protein stability as the melting temperature of both GST-Asp1$^{365\text{-}920}$ apo (0 Fe/monomer, Figure 36) and GST-Asp1$^{365\text{-}920}$ as-isolated (1 Fe/monomer, Figure 36) remains constant at 48.3 °C. As soon as the 2 Fe/monomer state is achieved the protein remains stable until 51 °C, indicating the coordination of one further iron atom improves stability. The melting temperature is increased up to 51.8 °C when two more iron atoms per monomer protein are present, which is a slightly minimal shift despite the total metal atom amount being doubled. Summarizing it can be said the one Fe-S cluster of the *S. pombe* Asp1 pyrophosphatase domain has a rather significant influence on the overall protein stability. The sample with 4 Fe per monomer generated via chemical reconstitution with 10x excess of iron and sulfur sources shows no significant increase in protein stability, this being a hint that it might be localized in the active site of the pyrophosphatase domain and might play a role in the regulation of the pyrophosphatase activity.

The quadruple mutant GST-Asp1$^{365\text{-}920\ QM}$ was used to investigate of the Fe-S cluster influences the secondary structure of the Asp1 Fe-S. To that purpose single spectra of the two apo (GST-Asp1$^{365\text{-}920\ QM}$ 0 Fe/monomer, GST-Asp1$^{365\text{-}920}$ 0 Fe/monomer) variants were recorded under similar conditions and compared (Figure 37). It can be observed that the signals of apo GST-Asp1$^{365\text{-}920}$ (0 Fe/monomer, Figure 37) and GST-Asp1$^{365\text{-}920\ QM}$ (0 Fe/monomer, Figure 37) almost coincide, including the position of the two minima at approx. 209 nm and 222 nm. The slight difference in the ellipticity signal amounts to approx. 11 deg*cm^2*dmol^{-1} and can be interpreted as insignificant in terms of changes in the distribution of secondary structure elements. This is further supported by the data shown in Supplementary table 6, with only a 2.6% and 4.6% difference in the helix and antiparallel β-sheet amount, respectively. For GST-Asp1$^{365\text{-}920\ QM}$ the total helix and β-sheet amounts (parallel and antiparallel) are balanced at approx. 25%. The highest percentage in the secondary structure consists in random coil with 38.9%, indicating a certain flexibility in the overall tertiary structure of apo GST-Asp1$^{365\text{-}920}$ variants (0 Fe/monomer QM, 0 Fe/monomer, Figure 37). The substitution of cysteine residues C607S, C663S, C864S and C879 leads to the complete loss of cofactor in GST-Asp1$^{365\text{-}920}$ making it structurally similar to a GST-Asp1$^{365\text{-}920}$ version where alle four cysteines are intact but only the cofactor is removed (GST-Asp1$^{365\text{-}920}$ apo, 0 Fe/monomer).[3] Based on the results and a previous study it appears the four cysteines contribute in some amount to cofactor binding.

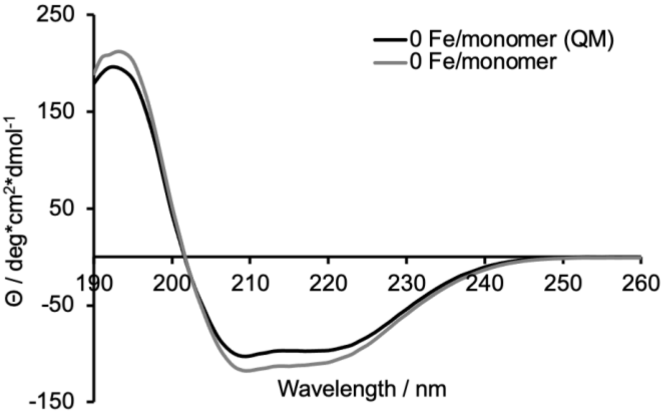


Figure 37: Comparison of single CD spectra of GST-Asp1$^{365\text{-}920\ QM}$ and GST-Asp1$^{365\text{-}920}$ wildtype. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm.

As in the case of the previously described states, the determination of the melting temperature of GST-Asp1$^{365\text{-}920\ QM}$ was of interest to investigate if the Asp1 Fe-S affects the protein stability and thus influence the secondary structure.

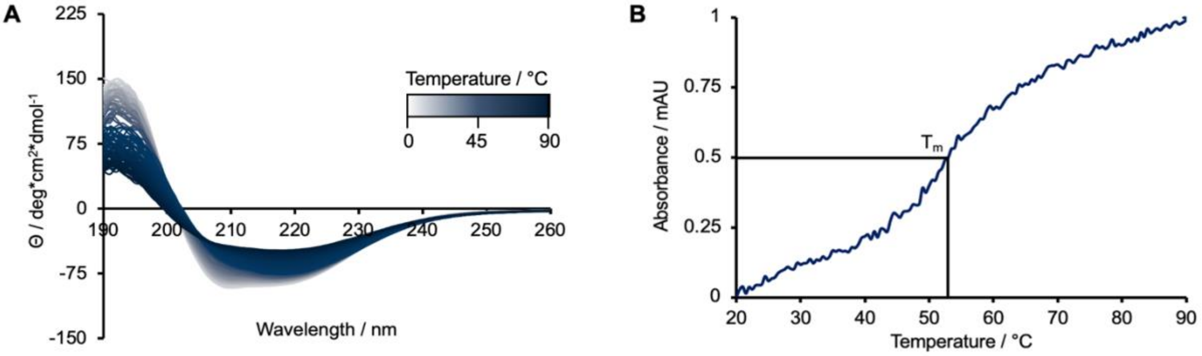


Figure 38: Denaturation curves for GST-Asp1$^{365\text{-}920\ QM}$ (A) and normalized absorbance for the determination of the melting temperature (B). A) Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm. B) The normalized absorbance was plotted as a function of temperature resulting in saturation curves. The melting temperature (T$_m$) is the corresponding value on the x-axis of the half-maximal absorbance (turning point).

The time interval measurement was conducted in a similar manner however, the data did not yield a fully saturated curve for the determination of the melting temperature. Multiple measurements of this apo construct were conducted to optimize the resulting data by varying different parameters as sample concentration and temperature interval. The further attempts were unsuccessful as a well-established plateau in the saturation curve was not achieved. This parameter might be lower and coincide with the data set of apo GST-Asp1$^{365-920}$ as a fully saturated curve would shift its turning point more to the left. Summarizing, the acquired data for GST-Asp1$^{365-920\ QM}$ is not as earlier-mentioned data sets for different GST-Asp1$^{365-920}$ wildtype states, thus making the determination of the melting temperature for this quadruple mutant variant impossible (Figure 38).

The characterization of the substrate-bound state of GST-Asp1$^{365-920}$ became of equal interest, as the previous investigations of the secondary structure were conducted by removing the cofactor. For that, reconstituted GST-Asp1$^{365-920}$ (4 Fe/monomer) was incubated with 10-fold excess of 1,5-methylene bisphosphonate inositol tetrakisphosphate (1,5-$(PCP)_2$-IP_4), which was kindly provided by the Fiedler group. The acquisition of the single spectra was performed similarly.

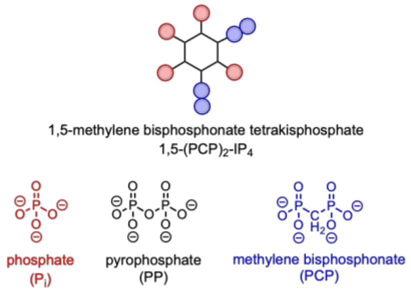


Figure 39: Chemical structure of 1,5-$(PCP)_2$-IP_4 and different species of phosphate functional groups that occur in this molecule.

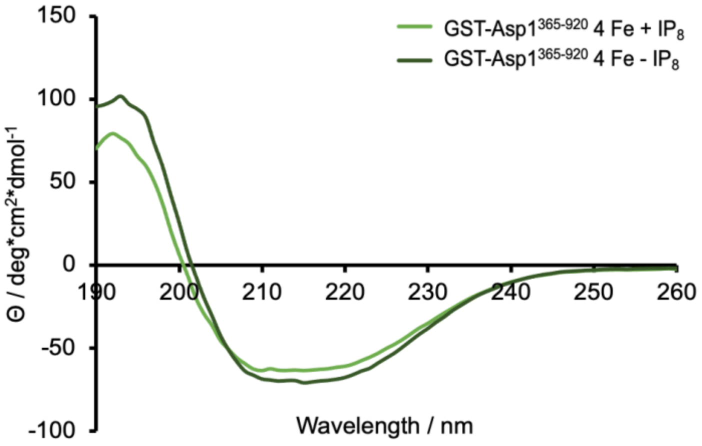


Figure 40: Comparison of single CD spectra of GST-Asp1$^{365-920}$ 4 Fe/monomer with and without 10-fold excess of 1,5-$(PCP)_2$-IP_4. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm.

One of the natural substrates of GST-Asp1$^{365-920}$ is 1,5-bisphosphoinositol tetrakisphosphate, usually abbreviated as 1,5-IP_8 because of harbouring a total of eight phosphate groups. Four of the phosphate groups occur as pyrophosphate functional groups at positions 1 and 5 of the hexose ring. A characteristic of the pyrophosphates is the highly energetic bond between the two phosphorus atoms which is categorized as an anhydride bond. Due to exothermic reactions this bond is usually hydrolysable and such compounds tend to be highly unstable during both synthesis and storage.[163] This is the rationale behind using protective groups on such labile chemical functionalities, in this case using a substitution of the β-oxygen atom with a carbon and thus transitioning from a pyrophosphate to a methylene bisphosphonate (1,5-methylene bisphosphonate tetrakisphosphate). Figure 39 shows in red the phosphate groups which are present at all positions besides 1 and 5. The methylene bisphosphonate protective groups are depicted in blue and substitute the naturally occurring pyrophosphate groups. The denaturation curves of GST-Asp1$^{365\text{-}920}$ with 4 Fe/monomer in the presence of 10-fold excess of 1,5-(PCP)$_2$-IP$_4$ appear strikingly similar to the ones without any substrate (Figure 41, Figure 35). The only slight difference that can be observed is an even more prominent flattening of the curves from 205 nm to 225 nm making the minima at 209 nm and 222 nm unremarkable.

Furthermore, the analysis of the temperature interval measurement yielded a saturation curve with a low signal to noise ratio as was the case for the GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$ (0 Fe/monomer, Figure 38) sample. In the same manner, the melting temperature at 49.5 °C is not a reliable result as full saturation is not achieved. The single spectra remain the only data set of GST-Asp1$^{365\text{-}920}$ in the presence of the substrate and in conclusion it does not indicate any significant changes in its secondary structure.

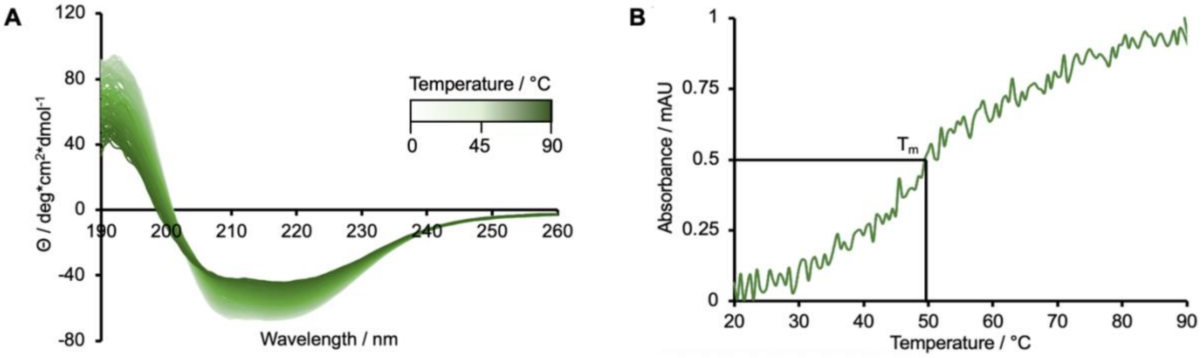


Figure 41: Denaturation curves for GST-Asp1$^{365\text{-}920}$ 4 Fe/monomer with 10-fold molar excess of 1,5-(PCP)$_2$-IP$_4$ (A) and normalized absorbance for the determination of the melting temperature (B). A) Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm. B) The normalized absorbance was plotted as a function of temperature resulting in saturation curves. The melting temperature (T_m) is the corresponding value on the x-axis of the half-maximal absorbance (turning point).

CD spectroscopy was used to characterize the His-tagged versions of the pyrophosphatase domain as well. Due to the small size of the affinity tag, these constructs offer an insightful overview of the secondary structure elements without the interference of some attributed to the GST fusion protein. The two states of the His-tagged constructs that were studied are described as apo and holo, since no ICP-MS samples were prepared and thus the exact number of iron ions per monomer cannot be specified.

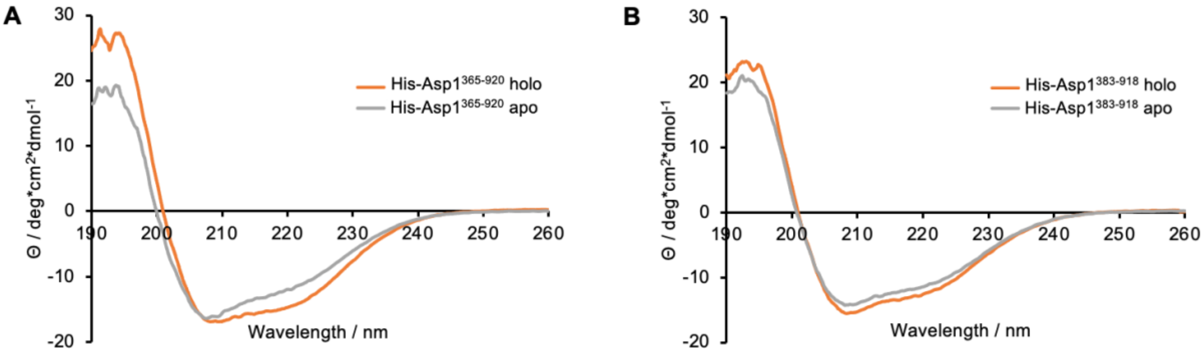


Figure 42: Comparison of single CD spectra of His-Asp1$^{365-920}$ and His-Asp1$^{383-918}$ in the apo and holo state. A) Single CD spectra of apo and holo His-Asp1$^{365-920}$. B) Single CD spectra of apo and holo His-Asp1$^{383-918}$. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm.

Figure 42.A shows the CD spectra of His-Asp1$^{365-920}$ in the apo and holo state, this protein variant exhibiting the same spectral features mentioned before: a minimum at 209 nm and a more implied shoulder at 222 nm. Among these two samples there is only a small noticeable deviation of the signal in the area 210 to 230 nm (Figure 42). Moreover, an increase of approx. 5% in the α-helix and a simultaneous decrease of approx. 4% in the β-sheet amounts occur based on the increase of iron in the overall sample (shift from apo to holo, Supplementary table 7). Even if the differences are low, the same trend was discernible in case of the GST-tagged samples of Asp1$^{365-920}$, thus confirming this to be a cluster-induced effect. The CD spectra of the apo and holo His-Asp1$^{383-918}$ samples are almost identical and exhibit the same two minima at 209 and 222 nm (Figure 42.B). The same trend regarding the α-helix and β-sheet percentages could be observed as well with the differences amounting to maximum 5% between the apo and holo states (Figure 42.B).

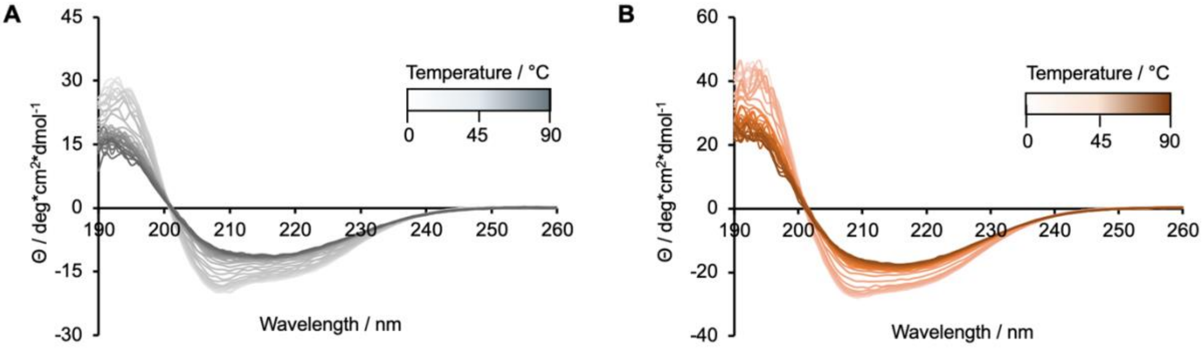


Figure 43: Denaturation curves of His-Asp1$^{383-918}$ in apo (A) and holo (B) states. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm. A) His-Asp1$^{383-918}$ apo. B) GST-Asp1$^{383-918}$ holo. The intensity of the colours correlate with increasing temperatures at which the data sets were recorded.

Figure 43 depicts the denaturation curves measured from apo and holo His-Asp1$^{383-918}$ in the near-UV area, where a flattening of the curves can be observed with increasing temperature. The α-helix with its specific signal at 222 nm exhibits a minimum which disappears until reaching 90 °C and full denaturation of the protein. Similar, the minimum at 209 nm is shifted to 214 nm during the temperature interval measurement, this enabling the use of the normalized absorbance at 209 nm to determine the melting temperature. The two sets of denaturation curves appear similar despite an increase in the overall signal when the cofactor is present. Furthermore, the amount of α-helical structure increases with bound Fe-S cluster by the same percentage that the β-sheet amount decreases (supplementary table 8).

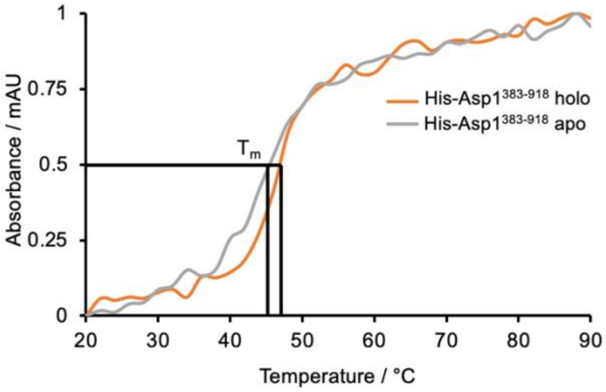


Figure 44: Normalized absorbance at 209 nm for the determination of the melting temperature of His-Asp1$^{383-918}$ in different cluster bound states. The normalized absorbance was plotted as a function of temperature resulting in saturation curves. The melting temperature ™ is the corresponding value on the x-axis of the half-maximal absorbance (turning point).

As per Figure 44 it can be observed that the calculated melting temperatures of apo and holo His-Asp1$^{383-918}$ are similar with 46 °C and 48 °C and in good agreement with the ones measured for the GST-tagged versions. The difference between the apo and holo state of GST-Asp1$^{365-920}$ amounted to approx. 3 °C which depicts the greatest measured difference between any two states of the Asp1 pyrophosphatase domain.

In the next part of the CD-spectroscopy-based characterization of the Asp1 pyrophosphatase domain the GST-Asp1$^{365-920}$ variant was digested using PreScission protease to remove the GST-tag. This experimental work-up provided a high-concentrated sample of the pyrophosphatase domain without any affinity tag which was then measured under the same conditions as the previously analyzed samples. This way a comparison between the tagged and untagged version was possible and the difference in the secondary structure made out by the large fusion protein could be quantified. Figure 45 shows a comparison between the tagged and untagged spectrum of Asp1$^{365-920}$ with the only perceptible difference being an increase in the signal (15 deg*cm^2*dmol^{-1}) of GST-Asp1$^{365-920}$ in the 190-200 nm area. The presence of the affinity tag leads to an increase of 10% and 4% in the amount of α-helical structure and parallel β-sheets as well as a decrease in the overall percentage of antiparallel β-sheets by 8% (Supplementary table 9). This occurs due to the GST-tag being a protein with secondary structure elements itself.

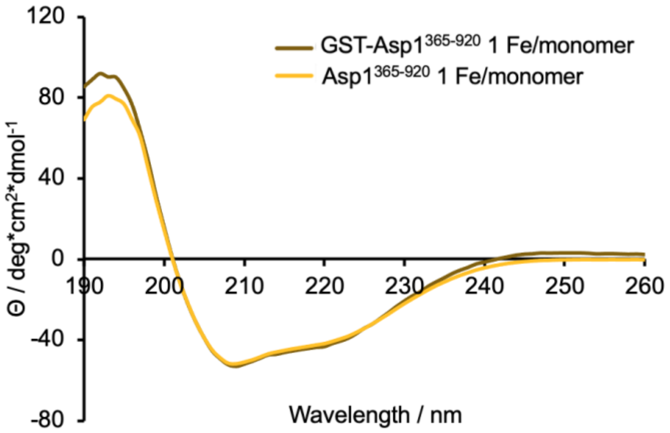


Figure 45: Comparison between CD spectra of the GST-tagged and cleaved Asp1$^{365-920}$ pyrophosphatase domain. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm. The CD spectrum was GST-Asp1$^{365-920}$ was recorded at 1.7 µM whereas the one of the untagged Asp1$^{365-920}$ was measured at 5 µM.

The denaturation curves of the untagged Asp1 pyrophosphatase domain appears the most stretched among all temperature interval measurements (Figure 46). The minimum at 209 nm is very prominent in comparison to the second one at 222 nm, these features amounting to a total of 19.8% of α-helical structure (supplementary table 9).

The amount of the antiparallel β-sheets is balanced with 22.2% whereas the remaining structural elements are contained in similar amounts as for the GST-tagged protein version (Supplementary table 9). Interestingly, the melting temperature of the untagged variant is identical with the one of the apo and as-isolated GST-Asp1$^{365-920}$. It appears that the presence of the GST-tag does not influence the melting temperature despite its considerable size of 26 kDa.

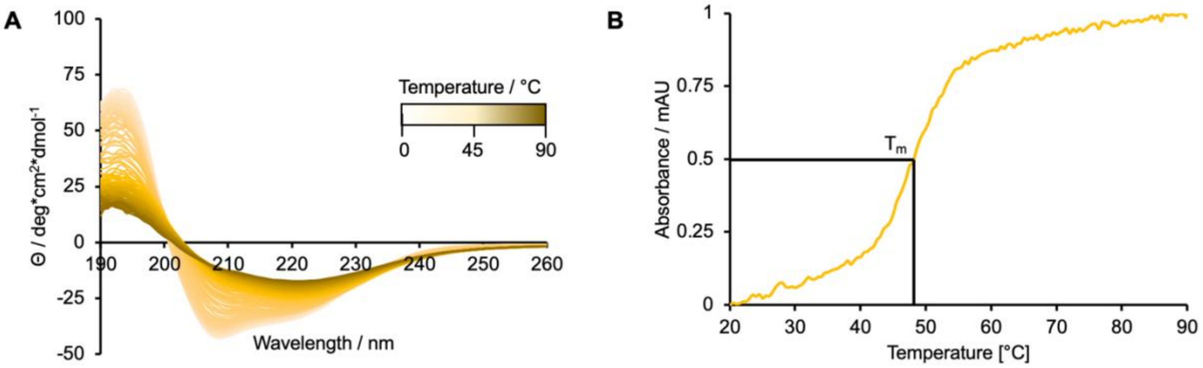


Figure 46: Denaturation curves for Asp1$^{365-920}$ with 1 Fe/monomer (A) and normalized absorbance for the determination of the melting temperature (B). A) Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.4. The data sets were converted to mean residues molar ellipticity values in deg*cm^2*dmol^{-1} (θ) which were plotted as a function of the wavelength in nm. B) The normalized absorbance was plotted as a function of temperature resulting in saturation curves. The melting temperature (T_m) is the corresponding value on the x-axis of the half-maximal absorbance (turning point).

4.9 Characterization of reconstituted wildtype Asp1$^{365-920}$ by Mössbauer spectroscopy

Mössbauer spectroscopy is a technique that enables the detection and characterization in distinct metal species in a sample and is mostly applied to studies of the ^{57}Fe isotope. It relies on the Mössbauer effect which consists in the recoilless resonance absorption and emission of nuclear γ-irradiation.[164] The γ-photon used for ^{57}Fe Mössbauer spectroscopy is obtained from the radioactive decay of the ^{57}Co isotope, yielding a metastable ^{57}Fe that decays through a series of transitions. Provided information usually involves the Fe oxidation number and spin state, nuclearity of a Fe-S cluster with at least one metal ion and delocalization properties in mixed-valent clusters.[164] The ^{57}Fe isotope has a low natural abundance (2%) and the other two iron isotopes do not have a Mössbauer effect. In order to obtain a good signal to noise ration the samples either have to be enriched in vivo or reconstituted in vitro with $^{57}FeCl_3$.

To obtain an isotope-labeled sample aerobically purified GST-Asp1$^{365-920}$ from *E. coli* BL21(DE3) *ΔiscR* cells with a concentration of 17 mg/mL was chemically reconstituted with 5-fold molar excess of $^{57}FeCl_3$ under anaerobic conditions. The sample underwent desalting via SEC to remove iron-sulfur aggregates which can superimpose the signals from the iron species of the Asp1 Fe-S cluster. This experiment also confirmed that the prepared GST-Asp1$^{365-920}$ sample is dimeric dispersed (Figure 100). An electronic spectrum recorded prior to the Mössbauer measurement showed a full occupancy of the presumed [2Fe-2S] cluster with a molar extinction coefficient of 11.59 mM^{-1}*cm^{-1}at 410 nm. The value, which slightly higher than values reported in literature (1-10 mM^{-1}*cm^{-1}at 410 nm), may be attributed to a slight peak at 600 nm indicating some adventitiously bound iron which does not undergo ligation by any amino acid residues of the protein (Figure 47).

To investigate this, a sample of the *S. pombe* pyrophosphatase domain reconstituted with 5-fold excess of iron and sulfur sources was digested with the suitable protease and analyzed via EAS. The electronic absorption spectrum of the tagged and untagged Asp1$^{365-920}$ variant are visible in Figure 47. The already known spectroscopic features are noticeable in a more prominent form when the affinity tag is removed. The two characteristic bands at 410 nm and 470 nm are present and the shoulder at 320 nm becomes a very well-defined peak. The molar extinction coefficient values are at 12 and 8 mM^{-1}*cm^{-1}, this being in good agreement with the Asp1-related EAS data for a fully-occupied 2Fe-2S cluster (Figure 47).[165,166] The significant difference through the removal of the tag is the decrease of the overall background signal that ultimately leads to more defined and prominent spectral properties.

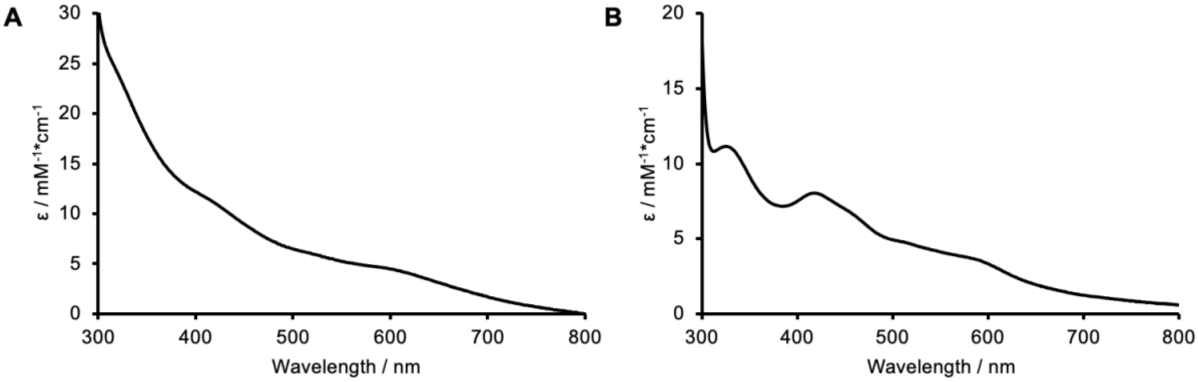


Figure 47: EA spectrum of GST-Asp1$^{365\text{-}920}$ (A) and digested Asp1$^{365\text{-}920}$ (B) after chemical reconstitution with 5-fold excess of iron (^{57}Fe) and sulfur sources. The molar extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). A) The measured sample was reconstituted GST-Asp1$^{365\text{-}920}$ which was used for Mössbauer measurements. B) The measured sample was reconstituted Asp1$^{365\text{-}920}$ which was used for Mössbauer measurements. The reconstituted sample was digested after completing the first Mössbauer measurement with the affinity tag.

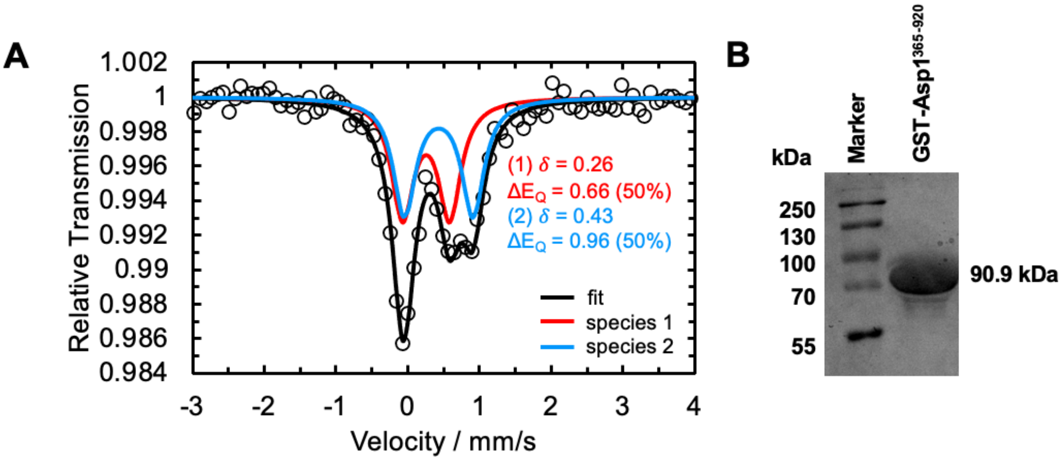


Figure 48: Mössbauer spectrum and SDS-PAGE of the chemically reconstituted GST-Asp1$^{365\text{-}920}$. A) The relative transmission was plotted as a function of velocity (mm/s) with the data points illustrated as white circles and the fit as a black line. Depicted in red and blue are the two components of the Asp1 Fe-S cluster which are attributed to two different iron species. B) SDS-PAGE showing the purity of the measured Mössbauer sample.

A suitable fit for the experimental data set was achieved with two components each with a relative intensity of 50%, exhibiting two quadrupole doublets with low isomer shifts (δ = 0.26 and δ = 43 mm/s, Figure 48) and small quadrupole splittings (ΔE$_Q$ = 0.66 and (ΔE$_Q$ = 0.96 mm/s, Figure 48). The recorded data is in good agreement with the Mössbauer data from a previous Asp1 study.[3] Additionally, it appears that the GST-Asp1$^{365\text{-}920}$ sample measured in this study is of higher purity. Three additional components with large quadrupole splittings (up until 4 mm/s) noticeable in the earlier Mössbauer spectrum and which attributed to residual Fe-S aggregates are not visible in this spectrum (Figure 48), this highlighting the impact of upgrading the desalting procedure from disposable desalting columns to SEC columns and the increase of sample purity.[3]

The isomer shifts and quadrupole splittings are characteristic of Fe-S clusters, however, do not coincide with the iron species from strictly cysteine ligated [2Fe-2S] clusters. The parameters appear slightly larger than those recorded for ferric ions ($\delta = 0.27$ mm/s, $\Delta E_Q = 0.60$ mm/s) but not as large as the ones observed for ferrous ions ($\delta = 0.60$ mm/s, $\Delta E_Q = 2.70$ mm/s) for which large quadrupole splittings are typical.[164,167] Remarkably, the isomer shifts and quadrupole splittings of the His-ligated clusters appear to fit better to the Mössbauer spectrum of Asp1. Such example is the Cis-ligated [2Fe-2S] cluster of IscU with an isomer shift of $\delta = 0.27$ mm/s and a quadrupole splitting of $\Delta E_Q = 0.66$ mm/s for the Cys-ligated Fe^{3+} ion as well as an isomer shift of $\delta = 0.32$ mm/s and a quadrupole splitting of $\Delta E_Q = 0.94$ mm/s for the His-ligated Fe^{3+} species.[168] These values align very well with the experimental values recorded for GST-Asp1$^{365\text{-}920}$ except for the isomer shift of component 2 which is slightly larger than in literature. Moreover the catalytically relevant histidine residue at position 397 appears to be in close proximity to cysteine residues C864 and C879 which play a role in the coordination of the Asp1 Fe-S cluster. This indicates that a histidine could certainly be ligating the Fe-S cluster. Recently coordination pattern where four cysteines and one histidine have been reported for glutaredoxins, which are highly potent redox active proteins. However, EXAFS data from the previous Asp1 study show that no nitrogen-containing amino acid could be measured in the first two shells surrounding the Fe-S cluster, this making the Mössbauer signal presumably attributed to a histidine ligation more intricate.

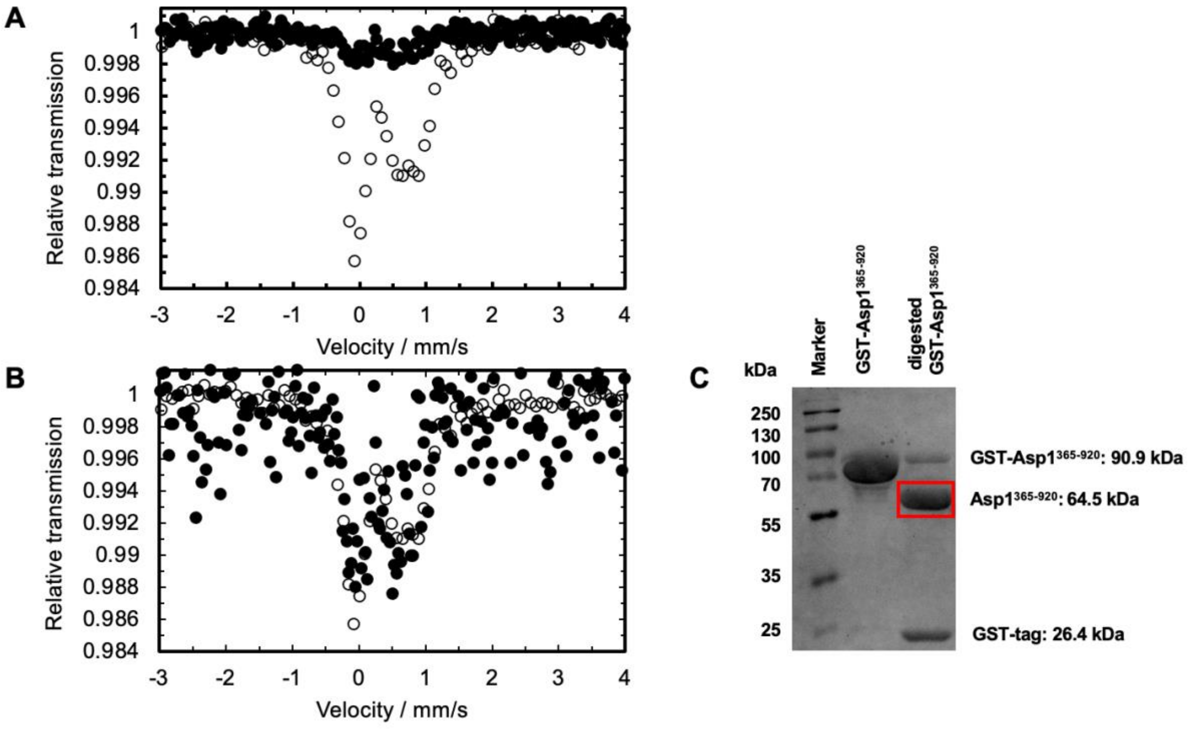


Figure 49: Mössbauer spectrum and SDS-PAGE of the chemically reconstituted Asp1$^{365\text{-}920}$. The relative transmission was plotted as a function of velocity (mm/s) with the data points illustrated as white circles and the fit as a black line. Depicted in red and blue are the two components of the Asp1 Fe-S cluster which are attributed to two different iron species. A) Overlayed Mössbauer spectra of tagged and cleaved versions of Asp1$^{365\text{-}920}$ reconstituted with 5-fold excess of ^{57}Fe. B) Close-up version of the two Mössbauer spectra. C: SDS-PAGE showing the purity of the measured Mössbauer sample.

The reconstituted sample of GST-Asp1$^{365\text{-}920}$ measured with Mössbauer spectroscopy was digested using PreScission protease and prepared for a new measurement with the goal of obtaining a data set of the pyrophosphatase domain without an affinity tag (Figure 49.C). Due to the preparative work-up the sample concentration amounted to only 80 µM. Figure 49.A depicts the results of the second Mössbauer measurement, where the newly-recorded spectrum is illustrated with black dots and the previous one of the tagged pyrophosphatase with white dots. No fit could be applied on this data set as the signal amplitude was not high enough due to a too low sample concentration. Figure 49.B shows a scaled-up version of both spectra for the purpose of comparison. The spectrum of the digested Asp1 pyrophosphatase domain appears similar to the GST-tagged one, however, the two cannot be determined identical with certainty due to insufficient concentration.

4.10 Spectroscopic characterization of Asp1$^{365-920}$ mutant variants

In this section the electronic properties of different GST-Asp1 constructs were analysed using EAS to gain insights into which amino acid residues play a role in the coordination of the Fe-S cluster. The electronic spectra show the anaerobic isolated GST-Asp1$^{365-920}$ and three mutant variants exhibit molar extinction coefficient (ε) values at 410 nm up to 5 mM^{-1}*cm^{-1}, values that are slightly lower than the ones found in literature.[165] The curve of the wildtype GST-Asp1$^{365-920}$ shows one peak at 410 nm with a shoulder at 460 nm which are characteristic for [2Fe-2S] proteins (Figure 50.A). However, the width of the peak is not typical for this type of cluster geometry and based on this criterion it could appear like the typical single broad peak displayed by proteins binding [4Fe-4S] clusters. Furthermore, the shoulder at 320 nm represents a spectroscopic feature often found in electronic absorbance spectroscopic data of well-characterized 2Fe-2S proteins (Figure 50.A). The AlphaFold2 shows that histidine 397 is located close to many cysteines which are thought to be potential binding sites for the Fe-S cluster. Biological data showed that a mutation of this amino acid leads to a reduced pyrophosphatase activity. The H397A mutant appears to have the same spectroscopic bands however at half the extinction coefficient observed for the wildtype. The C663S mutant exhibits only the peak at 410 nm whereas the C607S mutant shows the lowest signal among all analyzed variants and no distinct spectroscopic features (Figure 50.A).

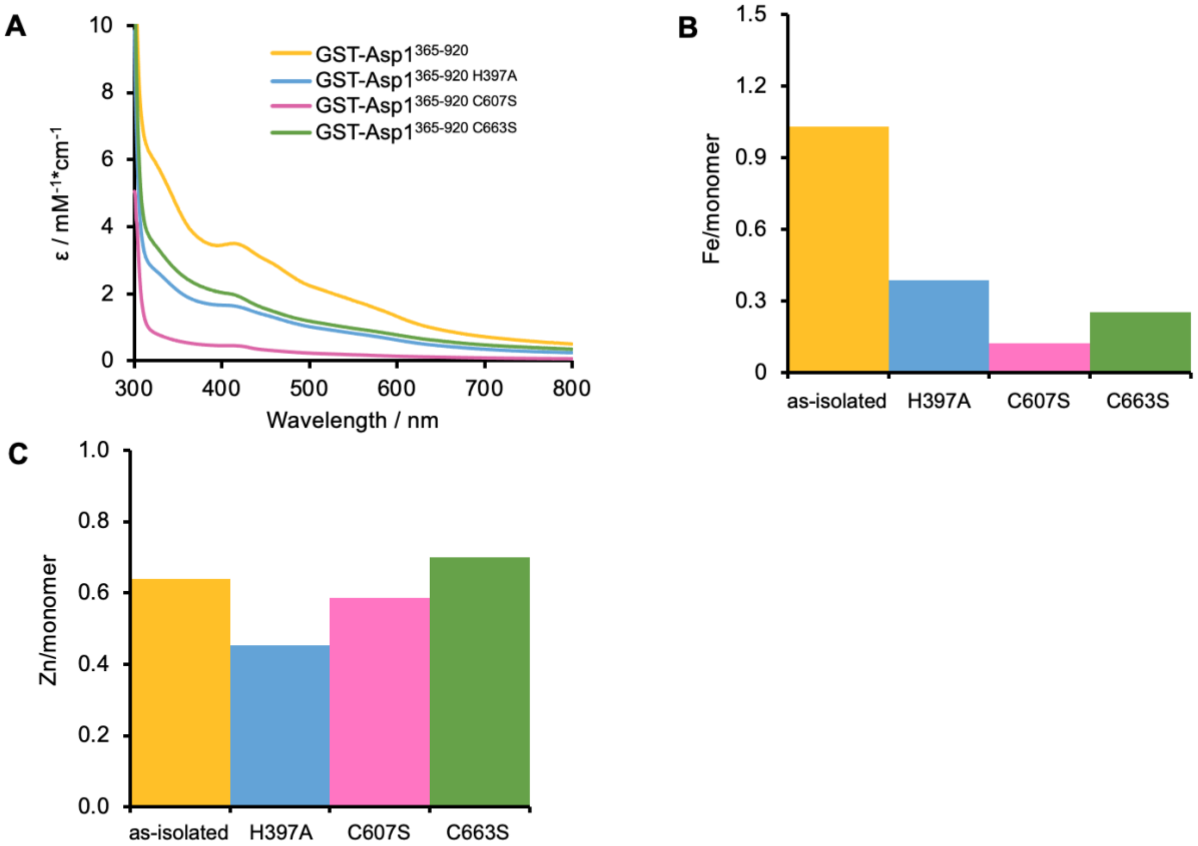


Figure 50: Overview of EA spectra and iron content of GST-Asp1$^{365-920}$ wildtype and single mutant variants from expression without supplemental iron and sulfur. A) The molar extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). B) The measured iron levels of the different samples are shown in corresponding colours. C) The measured zinc levels of the different samples are shown in corresponding colours. The measurements consisting of biological triplicates have additional error bars depicting the standard deviation.

The ICP-MS data shows that GST-Asp1$^{365\text{-}920}$ binds only 1 Fe atom per monomer in comparison to the three mutant variants (GST-Asp1$^{365\text{-}920\ \mathrm{H397A,}}$ GST-Asp1$^{365\text{-}920\ \mathrm{C607S}}$ GST-Asp1$^{365\text{-}920\ \mathrm{C663S}}$) that exhibit much lower values at approx. 0.3 Fe/monomer which indicates almost no iron is present. Interestingly, it appears that the applied expression protocol using *E. coli* strains specialized for upregulated Fe-S proteins production and the anaerobic isolation led to iron values which are not in agreement with previous studies. Possible rationales are that either half of the protein amount is fully-loaded with cofactor (100% holo) or all of the protein is only partially loaded with cofactor (50% holo), both theories amounting to 1 Fe/monomer overall as observed in Figure 50.B. Nonetheless full maturation of all Asp1 pyrophosphatase domain was not achieved in the absence of supplementary sources of iron and L-cysteine.

The advantage of ICP-MS is that not only one element is analyzed during a measurement but almost all biologically-relevant elements. It proved very interesting to analyze the zinc content per protein monomer as well as zinc is a bivalent ion and can compete with an iron atom for a biding site of a cluster. Remarkably, one zinc per monomer of Asp1$^{365\text{-}920}$ could be detected in the wildtype sample (Figure 50.B). Similar values were measured in the mutant variants H397A, C607S and C663S as well. Moreover, the reason why the ICP-MS values do not amount to 1 is the 20% standard deviation resulting from the protein concentration determination using the Bradford assay, that needs to be considered when analyzing this kind of data.

The expression of the pyrophosphatase domain was since then conducted using 2 mM ferric ammonium citrate before induction and 2 mM L-cysteine after induction with IPTG. Usually, it could be observed that cultures turned darker due to the presence of the chemicals in the culture media. Wildtype pyrophosphatase domain as well as mutant variants H397A, C607S, C643S, C663S and the quadruple mutant (QM) were expressed, anaerobically isolated and later analyzed via EAS and ICP-MS. As mentioned earlier, the GST-Asp1$^{383\text{-}918\ \mathrm{C643S}}$ mutant was added to the overview based on a sequence alignment that indicated that the cysteine 643 is the only conserved residue among all species of the Asp1 pyrophosphatase domain. To facilitate crystallization, the mutation was introduced on the shorter version of the *asp1*$^{+}$ gene. All experiments were conducted as biological triplicates to ensure reproducibility. The EAS and ICP-MS data sets of the chemically reconstituted pyrophosphatase domain were added as comparison. The reconstitution experiment was performed according to literature and as mentioned in section 3.9 using DTT as reductant and 5-fold molar excess of ferric chloride and disodium sulfide as iron and sulfur sources. This ratio proved to be favourable for the reconstitution of other [2Fe-2S] proteins and yielded a fully loaded cluster per monomer protein.[3,166]

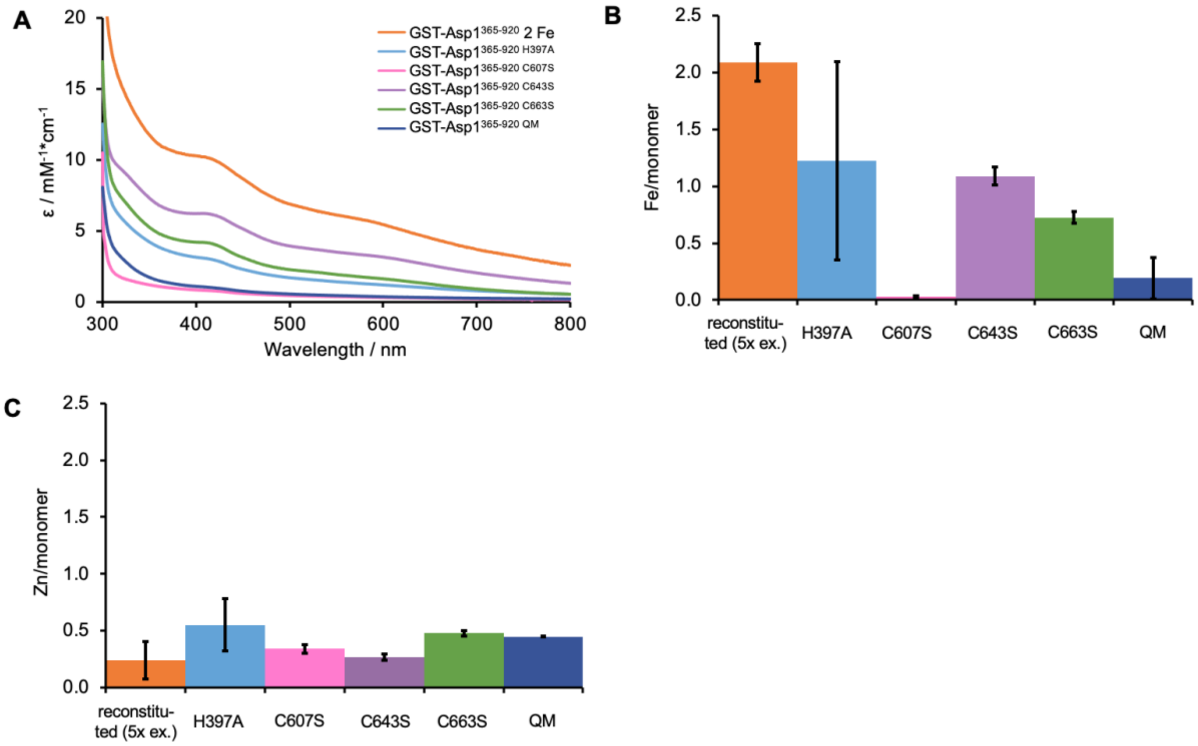


Figure 51: Overview of EA spectra and iron content of GST-Asp1$^{365\text{-}920}$ wildtype and single mutant variants from expression with supplemental iron and sulfur. A) The molar extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). B) The measured iron levels of the different samples are shown in corresponding colours. C) The measured zinc levels of the different samples are shown in corresponding colours. The measurements consisting of biological triplicates have additional error bars depicting the standard deviation.

The electronic absorption data of analyzed constructs with mutant variants exhibit molar extinction coefficients at 410 nm between 1-7 mM^{-1}*cm^{-1}. The characteristic spectroscopic elements mentioned based on Figure 50 and in Figure 51 like the two peaks in the 410 -460 nm region and the shoulder at 320 nm are present in the same samples comprised in Figure 51 and appear more visible due to the higher signal intensity. On the other hand, the value of the reconstituted wildtype pyrophosphatase domain is approx. double and amounts to 12 mM^{-1}*cm^{-1}. The electronic absorption data is well-complemented by the ICP-MS data summarized in Figure 51.B. For wildtype Asp1$^{365\text{-}920}$ a fully-loaded 2Fe-2S cluster was identified. Mutant variants H397A and C643S show high amounts of iron at approx. 1 Fe/monomer whereas the C607S construct and the quadruple mutant exhibit only residual to no iron traces.

In case for GST-Asp1$^{365\text{-}920\ C663S}$ 0.75 Fe/monomer were measured, which likely indicates the presence of one iron atom per monomer protein as well due to the calculated measurement error. ICP-MS values not completely overlapping with metal atom stoichiometry (0, 1, 2, 3 or 4 metal atoms per cluster) highlight that there might be adventitiously bound Fe on the surface of the protein that does not contribute to the cluster geometry. In comparison, in Figure 51.B shows that only residual amounts of zinc (values under 0.5) are present in of all the samples this confirming that the supplemental iron and sulfur sources are sufficient for the iron substitution by zinc to be suppressed during gene expression.

Optimization of expression in the presence of supplemental iron and sulfur was attempted by prolonging the induction time from 3 h to 24 h. This adjustment was shown beneficial in the maturation of various metalloproteins.[166] Even though expression tests indicated the yields appear to be significantly lower during overnight expression, this step might represent an improvement if full occupancy of the [2Fe-2S]-cluster is achieved. The significant temperature decrease might shift the focus *of E. coli* metabolism on cell growth and division to overexpression and correct folding of the target metalloprotein. Figure 52 comprises the electronic absorption spectrum and ICP-MS data for overnight expressed GST-Asp1$^{365\text{-}920}$, which show the presence of less cofactor than after 3 h induction. The molar extinction coefficient reaches values of 1 mM^{-1}*cm^{-1} with the ICP-MS data showing only residual trace amounts of iron (0.33 Fe/monomer, Figure 52). This data indicates that neither full maturation nor coordination of one iron ion per monomer protein is achieved under such prolonged expression conditions.

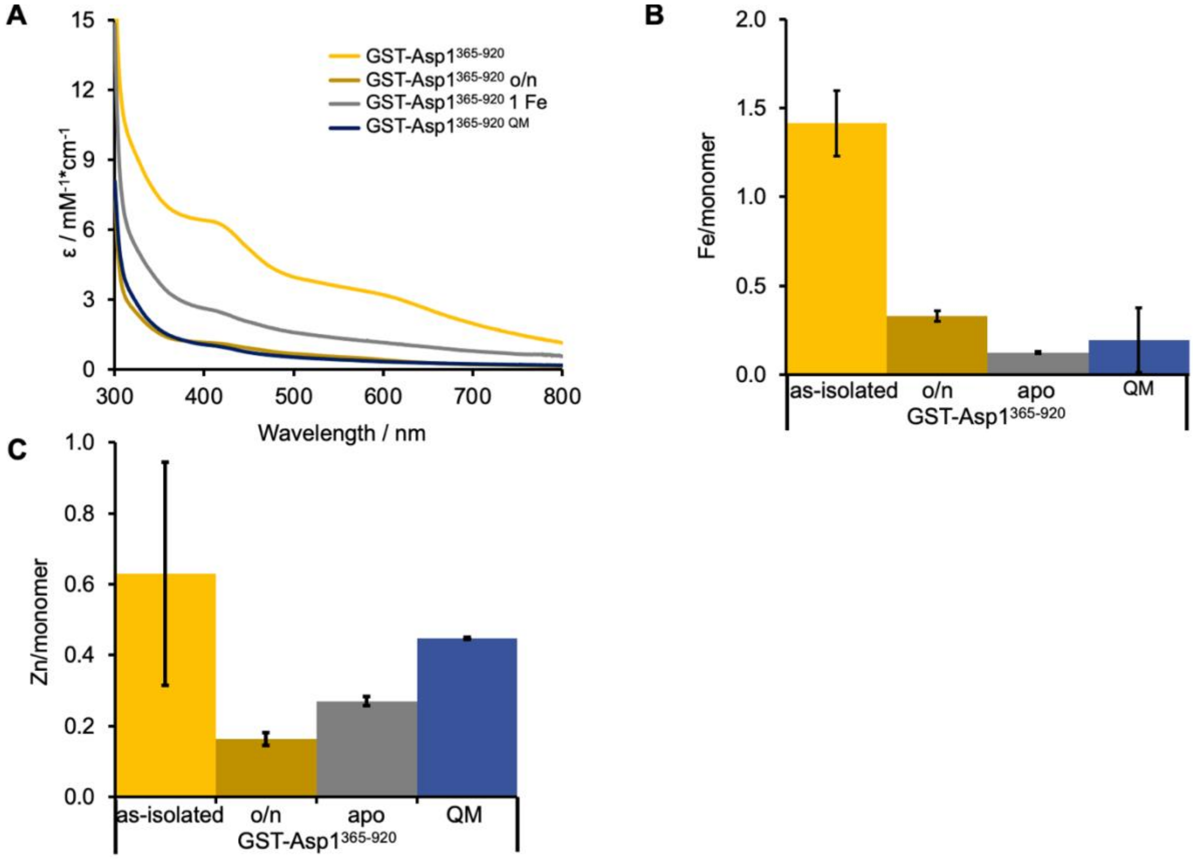


Figure 52: Comparison of EA spectra and iron content of different GST-Asp1$^{365\text{-}920}$ variants (as-isolated, as-isolated o/n, apo, QM). A) The molar extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). B) The measured iron levels of the different samples are shown in corresponding colours. C) The measured zinc levels of the different samples are shown in corresponding colours. The measurements consisting of biological triplicates have additional error bars depicting the standard deviation.

The presented values for the wildtype as-isolated and apo pyrophosphatase domain and the quadruple mutant were additionally compared in Figure 52 to determine if the named mutant can be used as a negative control in this study. Figure 52.A shows that the spectra of the two protein variants without cluster are very similar, the only difference being the values of the molar extinction coefficient which are twice as high for the apo variant than for the quadruple mutant.

However, none of the spectral signals observed in the as-isolated state are visible in the apo variants. Furthermore, the ICP-MS data in Figure 52.B clearly show no amount of iron present in the samples, this enabling a verification of the cluster absence and the quadruple mutant as negative control. The ICP-MS data regarding the zinc content present in this sample show the same tendency as previously mentioned, that zinc (0.6 Zn/monomer) prevents the coordination of the second iron atom in the [2Fe-2S] cluster form and thus full maturation (Figure 52). In addition, there is no zinc present in the apo and quadruple mutant variant which deem the protocols used in the preparation of these two samples suitable to achieve variants of the Asp1 pyrophosphatase domain that do not bind an Fe-S cluster (Zn/monomer protein < 0.5, Figure 52). These initial experiments showed that the iron amounts naturally present in cells and culture media are not enough to facilitate full maturation of the overexpressed *S. pombe* Asp1 [2Fe-2S] cluster. This fact also paired with insights gained from the previous study led to the addition of supplemental iron and L-cysteine to the culture media in the same manner as described in section 3.5.2.

In the early process of designing new mutant variants of the pyrophosphatase domain the cysteine residue at position 868 and the cysteine pair located as positions 643 and 644 were considered for site-directed mutagenesis. These two constructs were cloned in a manner similar to the C643S single mutant variant as mentioned in section 3.3.1. The chosen template was the shorter construct of the pyrophosphatase domain (GST-Asp1$^{383\text{-}918}$) and after validation via sequencing the transformation with *E. coli* BL21(DE3) *ΔiscR* cells was conducted. Expression of GST-Asp1$^{383\text{-}918\ \mathrm{C868S}}$ and GST-Asp1$^{383\text{-}918\ \mathrm{C643S\ C644S}}$ was conducted according to the optimized protocol (section 3.5.2) with 3 h induction time as well as additional sources of iron and sulfur. Anaerobic isolation was performed several times and according to protocol, however both constructs could not be isolated in the soluble fraction using these expression and purification conditions. SDS-PAGE analysis of samples collected during the failed isolations showed the GST-Asp1$^{383\text{-}918\ \mathrm{C868S}}$ and GST-Asp1$^{383\text{-}918\ \mathrm{C643S\ C644S}}$ are expressed nearly completely insoluble based on the large and intense band present at approx. 90 kDa in the pellet fraction (Figure 53). At the same time, the presence of many bands in the pellet fraction indicates that the cell disruption was conducted successfully as proteins of all sizes could be detected (Figure 53). Consequently, the cell disruption step during protein purification is not performed insufficiently and does not constitute the cause for not obtaining soluble protein.

There are multiple reasons that might lead to mutant protein variants that are not expressed soluble, one being incorrect folding during expression due to suboptimal growth parameters. Even though the expression protocol has been optimized in detail the loss-of-function mutation of one cysteine to a serine could have significant consequences on the properties of the protein. Cysteine residues are not only known to bind ancient cofactors as Fe-S clusters but also act as structural elements and undergo disulfide bonds that increase rigidity of certain structural areas of a protein. If such a mutated residue is structurally relevant its substitution could lead to incorrect folding during gene overexpression and subsequent precipitation in the pellet fraction. Several strategies were tested to increase the solubility of these two variants of the pyrophosphatase domain, the first being a change in the induction temperature from 30 °C to 18 °C. The reduction in temperature could be beneficial as it triggers the co-expression of chaperones, which are known to facilitate folding processes. If expression at such a low temperature does not prove favourable in any way for the target protein, the co-expression of chaperones could be initiated by incubation of the cultures at 4 °C for 1 h prior to induction and subsequent induction at a higher temperature (30 °C) so that the *E. coli* metabolism is not as affected. However, these changes in the growth parameters did not yield any improvements and the two mutant variants of the pyrophosphatase (GST-Asp1$^{\text{383-918 C868S}}$, GST-Asp1$^{\text{383-918 C643S C644S}}$) domain were still expressed predominantly insoluble. In the next step the optimization of the purification protocol was attempted. The addition of glycerol as a stabilizer was proven successful in establishing a first isolation protocol for the wildtype pyrophosphatase domain, so this parameter remained unchanged during the following purification tests. The different purification approaches summarized in Supplementary table 10, did not lead to any improvements in the yields of the two mutant variants. Despite increasing the glycerol concentration up to 20%, testing two versatile buffer systems and adding urea to buffers to diminish the hydrophobic effect and stabilize non-polar side chains of the protein in aqueous solution no positive effect could be either observed or measured (Supplementary table 10).

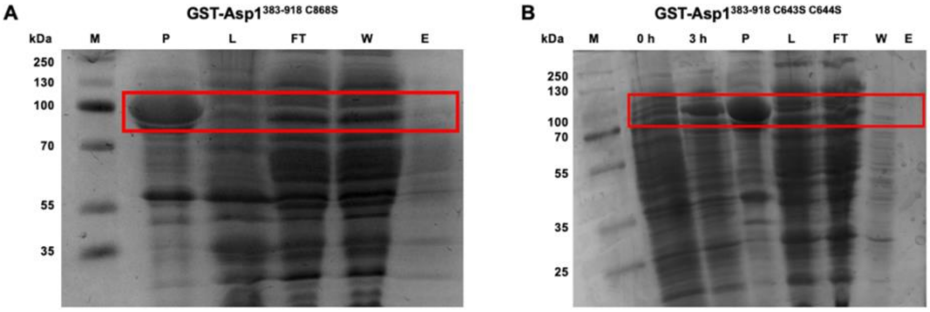


Figure 53: SDS-PAGE showing different fractions from the anaerobic isolations of GST-Asp1$^{\text{383-918 C868S}}$ (A) and GST-Asp1$^{\text{383-918 C643S C644S}}$ (B). A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, W = wash, E = elution.

During one purification using the buffer composition described as condition (Supplementary table 10) a very small amount of the GST-Asp1$^{\text{383-918 C868S}}$ was isolated with a concentration of approx. 8 µM. The concentration could be not increased any further during the experimental work-up. The sample was analyzed using EAS and the recorded data does not show any of the characteristic signals at 320 nm and in the 410-460 nm region. The concentration however is too low to obtain a certain signal to noise ration and to be able to detect any spectral features. The purification using condition was reproduced several times but did not yield any more of the C868S mutant variant. None of the remaining purification conditions improved the yield and sample quality of either GST-Asp1$^{383\text{-}918\ C868S}$ or GST-Asp1$^{383\text{-}918\ C643S\ C644S}$. It could still be possible that the tested conditions are not the right ones to produce the two mutant protein variants in a soluble state. Furthermore, based on the EAS data collected for GST-Asp1$^{365\text{-}920\ C643S}$, it can be observed that this single residue is not involved in the coordination of the Fe-S cluster since iron per monomer is still detectable after mutating cysteine C643. These findings paired with earlier mentioned EAS and ICP-MS data show the cysteines 868, and 644 are most probably structurally relevant residues and do not contribute to the coordination of the Fe-S cluster. Cysteine 643 does not appear to influence the binding of a cofactor and or to play a role in maintaining the protein structure since the expression of its corresponding single mutant does not lead to insoluble protein.

4.11 Crystallization of Asp1$^{365-920}$ and Asp1$^{383-918}$ variants

Sparse matrix screens were prepared to screen for optimal crystallization conditions for the Asp1 pyrophosphatase domain. Different versions of the pyrophosphatase domain were isolated anaerobically to facilitate the retention of the natural cluster architecture. The protein constructs used in crystallization are the GST-tagged wildtype versions of the pyrophosphatase domain Asp1$^{365-920}$ and Asp1$^{383-918}$. Crystallization screens were conducted using tagged protein variants despite this being an inconvenience in the crystallization process. The affinity tags were kept in some of the screens because they facilitate protein stability. Another strategy was the use of the same protein variants with the small His-tag to avoid any inconveniences that might be caused by a too large affinity tag and to preferably obtain a crystal structure of the pyrophosphatase domain without any ligands or fusion proteins. Moreover, the shorter versions of the wildtype pyrophosphatase domain were designed especially to promote the crystallization process by the absence of any unstructured regions in the protein sequence.

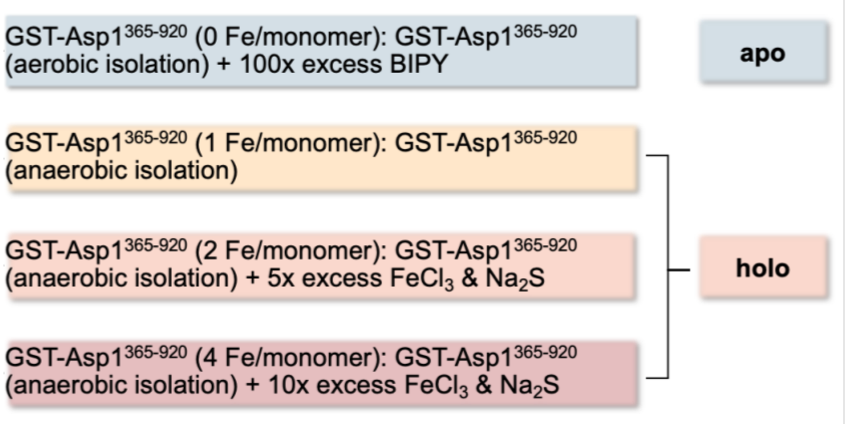


Figure 54: Overview of the different procedures to achieve the four known states of GST-Asp1$^{365-920}$ for crystallization purposes. The same workflow was applied if His-tagged constructs were to be used for crystallization. The chosen colours further indicate the maturation state of the Fe-S cluster.

The followed work sequence is illustrated in Figure 54 on the example of GST-Asp1. Depending on which construct, and which state were needed for crystallization different methods were used to achieve the target combination of parameters. Depending on which construct was needed for crystallization. If none of these strategies were fruitful, the co-crystallization of Asp1$^{365-920}$ in complex with one of the substrates (1,5-IP_8 or 1-IP_7) could be attempted. Chemically synthesized and protected versions of the two substrates (1,5-$(PCP)_2$-IP_4 and 5-$(PCP)_2$-IP_3) were kindly provided by the Fliedler group. Usually the co-crystallization of proteins in complex with the substrate can yield great results as the presence of the substrate forces the protein to assume its correct structure thus facilitating the overall crystallization process that might not occur for the protein alone. Finally, the used crystallization screening solutions represent another parameter that can be adapted throughout the process. The JBScreen Classic solutions (1-8) were used routinely in the preparation of crystal plates (Supplementary table 11). Furthermore, the sample buffer was changed to check if any buffer system would prove especially favourable in the crystallization process.

All the above-mentioned strategies were employed for the goal of crystallizing the Asp1 pyrophosphatase domain yet none of them yielded any crystals. Two major buffer systems were used in the crystallization attempts these being the Tris and sodium phosphate buffer systems. Crystallization in the presence and absence of glycerol was tested as well, as this compound can sometimes interfere in the process of obtaining diffracting crystals. Furthermore, the tested concentrations ranged up to 30 mg/mL of pure protein as Asp1 was not stable enough at even higher concentrations. Sodium phosphate as buffer was used in both concentrated (100 mM) and diluted form (10 mM, Supplementary table 11). The concentrated version of the buffer proved beneficial for long-term protein stabilization however, the main disadvantage of using sodium phosphate as a buffer system during crystallography is the effortless entering in complex of the phosphate with different salts from the screening kits. Phosphate can form inorganic salt crystals with almost any mono- or divalent alkali metal, making the identification of real protein crystals very difficult. Additionally, even if a concentration of 100 mM sodium phosphate might be stabilizing for Asp1, this concentration is far higher than any that might occur in vivo. The only crystals ever detected were indeed inorganic lithium phosphate crystals which were identified as such by testing the absorbance at 280 nm and even for diffraction (Figure 101).

4.12 Expression and isolation of His-Asp1$^{1-902}$ and His-Asp1$^{33-902}$

An alternate strategy to study the Asp1 pyrophosphatase domain was possible by characterizing Asp1 in its full length. Studying a domain in its separate form can bring many advantages but since obtaining a crystal structure with constructs harbouring only the pyrophosphatase domain was not yet successful, it became apparent that the protein in its native length might naturally crystallize better. To increase the chances of obtaining crystals, a second version of full-length Asp1$^{1-920}$ was designed which lacked the first 32 and last 18 amino acids, thus removing all proposedly unstructured elements from the working sequence. Since Asp1 is a protein large, a decahistidine-tag was a better choice for these constructs. After the full-length *asp1+* gene was integrated into the pET16b-TEV vector using In-Fusion cloning, competent *E. coli* BL21 *ΔiscR* cells were transformed with the freshly prepared DNA. Positive transformants were used for the inoculation of overnight cultures for the conduction of small-scale expression tests to identify optimal expression conditions.

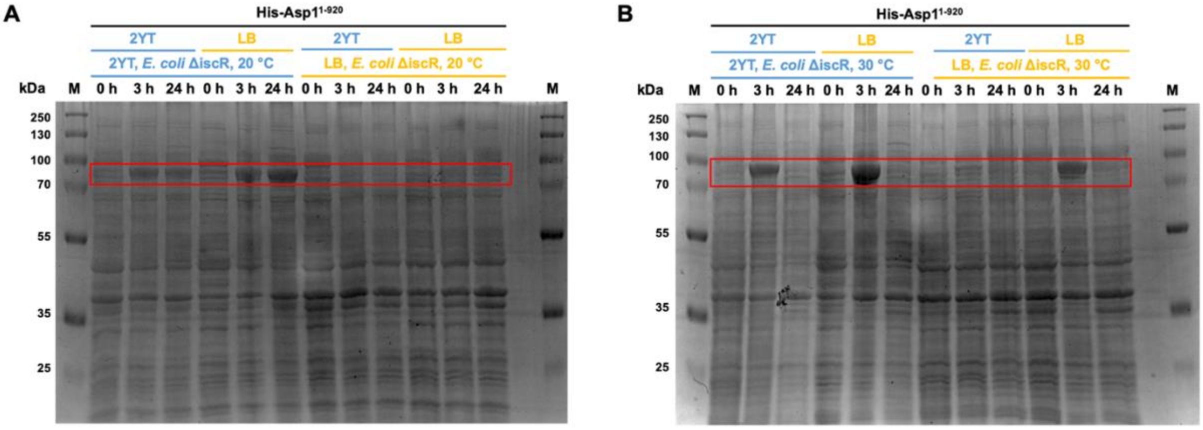


Figure 55: Expression tests using full-length His-Asp1$^{1-920}$ constructs and *E. coli* BL21(DE3) *ΔiscR* cells in two different cultivation media and induction temperatures. A: His-Asp1$^{1-920}$ at 20 °C during induction, B: His-Asp1$^{1-920}$ at 30 °C during induction. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker.

It can clearly be observed that expression of His-Asp1$^{1-920}$ achieves the maximum yield after 3 h induction at 30 °C following cultivation in 2YT medium and inoculation from an overnight culture prepared in LB medium. The choice of the overnight culture makes a considerable difference, since the band from the expression where a LB overnight culture is used is double in size and intensity compared to expression in 2YT medium from the beginning (Figure 55). The expression of His-Asp1$^{1-920}$ at 20 °C does not appear favourable as the overall protein levels are very low compared to the ones prior to induction. A similar tendency can be observed for the truncated version of full-length Asp1 (His-Asp1$^{33-902}$) with maximum expression levels after expression in 2YT for 3 h at 30 °C post induction and starting from an LB overnight culture. Interestingly, the expression could be conducted also under opposite conditions, starting from a 2YT overnight culture and transferring the cell suspension to main cultures in LB medium incubated for 3 h at 30 °C post induction. The expression at 20 °C after induction does not favour the construct His-Asp1$^{33\text{-}902}$ as well since the expression levels are much lower than the ones at 30 °C (Figure 56).

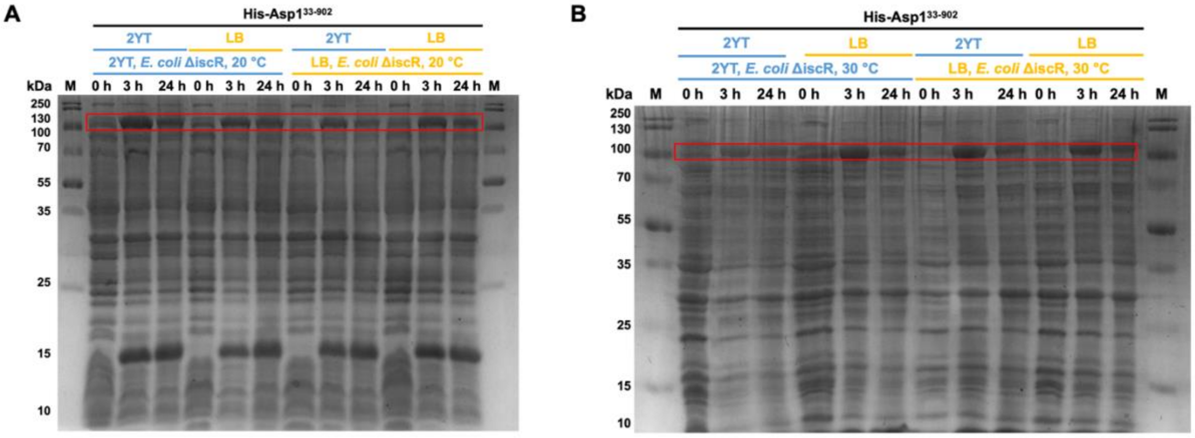


Figure 56: Expression tests using full-length His-Asp1$^{33\text{-}902}$ constructs and *E. coli* BL21(DE3) *ΔiscR* cells in two different cultivation media and induction temperatures. A: His-Asp1$^{33\text{-}902}$ at 20 °C during induction, B: His-Asp1$^{33\text{-}902}$ at 30 °C during induction. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker.

Since the expression protocol for the Asp1$^{365\text{-}920}$ pyrophosphatase domain proved to be applicable of the two full-length Asp1 constructs to an extent, the same strategy was applied for the establishment of a purification protocol for these newly designed variants. Cell disruption and isolation was conducted after the same protocol and using the 100 mM sodium phosphate buffer as well. The isolation was in part successful since protein could be isolated, however the purity was nowhere acceptable. Different protocols and combinations of IMAC and SEC were used to purify His-Asp1$^{1\text{-}920}$ and His-Asp1$^{33\text{-}902}$ but due to the His-tag and its similarity to histidine rich areas from various *E. coli* household proteins it was not possible to isolate the two target constructs in a pure fraction using this affinity tag. SDS-PAGE analyses showed a multitude of bands in the elution fraction with over 12 bands per fraction despite using adequate elution gradients and wash steps. Thus, the design, cloning and expression of two full-length version of *S. pombe* Asp1 were successful however, the purification needs to either be improved or be performed in a different affinity tag system altogether, like the GST-tag. The new cloning of GST-tagged constructs was not foreseen in this study anymore as it would result in too large constructs, whose biochemical and structural characterization would be limited by the considerable size.

5 Discussion

Codon optimization to accommodate heterologous gene expression, *E. coli* strain selection, and effective purification protocols were crucial for achieving high yields of both tagged and untagged Asp1 pyrophosphatase domain proteins.[169] As attempts were made to optimize of protein isolation protocols, the addition of glycerol proved to be key to increase the overall stability of the protein variants in solution in line with literature.[141] Theoretically, the glycerol concentration could be elevated in order to further increase the yield. However, the use of highly concentrated glycerol buffers complicates the chromatographic procedure due to viscosity issues, which have been recently addressed in literature.[140] Thus, the protein variants generated using an optimized expression and purification protocol were stabilized in solution by using 10% glycerol.

Electronic absorption spectra showed that the Asp1 pyrophosphatase domain binds a [2Fe-2S] cluster based on the distinct bands expected for [2Fe-2S] proteins such as ferredoxins.[170–172] The resulting extinction coefficient for the measured Asp1 samples is in good agreement in values reported in literature for [2Fe-2S] proteins.[165] However, the double peaks in the 410 nm to 470 nm range appear wider than usually encountered in electronic spectra of [2Fe-2S] proteins. This gives the impression that there is only one broad band in the 400 nm region, which is more characteristic of [4Fe-4S] proteins.[173] This observation gave rise to the question if the iron concentration per monomer could be increased in the pyrophosphatase domain and how would the electronic spectrum of the Fe-S look under these conditions. It could be shown that the iron content can be increased up to 4 Fe/monomer using chemical reconstitution. This method often leads to a considerable amount of impurities in form of Fe-S aggregates which are visible as a broad band at 600 nm. The observed [4Fe-4S] cluster could be an artifact from the chemical reconstitution. However, data from this thesis showed that the iron is coordinated to the protein due to the absence of signals in the 500 nm to 600 nm range, with 500 nm being the characteristic wavelength for free ferric iron.[174] Based on comparison with spectra of [4Fe-4S] proteins, it appears that the Asp1 pyrophosphatase domain can bind up to 4 Fe per monomer and exhibit a broad absorption band in the 400 nm range which is more characteristic of [4Fe-4S] proteins.[166,173,175]

The established protocol for expression and purification lead to a significant increase in the cell pellet and protein yield. However, electronic spectra showed that the expression and purification conditions do not seem to facilitate fully maturation of the cofactor even though iron, sulfur and fumarate are added in excess to the cultivation media. This might indicate that the iron uptake by *E. coli* is not optimal, even though ferric ion is predominantly present in the cultivation medium under aerobic conditions.[176] Under physiological conditions iron exists in either its reduced Fe^{2+} or its oxidized Fe^{3+} form.[177–180] However, the reduced form is unstable under aerobic conditions and could be oxidized to ferric iron during cultivation of bacterial cultures, overall making the loading of Asp1 with cofactor more difficult in vivo. In addition, ferric iron (Fe^{3+}) is usually sequestered in host proteins such as transferritin, lactoferritin and ferritin until it finds its further use as part of structural elements in proteins or in iron homeostasis.[176] The results showcase a possibility that the iron uptake in the *E. coli* strain might not proceed smoothly and thus limit the intracellular amount of iron available for Fe-S cluster assembly. This way, even though expression of the Asp1 variants was optimized, the proteins are not fully loaded with cofactor after their in vivo heterologous expression.

Chemical reconstitution experiments showed that different iron sources and reductants affect spectral properties, with GSH notably improving signal clarity. The possibility of GSH to play a special role as binding partner is further highlighted by the fact that GSH is a natural ligand in the coordination of Fe-S clusters in glutaredoxins.[150,181,182] At the same time, one ligand-free state and three different ligand-bound states could be identified for the Asp1 pyrophosphatase domain. It was shown that Asp1 can go from overall coordinating 1 Fe/monomer up to 4 Fe/monomer, a transition phenomenon that has been observed for another mitochondria-associated Fe-S protein ATE1 of *S. cerevisiae*.[183] Not only does the amount of iron per monomer fitting to every state but a progressional increase in the extinction coefficient at 320 nm and 410 nm is observed. In addition, the lack of absorption bands at 500 nm and 600 nm deems the samples free of free ferric ion or Fe-S aggregates.[174] Overall, the observation that the Asp1 pyrophosphatase domain can bind different amounts of iron sheds light on new possibilities of how the cluster could function in vivo. The coordination of 4 Fe/monomer which is attributed to a [4Fe-4S] cluster based on distinct absorption bands that are in good agreement with relevant studies.[173] At the same time, the observed [4Fe-4S] cluster could be an artifact from chemical reconstitution. However, this possibility cannot be confirmed since none of the typical absorption bands for free iron and Fe-S aggregates are present in the electronic spectrum.

Events of competitive inhibition of iron binding sites in the cluster by zinc were observed after gene overexpression using the expression protocol tailored for Asp1$^{365\text{-}920}$ production. This phenomenon has been observed for [2Fe-2S] proteins since the ferrous iron found as the predominant iron species in [2Fe-2S] clusters are divalent just as zinc.[146] An increase in the local iron abundance during chemical reconstitution remove zinc ions as competitive binders and confirm that only iron is involved in cluster coordination. This confirms that the designed protocol though leading to high yields of correctly folding protein does not promote full cluster maturation. This raises the question of promiscuity of the cofactor in regard to the coordinated metal, although zinc has not been observed to contribute to cluster assembly of any kind.

The incomplete cluster maturation was shown to promote oligomeric states. This phenomenon was reported in literature as well.[184] Further analyses confirm that the oligomeric state cannot be reversed to a dimeric or monomeric state by variation of reducing agents or of the glycerol concentration. In this work the only combination of methods that lead to a monomeric state are chemical reconstitution in vitro with significant iron and sulfur excess and subsequent SEC to remove Fe-S aggregates. The Asp1$^{365\text{-}920}$ Fe-S cluster could be the only element impacting the dispersion pattern of this domain. This might be a strong indication that this cofactor plays a structural role in correct folding, which implies a possible location of the cofactor at the surface of the protein.

The CD spectra presented in this thesis showed that the presence of iron impacts the secondary structure, with shifts between states rich in α-helix and random coil (apo and 2 Fe/monomer) and those with higher β-sheet content (1 and 4 Fe/monomer). Denaturation curves revealed that iron binding increases protein stability, with the melting temperature rising from 48.3 °C (apo) to 51.8 °C (4 Fe/monomer). A mutant variant lacking four iron-binding cysteine residues exhibited similar structural features to the apo form, validating its use as a negative control. In addition, the pyrophosphatase domain's interaction with a substrate analog, 1,5-$(PCP)_2$-IP_4, caused no significant changes in the CD spectra. This observation is very interesting since proteins are more flexible before entering a complex with their substrate. The binding of the substrate increases protein rigidity since the substrate-enzyme complex is usually stable and in an energetically favourable state. Analysis of His-tagged constructs confirmed the influence of iron on secondary structure and stability, with minimal differences compared to the GST-tagged versions. Taken together, it appears that the Asp1 [2Fe-2S] cluster stabilizes α-helical regions and contributes slightly to the overall stabilization of the protein.

The results of the Mössbauer data draw attention to the two different iron species present in the Asp1 [2Fe-2S] cluster. The measured isomer shifts and quadrupole splittings were consistent with Fe-S clusters but deviated from strictly cysteine-ligated [2Fe-2S] clusters. It suggested potential histidine involvement in the ligation of the cluster, supported by other studies involving His-ligated Fe-S clusters.[164,167] Further analysis after cleaving the protein's affinity tag indicated similar spectra, but the data couldn't be fully analyzed due to a low sample concentration. Interestingly, no nitrogen signal could be measured in a different EXAFS data set of a similarly prepared sample of the Asp1 pyrophosphatase domain. This reveals that histidine residues cannot be located in the first coordination sphere of the Fe-S cluster although the Mössbauer spectral features conclude this.[3] The only explanation for this phenomenon would be a transient effect of a histidine residue based on an intramolecular interaction in which the Fe-S cluster is involved. This theory is even more fascinating when considering the biological relevance of the H397 residue for the Asp1 pyrophosphatase domain.

Efforts to enhance solubility, including lowering induction temperature, co-expression of chaperones, and buffer optimization with glycerol, urea, and different buffer systems, were unsuccessful in the optimization and establishment of a purification protocol for Asp1 variants GST-Asp1$^{365\text{-}920\ C643S\ C644S}$ and GST-Asp1$^{383\text{–}918\ C868S}$. Limited soluble protein was obtained for GST-Asp1$^{383\text{–}918\ C868S}$, but its concentration was too low for reliable characterization via EAS. Based on previous data and the difficulty to obtain these two mutant variants in a pure soluble fraction, cysteines 868, 643, and 644 are very likely crucial for structural integrity rather than directly coordinating the Fe-S cluster.[185] This suggests that the loss of solubility might be due to the importance of these cysteines in maintaining the protein’s structure.[3] In addition, these findings are in line with describes roles of cysteines in the maintenance of protein structure.[185] Interestingly, the previously characterized four cysteines that are supposed to coordinate the Asp1 [2Fe-2S] cluster are localized in two separate structural areas of the pyrophosphatase domain. Cysteines C607 and C663 appear on a helix structure on the surface of the domain whereas the other two cysteine residues C864 and C879 are localized inside the protein core (Figure 57). This assumption that the four are coordinating the same Fe-S cluster rather questionable, as the two pairs are localized at distances at approx. 30 A which is high over the value of a typical cys-Fe coordinative bond. At the same time, it appears that the protein core is crowded with other cysteines and residues relevant for the pyrophosphatase activity.

The AlphaFold2 software which was used to generate this structure prediction does not take cofactors into account for the calculation of the most probable tertiary structure. However, the model confidence values for many relevant regions in the pyrophosphatase domain are high so the proposed structural layout by AlphaFold2 could be correct.[128,129,131] Thus, it could be possible for the four cysteines that were previously assumed to be in close proximity to each other and bind a [2Fe-2S] cluster to actually be located in two different sites of the pyrophosphatase domain altogether. Furthermore, the active site of Asp1$^{365-920}$ seems to harbour many amino acid residues with catalytic functions such as I808 and H397 and other cysteines with proposed functions as structural elements.

The measurements in yellow indicate that the distances among the residues are large enough for a [2Fe-2S] cluster to be situated there. Amino acids C607 and C663 have been investigated earlier as well and deemed as potential binding partners for a [2Fe-2S] cluster. This would be accomplished with two GSH molecules as external ligands. Interestingly, this hypothesis might be supported by the data showing the chemical reconstitution with GSH and two different iron sources, where it can be observed that the use of this reductant influences the shape and intensity of the spectral signals. Regarding the active site of the protein, a cluster could be coordinated by C864, C879 and two more cysteines in the active site. Residue H397 being in close proximity to the proposed cluster in the active site might explain the signal in the Mössbauer spectrum which identifies two different iron species. The electronegative pull from the benzol ring of H397 possibly affects only the electron density of the one iron atom in its vicinity. At the same time, H397 not being in the first coordination sphere of the iron ion in question, makes the nitrogen signal from the histidine residue not identifiable by XAS.

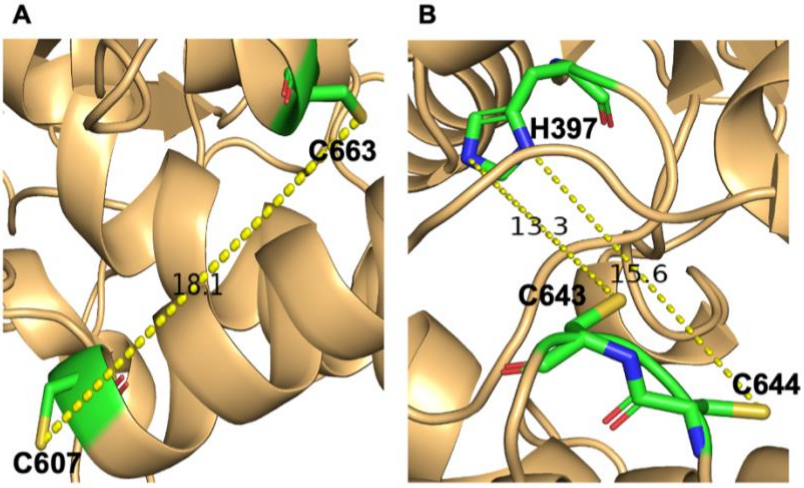


Figure 57: Close-up of the areas in the pyrophosphatase domain which key amino acid residues are located. A) Measurement between C607 and C663. B) Measurement between H397 and C643 and C644. Measurements in yellow indicate distances between amino acids.

The different data sets on the pyrophosphatase domain of *S. pombe* Asp1$^{365-920}$ could also be interpreted in a way that leans on the existence of two different [2Fe-2S] clusters located inside the active site and on the surface of the pyrophosphatase domain. The surface [2Fe-2S] cluster could possibly be responsible for all oligomerization events which were observed under different maturation conditions of the cofactors. The choice to use a larger affinity tag, as the GST-tag, is known to promote dimerization as well, but the SEC data stion of this study indicates that GST-Asp1$^{365-920}$ can exist predominantly monomeric dispersed when both clusters are fully maturated (4 Fe/monomer protein).[158] Moreover, the possibility of the cofactor to transition between different cluster architectures, which has been observed for the mitochondria-associated ATE1 of *S. cerevisiae*, was rarely attempted or investigated by chemical reconstitution with iron and sulfur excess ratios typical for [4Fe-4S] proteins.[183] This method of reconstitution is prone the build-up of impurities visible as a broad peak at 600 nm in EA spectra, however the EAS data presented in this study does not indicate any of the expected impurities. These were observed only during treatment of GST-Asp1$^{365-920}$ with radioactive-labeled ^{57}Fe and even in that case they were removed by gel filtration. At the same time, the CD data depicts very distinct structural changes where a similar conformation is adopted for states containing 0 and 2 Fe/monomer and another β-strand-rich one for states with 1 and 4 Fe/monomer. Taken together, the combined data indicate that chemical reconstitution leads to the formation of a [4Fe-4S] cluster that is biologically not relevant and does not coincide with the reported data on the as-isolated Asp1 pyrophosphatase domain and that the spectral properties of present [2Fe-2S] cluster cannot be fully explained with coordination patterns known in biology.

Certain amino acids which are prone to act as structural elements (C644, C868) maintain a necessary tertiary structure of the domain so that cluster coordinating residues can be in vicinity to each other. Moreover, the loop in the pyrophosphatase domain predicted by AlphaFold2 might be a hint towards a structurally flexible regulating element that could be involved in substrate binding. Molecules with two pyrophosphate groups such as 1,5-IP_8 have a considerable size and require certain space to reach the binding pocket and partake in the reaction. The necessary space and domain flexibility might be facilitated by making use of the loop which could serve as a gate between an open and closed substrate binding site. This possibility for the regulation of the pyrophosphatase domain is supported by the biological data showing that I808 has a tremendous effect and that the quadruple mutant variant might lead to a maximum of overregulated pyrophosphatase activity. The H397 residue appears to be biologically highly relevant and introduces complex questions regarding the coordination environment of the [2Fe-2S] cluster in the active site as well as its potential biological function. Unfortunately, the in-depth questions regarding the size of the binding pocket and potential steric hindrance around the substrate cannot be clarified as crystallographic attempts to obtain a tertiary structure of the pyrophosphatase domain in both substrate- and/or cluster-bound and/or unbound form were not successful. The data from this study opens the possibility of further experiments to obtain a more complete image of the Asp1$^{365\text{-}920}$ or even of the full-length Asp1$^{1\text{-}920}$ tertiary structure and the related regulation events.

The goal of sparse matrix screening was to obtain crystals and potentially obtain a crystal structure data set for the Asp1 pyrophosphatase domain. However, all attempted experiments were unsuccessful, and it was not possible to crystallize the Asp1 pyrophosphatase domain. Since this approach proved unsuccessful, a new strategy might be the employ of cryo-electron microscopy (cryo-EM) for which a full-length Asp1$^{1\text{-}920}$ construct with an affinity tag of considerable size is needed to meet the requirements of this method. A protein structure is very likely to shed light on all observed events regarding this regulatory domain so far. At the same time, it is of great importance to quantify the elevated 1,5-IP$_8$ levels produced by the quadruple mutant variant by HPLC profiling. Furthermore, the biological regulation of the Asp1$^{365\text{-}920}$ domain can be investigated by growing strains with known mutant variants on lower or higher iron levels than physiologically normal and observe how the iron homeostasis is influenced by this.

The biochemical and spectroscopic data presented in this thesis demonstrate that the Asp1 pyrophosphatase domain of *S. pombe* harbors an iron-sulfur (Fe-S) cluster, with evidence for dynamic coordination of up to four iron atoms per monomer. This observation raises intriguing questions about the molecular context in which the Fe-S cluster is assembled and maintained in vivo, especially given the incomplete cluster maturation observed following heterologous expression in *E. coli*. Since Fe-S cluster biogenesis in eukaryotes is tightly regulated by the CIA (cytosolic iron-sulfur assembly) machinery, a functional link between Asp1 and the CIA pathway becomes a relevant hypothesis to explore.

Although direct interactions between Asp1 and CIA components such as Nar1, Cfd1, Nbp35, or Mms19 were not identified in this study, several lines of evidence suggest a possible interface.[186,187] Firstly, the observed dependence of cluster coordination on iron availability and the promiscuity for iron versus zinc incorporation highlight that Asp1 may rely on a dedicated, regulated assembly system in its native cellular environment. The CIA machinery is known to selectively assemble [4Fe-4S] clusters on specific target proteins, including several nuclear enzymes involved in DNA metabolism and ribosome biogenesis, but it has also been implicated in the maturation of proteins involved in signaling and metabolic control. Given the dual localization of Asp1 (cytosolic and nuclear) and its central role in inositol pyrophosphate signaling, it is conceivable that Asp1 could belong to a subclass of CIA client proteins that are involved in cellular stress response or metabolic regulation.

Moreover, the Mössbauer and CD data reveal that the Asp1 Fe-S cluster likely adopts a non-canonical coordination environment, potentially involving histidine and/or GSH ligands. This atypical coordination, which may explain the functional lability and redox responsiveness of the cofactor, is reminiscent of other known CIA clients with unique Fe-S coordination geometries, such as the tRNA modification enzyme Cia1 or the mitochondrial outer membrane protein ATE1.[188] The similarity to

ATE1 is particularly noteworthy, as this protein also exhibits Fe-S cluster remodeling behavior and shares a role in iron-dependent protein regulation.

The requirement for GSH in cluster reconstitution and the dynamic range of iron incorporation (from 1 to 4 Fe per monomer) may point toward a regulatory rather than purely structural role for the Fe-S cluster in Asp1. This is further supported by the changes in secondary structure and stability linked to iron binding, suggesting that Fe-S cluster coordination could influence domain folding or oligomerization states, potentially regulating access to the substrate-binding site.[189,190] Such regulation mechanisms would be consistent with the function of a CIA client, where Fe-S cluster insertion serves to modulate protein activity or stability in response to cellular iron status.

From a structural perspective, the spatial separation of the cysteine residues proposed to coordinate the cluster, as shown by AlphaFold2 modeling, suggests the possibility of multiple cluster-binding sites or the involvement of larger structural rearrangements. These features could necessitate a highly specific and regulated assembly mechanism in vivo, potentially involving CIA chaperones or scaffold proteins.

Taken together, while no direct biochemical interactions with CIA components were detected in this study, the complex Fe-S cluster properties of Asp1, its iron-dependence, and the apparent structural plasticity strongly suggest that it may be influenced by or directly depend on the CIA machinery in *S. pombe*. Future studies could aim to test this hypothesis by assessing Asp1 functionality and Fe-S cluster maturation in strains lacking CIA components, or through proteomic co-immunoprecipitation approaches under iron-limiting conditions. Additionally, in vivo complementation assays using Asp1 cluster-binding mutants or iron supplementation experiments may help clarify the relevance of iron homeostasis and CIA activity for Asp1 regulation.

6 Supporting data

6.1 Supplementary tables

Supplementary table 1: *E. coli* strains used in this study.

Name	**Genotype**	**Reference**
BL21(DE3)	F^{-} *omp*T *hsd*S$_{\mathrm{B}}$ (r$_{\mathrm{B}}^{-}$, m$_{\mathrm{B}}$) *gal dcm* (DE3)	ThermoScientific C600003
NiCo21 (DE3)	*can::CBD fhuA2 [lon] ompT gal (λ DE3) [dcm] arnA::CBD slyD::CBD glmS6Ala ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	New England Biolabs C2529H
BL21(DE3) *ΔiscR*	F^{-} *omp*T *hsd*S$_{\mathrm{B}}$ (r$_{\mathrm{B}}^{-}$, m$_{\mathrm{B}}$) *gal dcm iscR*::kan$^{\mathrm{R}}$ (DE3)	Jones et al 2008
XL1-blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F$'$ *proAB lacI*$^{\mathrm{q}}$Z*ΔM15* Tn*10* (Tet$^{\mathrm{r}}$)]	Agilent Scientific #200249
605	*his3-D1, ade6-M210, leu1-32, ura4-D18,* h^{-}	K. Gould
2553	*asp1*$^{\mathrm{I808D}}$::kan$^{\mathrm{R}}$, *his3-D1, ade6-M210, leu1-32, ura4-D18,* h^{-}	U. Fleig
3292	*asp1*::kan$^{\mathrm{R}}$, *his3-D1, ade6-M210, leu1-32, ura4-D18,* h^{-}	U. Fleig
3294	*asp1*$^{\mathrm{C607S,C663S,C864S,C879S}}$::kan$^{\mathrm{R}}$, *his3-D1, ade6-M210, leu1-32, ura4-D18,* h^{-}	U. Fleig

Supplementary table 2: Overview of the yields obtained for each purified protein variant.

Protein variant	**Yield / g pellet**	**Yield / L culture**
GST-Asp1$^{365\text{-}920}$	60 mg	180 mg
GST-Asp1$^{383\text{-}918}$	46 mg	137 mg
GST-Asp1$^{365\text{-}920\ \mathrm{H397A}}$	35 mg	124 mg
GST-Asp1$^{365\text{-}920\ \mathrm{C607S}}$	25 mg	95 mg
GST-Asp1$^{365\text{-}920\ \mathrm{C643S}}$	47 mg	151 mg
GST-Asp1$^{365\text{-}920\ \mathrm{C663S}}$	59 mg	176 mg
GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$	29 mg	72 mg
His-Asp1$^{365\text{-}920}$	43 mg	129 mg
His-Asp1$^{383\text{-}918}$	52 mg	154 mg
His-Asp1$^{1\text{-}920}$	36 mg	133 mg
His-Asp1$^{33\text{-}902}$	48 mg	164 mg

6.2 Expression and purification of the Asp1$^{365\text{-}920}$ pyrophosphatase domain

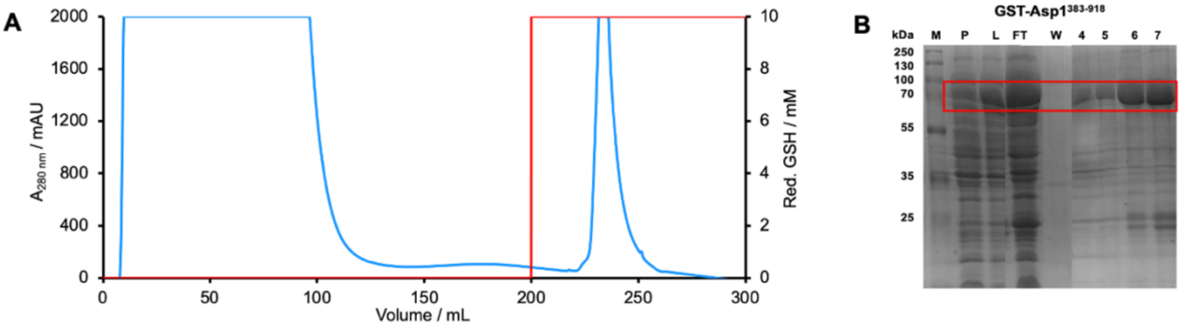


Figure 58: Overview of the purification results of wildtype GST-Asp1$^{383\text{-}918}$ variant via GSTrap affinity chromatography. A) Chromatogram displaying the course of the isolation. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis. The concentration of reduced glutathione (red curve) in mM used for eluting the target protein is listed on the secondary y-axis. B) SDS-PAGE showing fractions of the anaerobic purification of GST-Asp1$^{383\text{-}918}$. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, W = wash. The numbers indicate the different elution fractions.

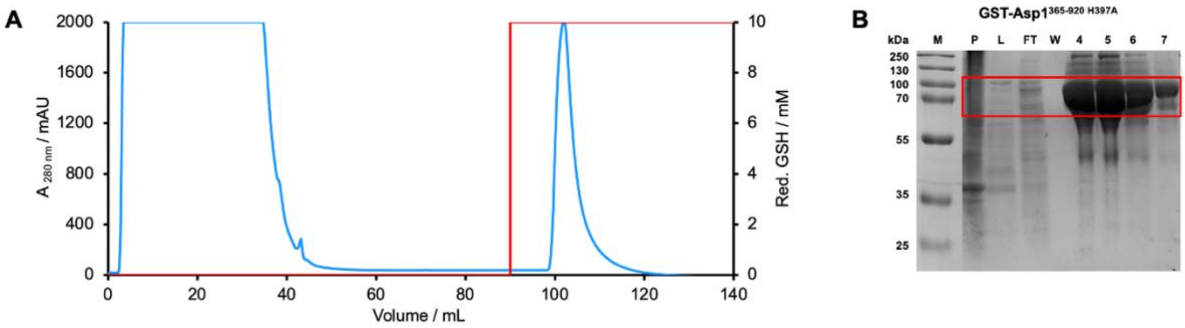


Figure 59: Overview of the purification results of GST-Asp1$^{365\text{-}920\ H397A}$ mutant variant via GSTrap affinity chromatography. A) Chromatogram displaying the course of the isolation. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis. The concentration of reduced glutathione (red curve) in mM used for eluting the target protein is listed on the secondary y-axis. B) SDS-PAGE showing fractions of the anaerobic purification of GST-Asp1$^{383\text{-}918}$. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, W = wash. The numbers indicate the different elution fractions.

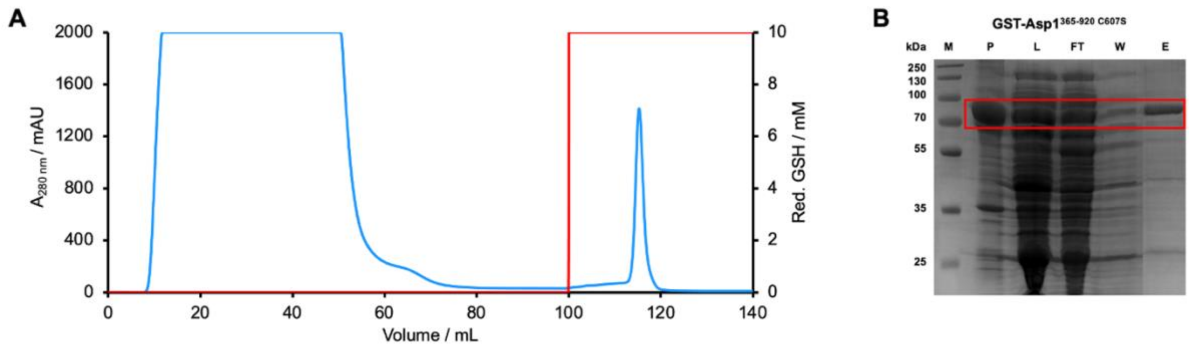


Figure 60: Overview of the purification results of GST-Asp1$^{365-920\ C607S}$ mutant variant via GSTrap affinity chromatography. A) Chromatogram displaying the course of the isolation. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis. The concentration of reduced glutathione (red curve) in mM used for eluting the target protein is listed on the secondary y-axis. B) SDS-PAGE showing fractions of the anaerobic purification of GST-Asp1$^{365-920\ C607S}$. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, W = wash, E = elution.

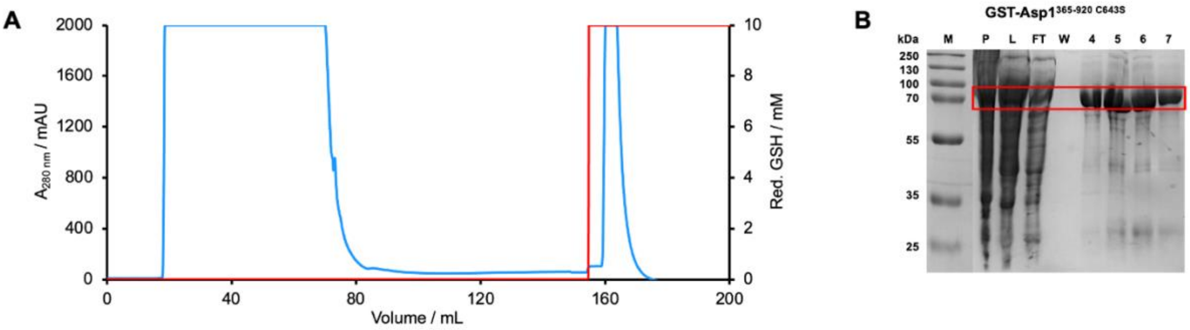


Figure 61: Overview of the purification results of GST-Asp1$^{365-920\ C643S}$ mutant variant via GSTrap affinity chromatography. A) Chromatogram displaying the course of the isolation. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis. The concentration of reduced glutathione (red curve) in mM used for eluting the target protein is listed on the secondary y-axis. B) SDS-PAGE showing fractions of the anaerobic purification of GST-Asp1$^{365-920\ C643S}$. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, W = wash. The numbers indicate the different elution fractions.

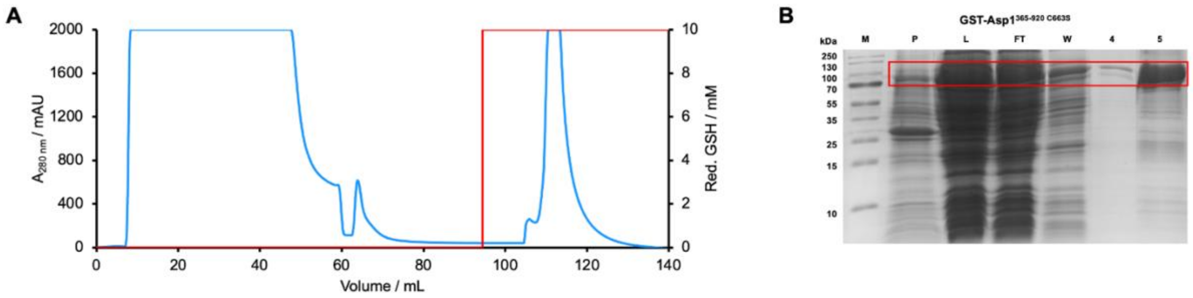


Figure 62: Overview of the purification results of GST-Asp1$^{365\text{-}920\ C663S}$ mutant variant via GSTrap affinity chromatography. A) Chromatogram displaying the course of the isolation. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis. The concentration of reduced glutathione (red curve) in mM used for eluting the target protein is listed on the secondary y-axis. B) SDS-PAGE showing fractions of the anaerobic purification of GST-Asp1$^{365\text{-}920\ C663S}$. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, W = wash. The numbers indicate the different elution fractions.

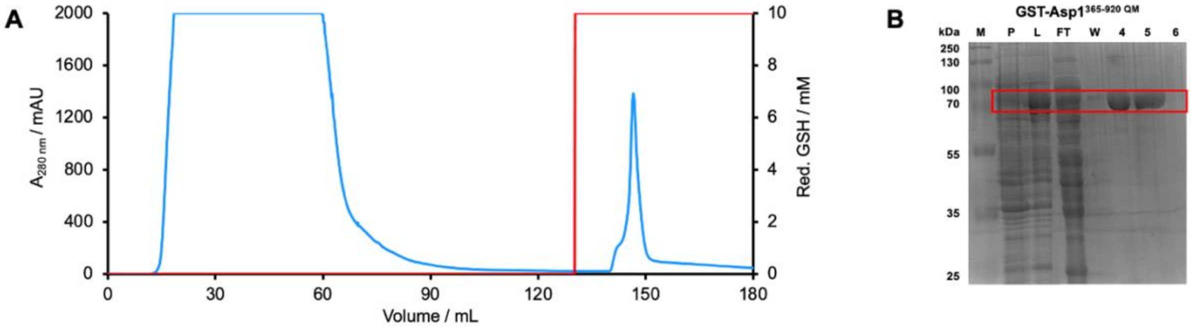


Figure 63: Overview of the purification results of GST-Asp1$^{365\text{-}920\ QM}$ mutant variant via GSTrap affinity chromatography. A) Chromatogram displaying the course of the isolation. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis. The concentration of reduced glutathione (red curve) in mM used for eluting the target protein is listed on the secondary y-axis. B) SDS-PAGE showing fractions of the anaerobic purification of GST-Asp1$^{365\text{-}920\ QM}$. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, W = wash. The numbers indicate the different elution fractions.

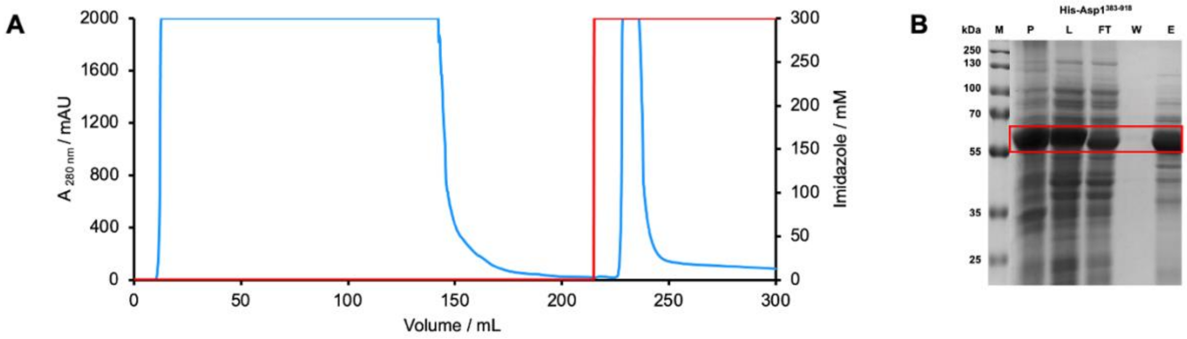


Figure 64: Overview of the purification results of the wildtype His-Asp1$^{365\text{-}920}$ variant via IMAC. A) Chromatogram displaying the course of the isolation. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis. The concentration of imidazole (red curve) in mM used for eluting the target protein is listed on the secondary y-axis. B) SDS-PAGE showing fractions of the anaerobic purification of GST-Asp1$^{383\text{-}918}$. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker. The fractions were abbreviated as following: M = marker, P = pellet, L = lysate, W = wash, E = elution.

6.3 Characterization of the Asp1$^{365\text{-}920}$ pyrophosphatase domain

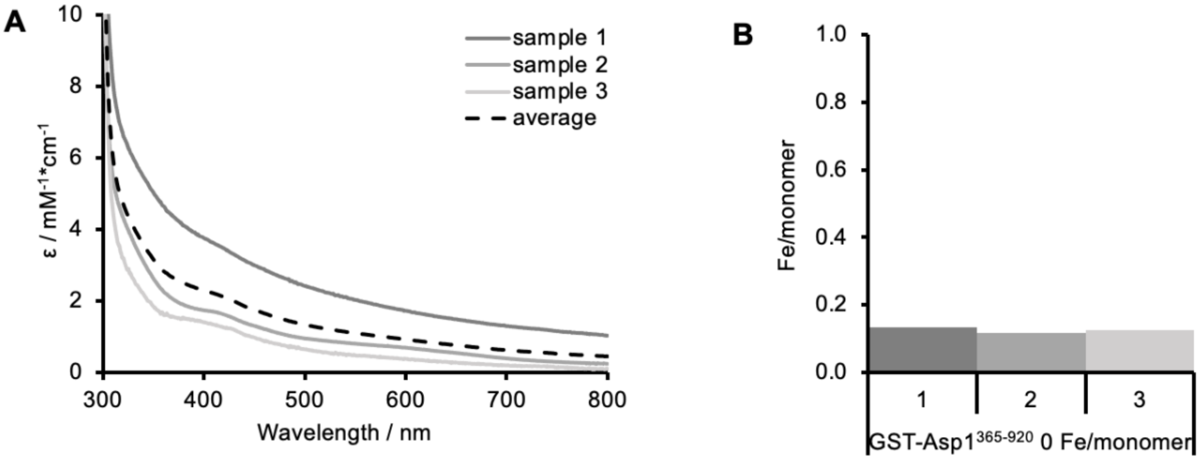


Figure 65: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365\text{-}920}$ with 0 Fe/monomer (apo) including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological triplicates are shown in a colour gradient.

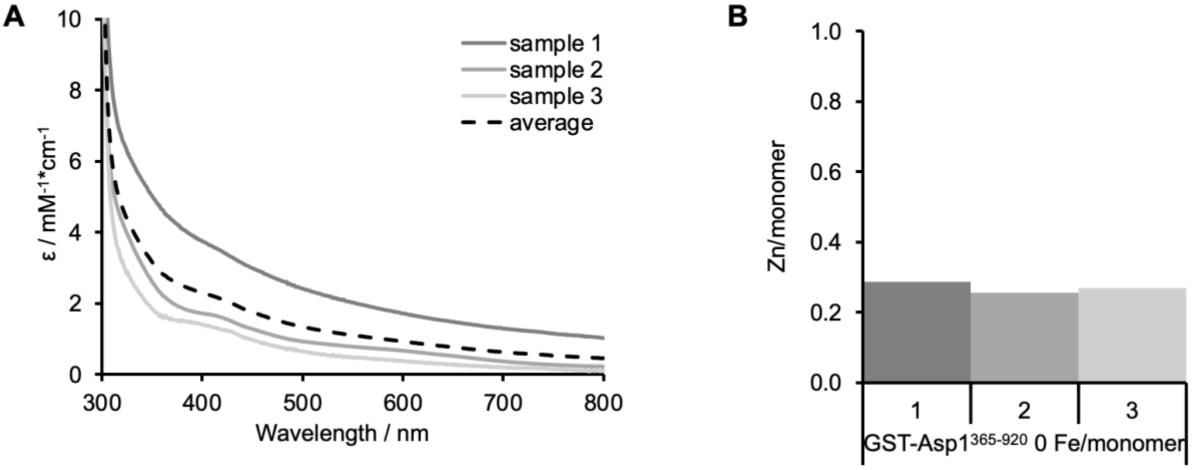


Figure 66: Overview of the electronic absorption spectroscopy and determined zinc content for GST-Asp1$^{365\text{-}920}$ with 0 Fe/monomer (apo) including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

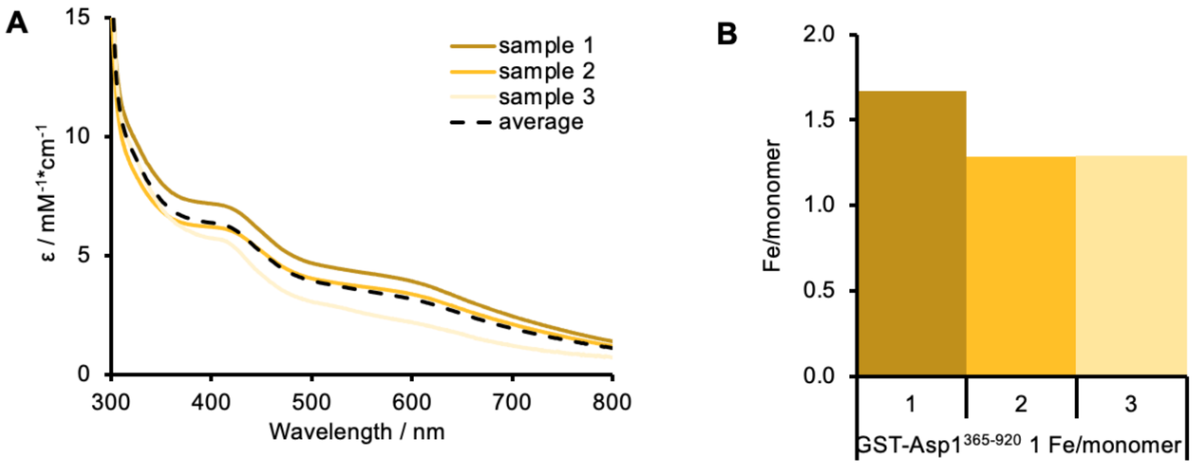


Figure 67: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365\text{-}920}$ with 1 Fe/monomer (as-isolated) including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) he measured Fe levels among the biological triplicates are shown in a colour gradient.

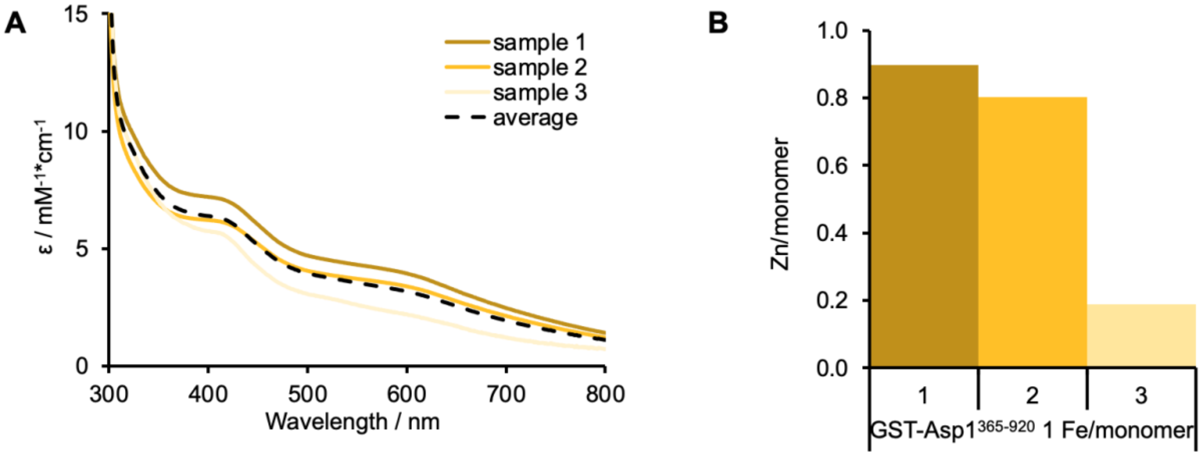


Figure 68: Overview of the electronic absorption spectroscopy and determined zinc content for GST-Asp1$^{365\text{-}920}$ with 1 Fe/monomer (as-isolated) including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

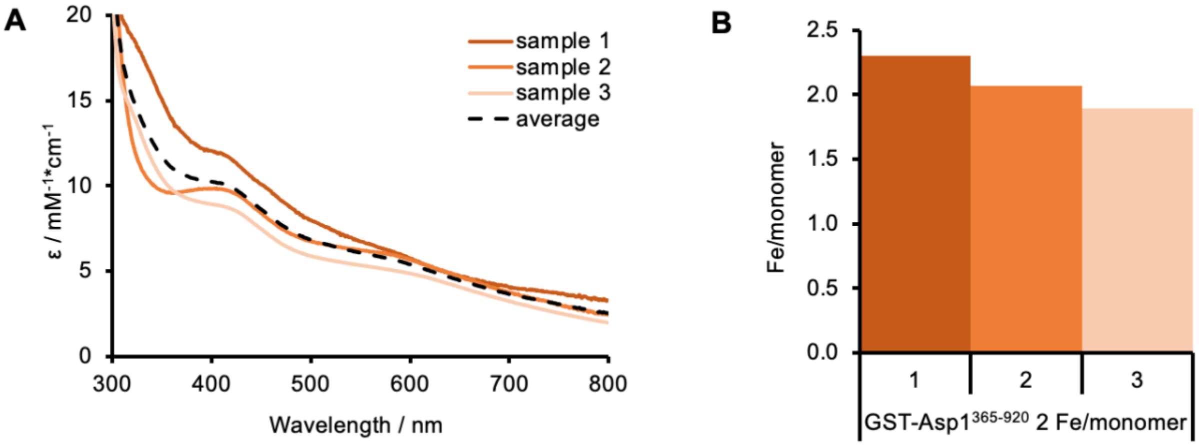


Figure 69: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365\ 920}$ with 2 Fe/monomer (5x excess reconstitution) including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological triplicates are shown in a colour gradient.

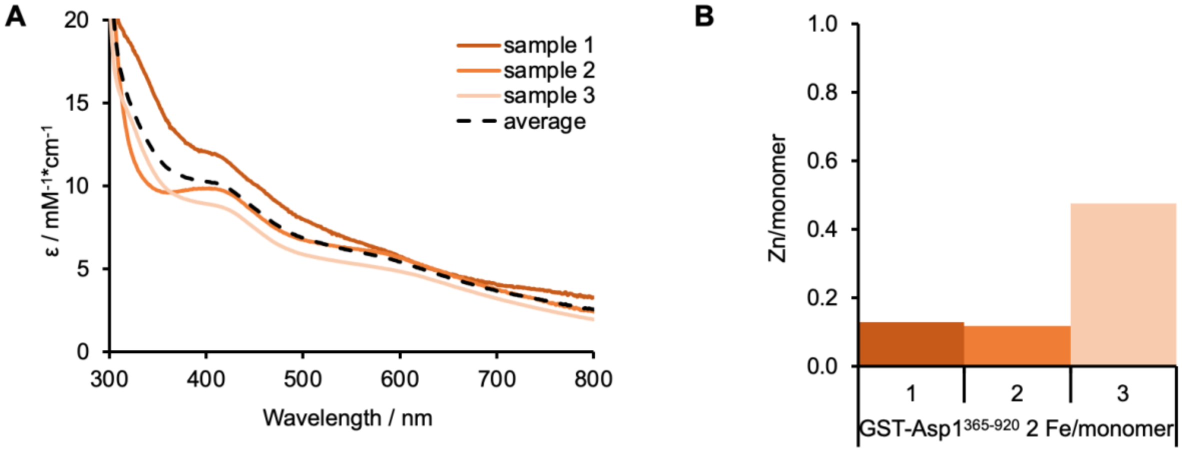


Figure 70: Overview of the electronic absorption spectroscopy and determined zinc content for GST-Asp1$^{365-920}$ with 2 Fe/monomer (5x excess reconstitution) including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

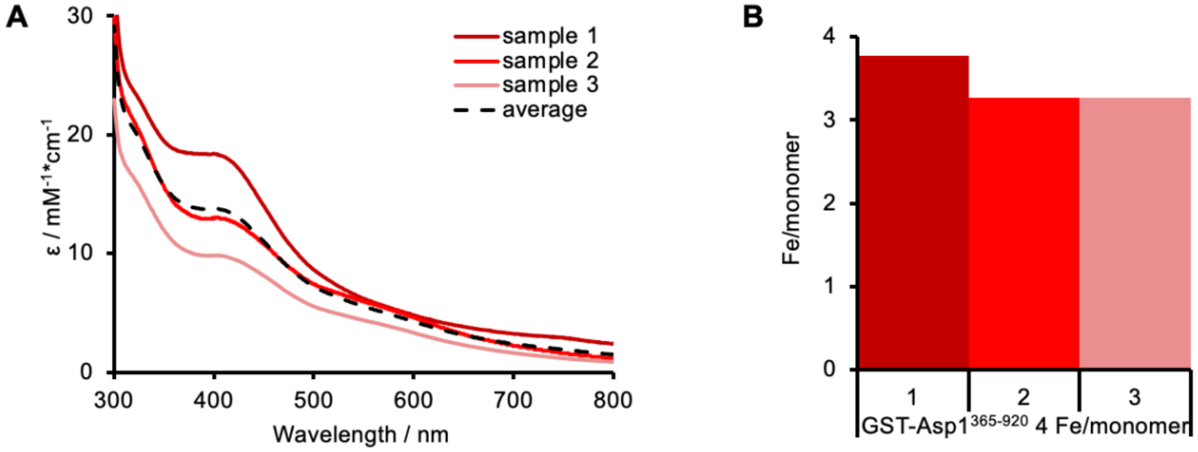


Figure 71: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365-920}$ with 4 Fe/monomer (10x excess reconstitution) including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological triplicates are shown in a colour gradient.

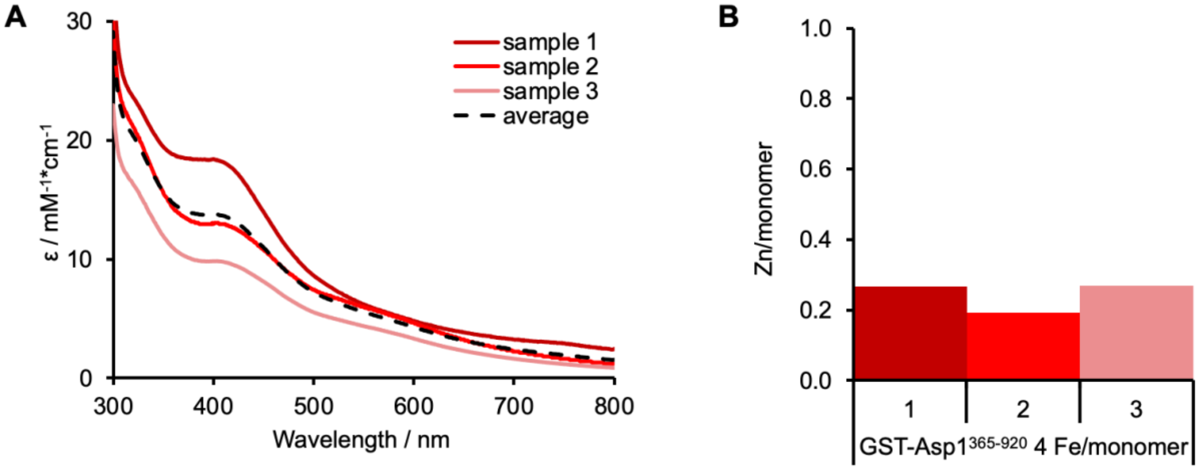


Figure 72: Overview of the electronic absorption spectroscopy and determined zinc content for GST-Asp1$^{365-920}$ with 4 Fe/monomer (10x excess reconstitution) including biological triplicates. A)The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

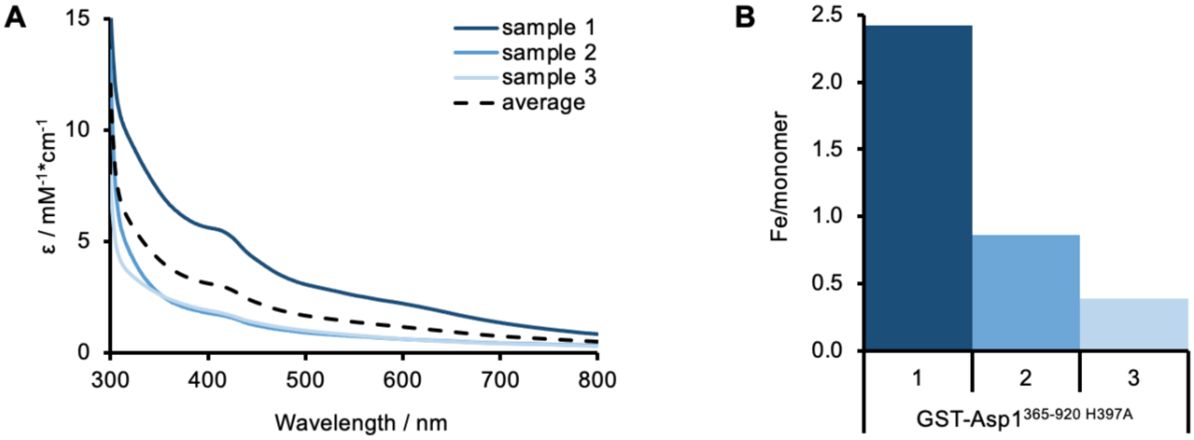


Figure 73: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365\text{-}920\ H397A}$ including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological triplicates are shown in a colour gradient.

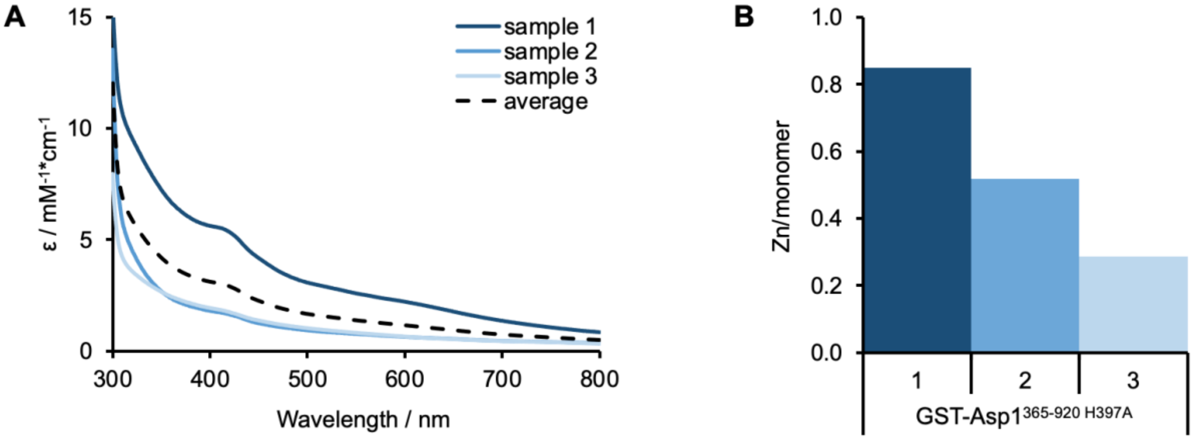


Figure 74: Overview of the electronic absorption spectroscopy and determined zinc content for GST-Asp1$^{365\text{-}920\ H397A}$ including biological triplicates. A)The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

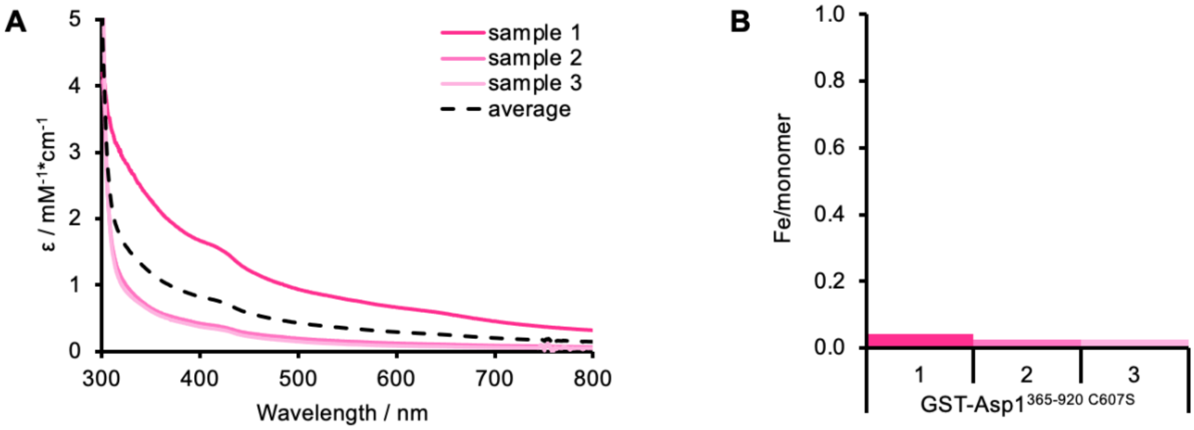


Figure 75: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365\text{-}920\ \mathrm{C607S}}$ including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological triplicates are shown in a colour gradient.

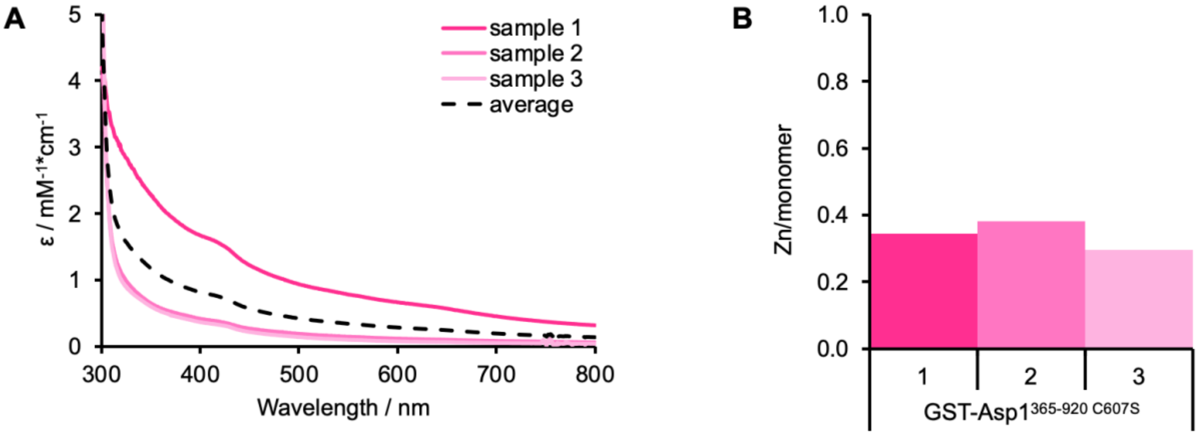


Figure 76: Overview of the electronic absorption spectroscopy and determined zinc copntent for GST-Asp1$^{365\text{-}920\ \mathrm{C607S}}$ including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

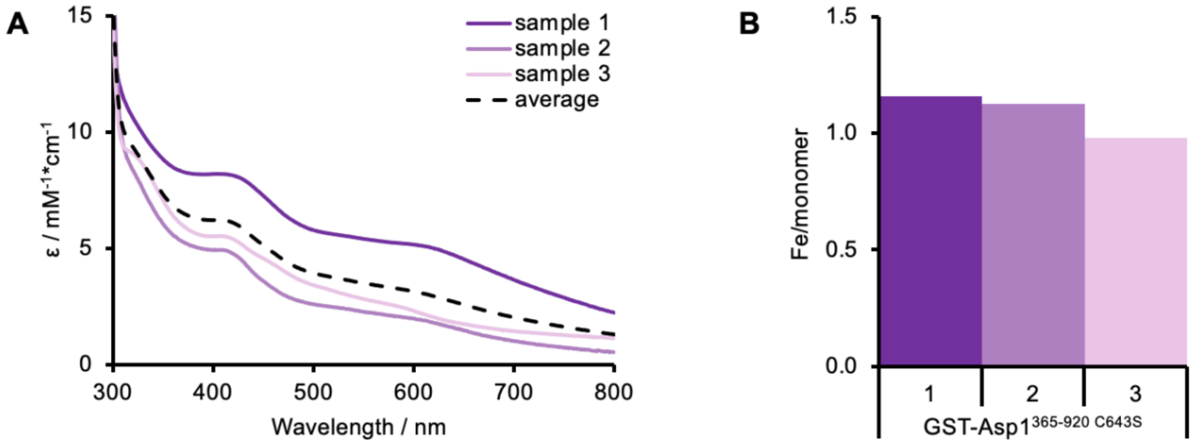


Figure 77: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365-920\ C643S}$ including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological triplicates are shown in a colour gradient.

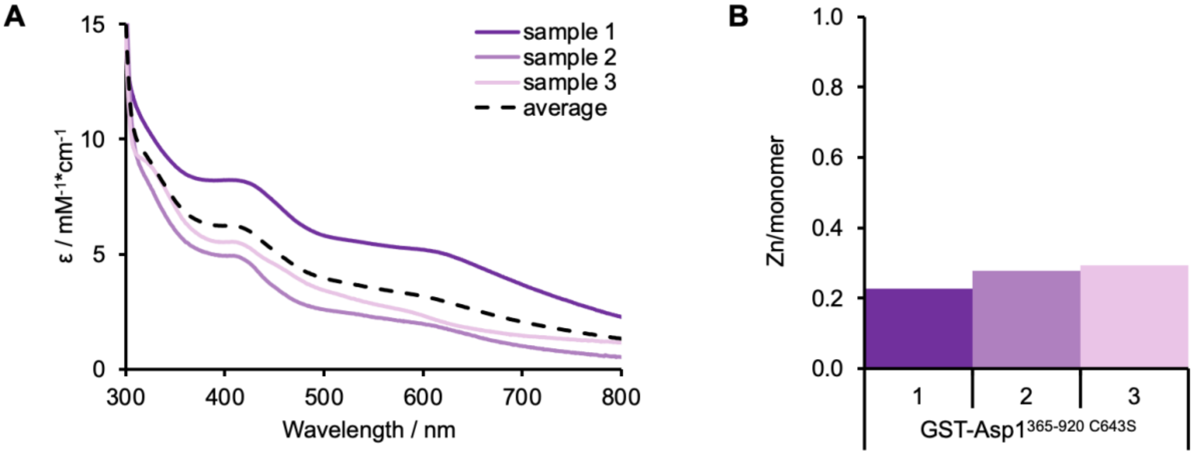


Figure 78: Overview of the electronic absorption spectroscopy and determined zinc content for GST-Asp1$^{365-920\ C643S}$ including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

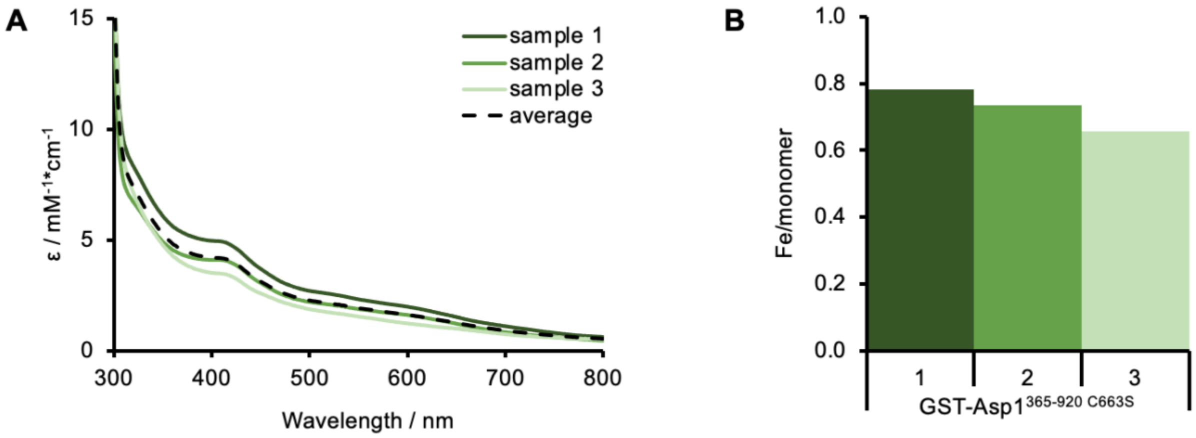


Figure 79: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365-920\ C663S}$ including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological triplicates are shown in a colour gradient.

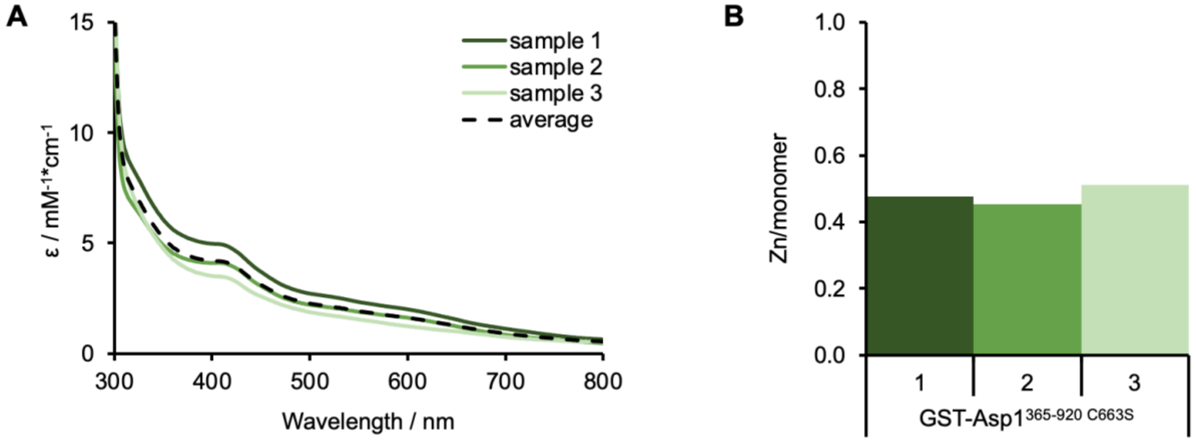


Figure 80: Overview of the electronic absorption spectroscopy and determined zinc content for GST-Asp1$^{365-920\ C663S}$ including biological triplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

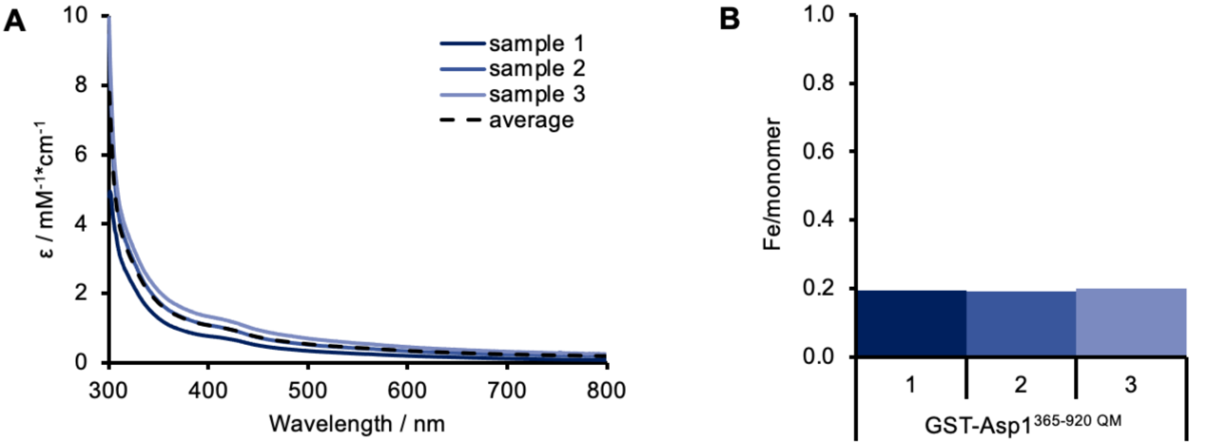


Figure 81: Overview of the electronic absorption spectroscopy and ICP-MS data for GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$ including biological triplicates. A) The extinction coefficient ε in $\mathrm{mM^{-1}{*}cm^{-1}}$ (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological triplicates are shown in a colour gradient.

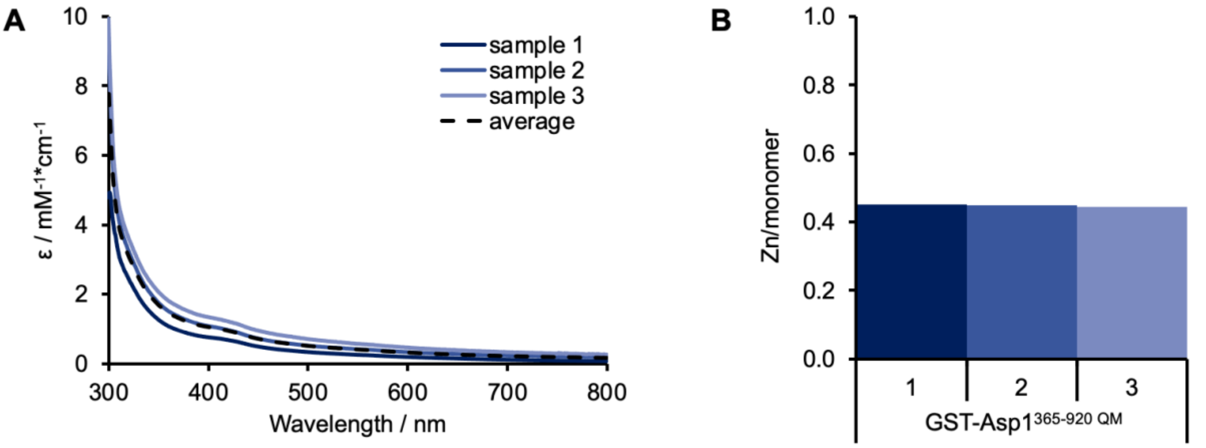


Figure 82: Overview of the electronic absorption spectroscopy and determined zinc content for GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$ including biological triplicates. A) The extinction coefficient ε in $\mathrm{mM^{-1}{*}cm^{-1}}$ (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological triplicates are shown in a colour gradient.

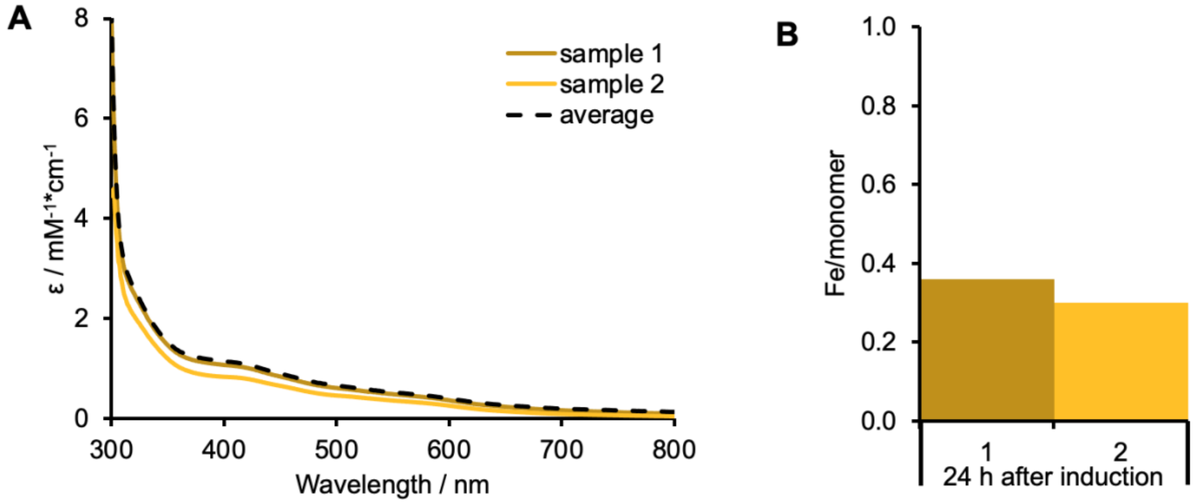


Figure 83: Overview of the electronic absorption spectroscopy and ICP-MS data for as-isolated GST-Asp1$^{365\text{-}920}$ after expression with supplemental iron and sulfur and overnight induction including biological duplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Fe levels among the biological duplicates are shown in a colour gradient.

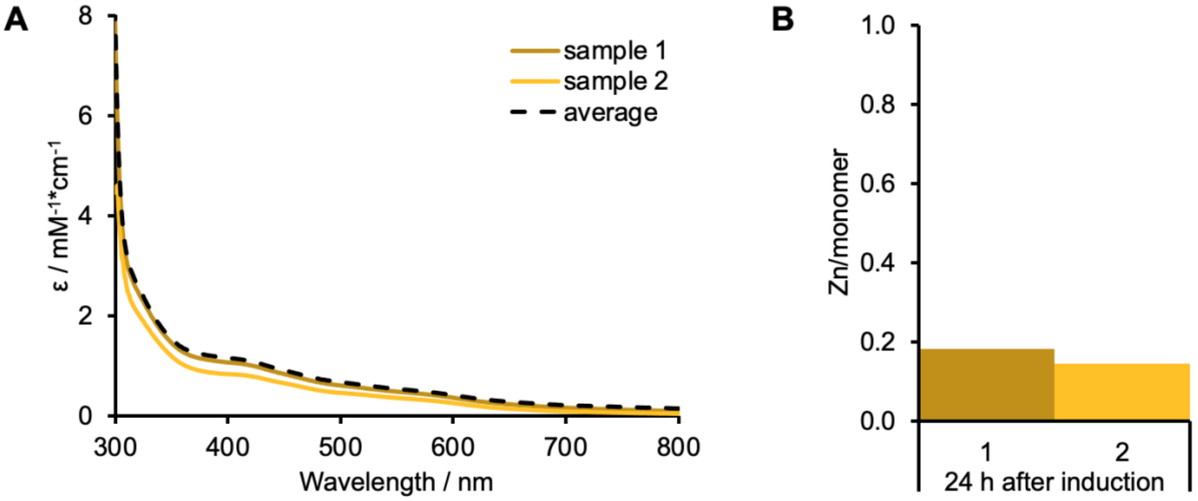


Figure 84: Overview of the electronic absorption spectroscopy and determined zinc content for as-isolated GST-Asp1$^{365\text{-}920}$ after expression with supplemental iron and sulfur and overnight induction including biological duplicates. A) The extinction coefficient ε in mM^{-1}*cm^{-1} (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines. B) The measured Zn levels among the biological duplicates are shown in a colour gradient.

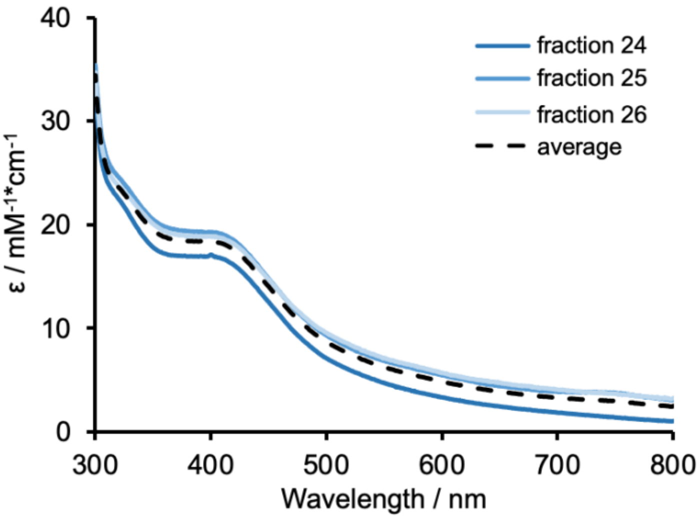


Figure 85: Overview of the electronic absorption spectroscopy data for the GST-Asp1$^{365\text{-}920}$ sample reconstituted with DTT and 10-fold excess of $FeCl_3$. The extinction coefficient ε in $mM^{-1}*cm^{-1}$ (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines.

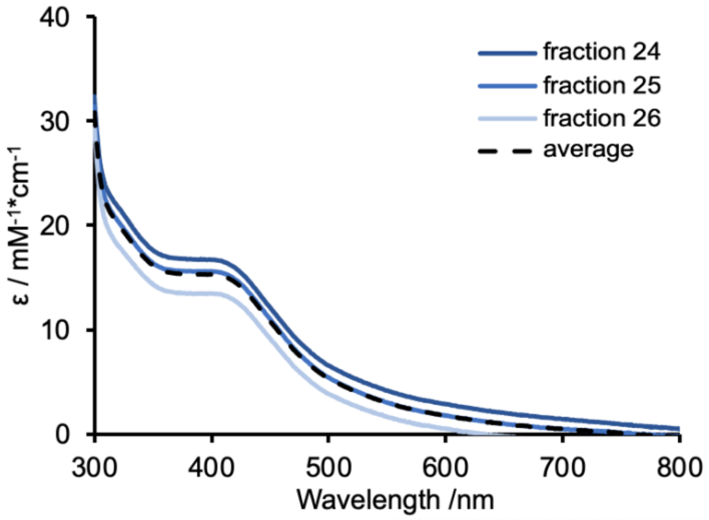


Figure 86: Overview of the electronic absorption spectroscopy data for the GST-Asp1$^{365\text{-}920}$ sample reconstituted with DTT and 10-fold excess of FeCit. The extinction coefficient ε in $mM^{-1}*cm^{-1}$ (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines.

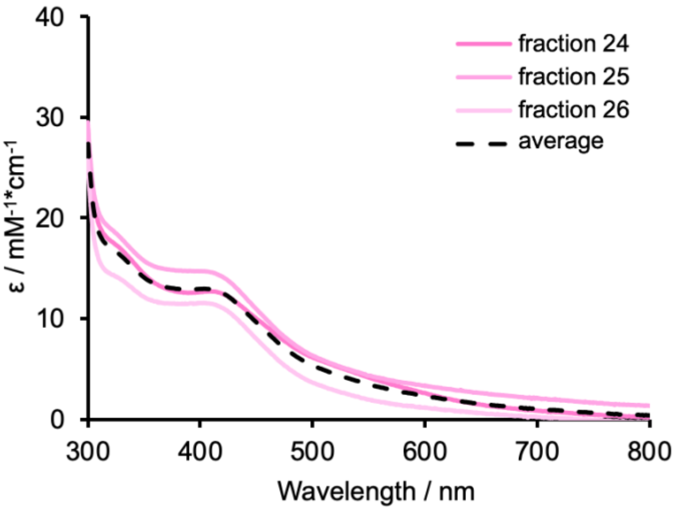


Figure 87: Overview of the electronic absorption spectroscopy data for the GST-Asp1$^{365-920}$ sample reconstituted with GSH and 10-fold excess of $FeCl_3$. The extinction coefficient ε in $mM^{-1}*cm^{-1}$ (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines.

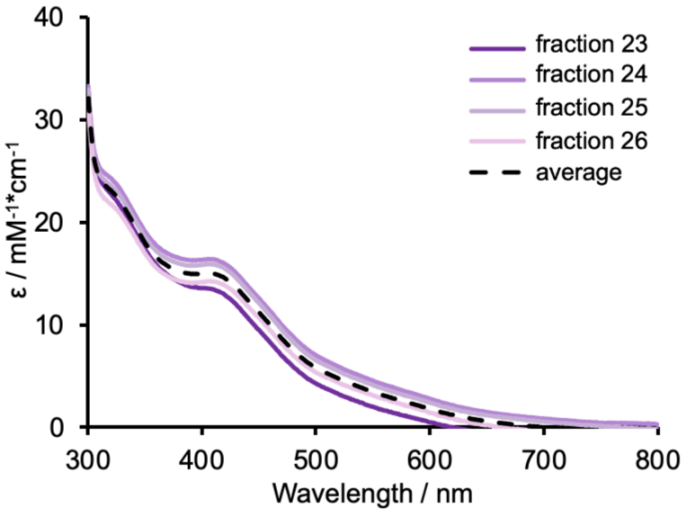


Figure 88: Overview of the electronic absorption spectroscopy data for the GST-Asp1$^{365-920}$ sample reconstituted with GSH and 10-fold excess of FeCit. The extinction coefficient ε in $mM^{-1}*cm^{-1}$ (y-axis) was plotted as a function of the wavelength in nm (x-axis). The three biological triplicates are depicted as a colour gradient, whereas the average curve is illustrated in black interrupted lines.

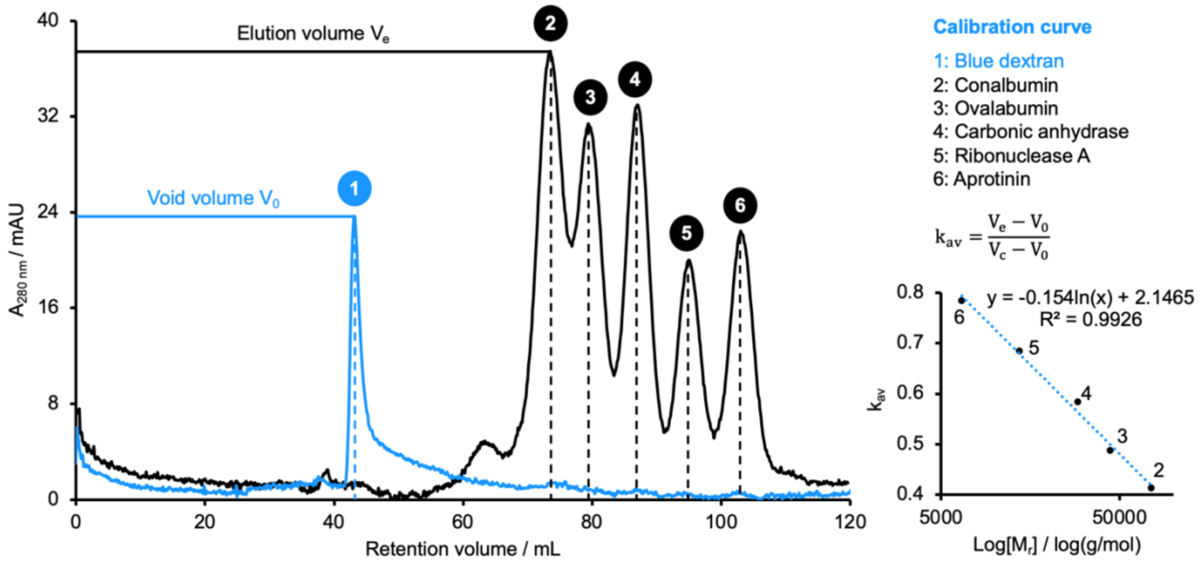


Figure 89: Calibration of the HiLoad 16/600 Superdex 200 pg SEC column using the Cytiva LMW kit. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis.

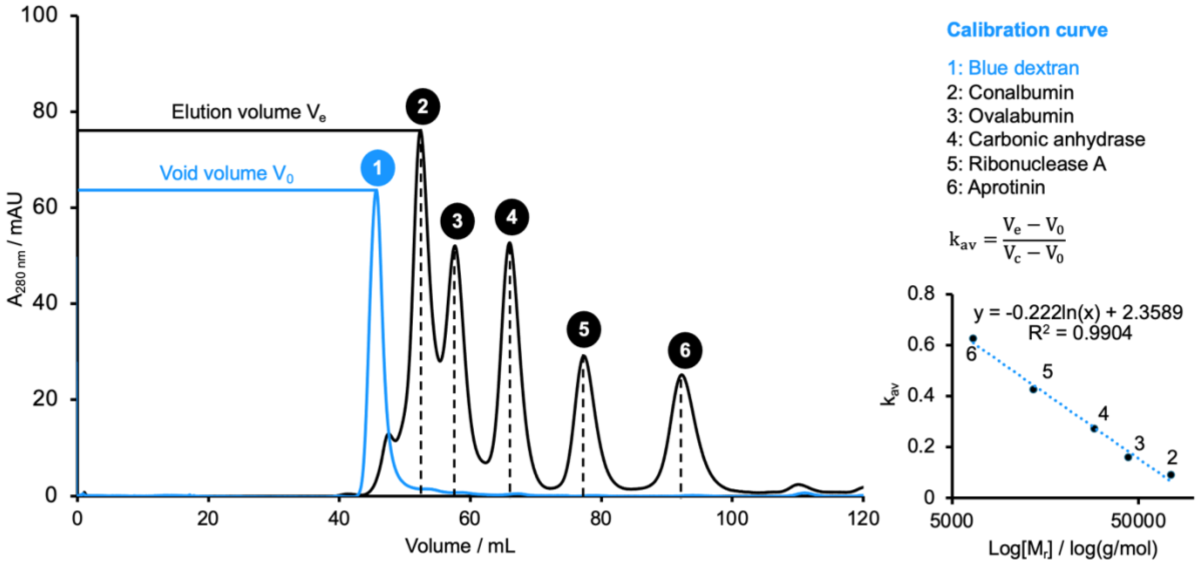


Figure 90: Calibration of the HiLoad 16/600 Superdex 75 pg SEC column using the Cytiva LMW kit. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis.

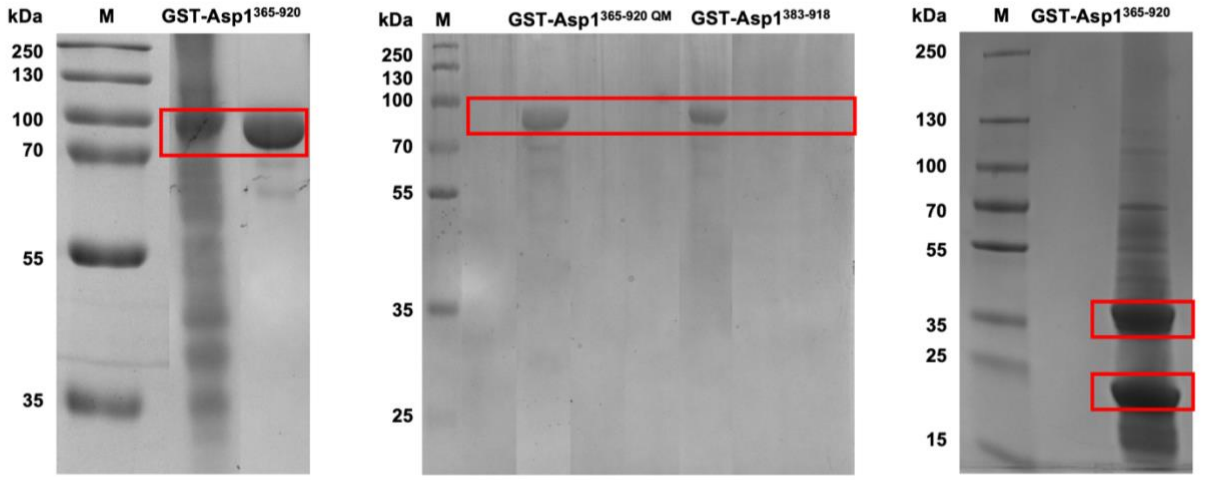


Figure 91: SDS-PAGE showing the purity of the fractions collected during different SEC runs. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker.

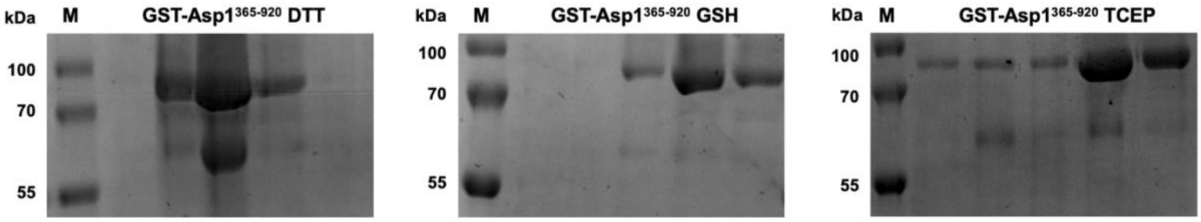


Figure 92: SDS-PAGE showing the protein composition of the peaks recorded in the experimental set-ups with three different reducing agents. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min. Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker.

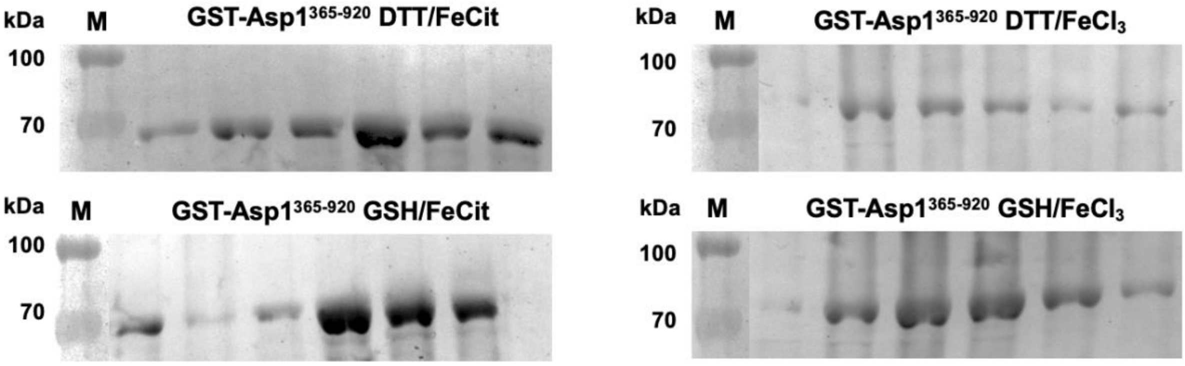
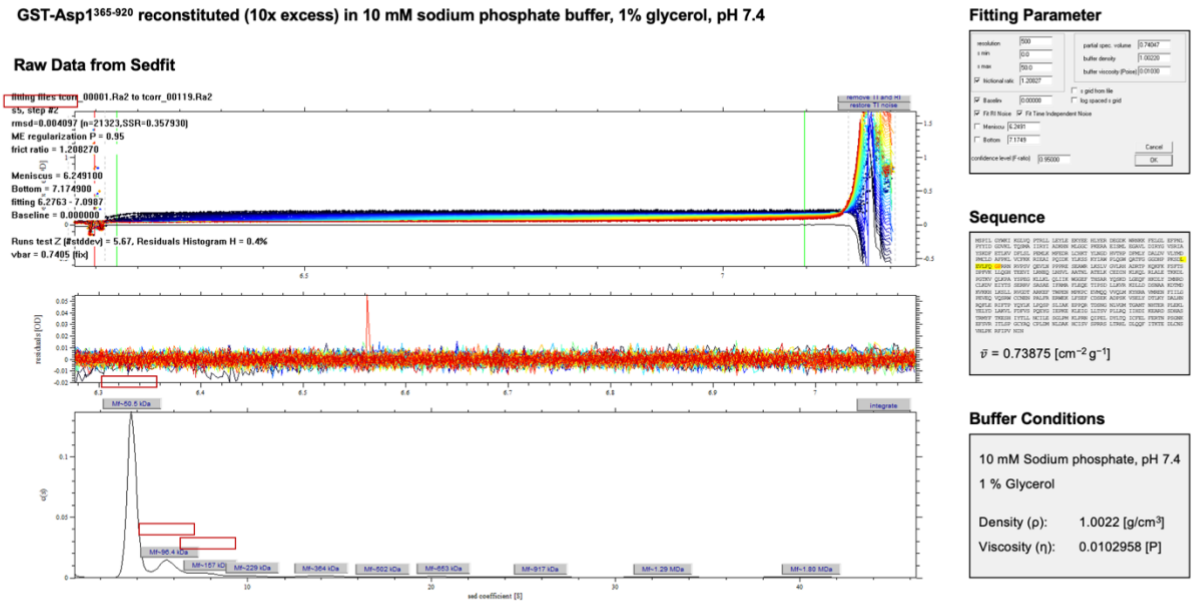


Figure 93: SDS-PAGE showing the content of the collected fractions during desalting of the reconstituted samples with 10-fold excess $FeCl_3$ or FeCit. A 15% Tris-glycine separating gel and a 2% Tris-glycine stacking gel were prepared according to the protocol. A volume of 10 µL sample was loaded on the gel, as well as 3 µL ladder. The gel electrophoresis was conducted at 140 mA for 1.5 h. The 15% gel was stained in prewarmed Coomasie Brilliant Blue G-250 staining solution for 30 min.

Next the page was decolourised in water o/n. The PageRuler™ Plus Prestained Protein Ladder of ThermoScientific was used as marker.



fitting files tcorr_00001.Ra2 to tcorr_00119.Ra2
s5, step #2
rmsd=0.004097 (n=21323,SSR=0.357930)
ME regularization P = 0.95
frict ratio = 1.208270
Meniscus = 6.249100
Bottom = 7.174900
fitting 6.2763 - 7.0987
Baseline = 0.000000
Runs test Z (#stddev) = 5.67, Residuals Histogram H = 0.4%
vbar = 0.7405 (fix)

resolution 500
s min 0.0
s max 50.0
☑ frictional ratio 1.20827
partial spec. volume 0.74047
buffer density 1.00220
buffer viscosity (Poise) 0.01030
☐ s grid from file
☐ log spaced s grid
☑ Baseline 0.00000
☑ Fit RI Noise ☑ Fit Time Independent Noise
☐ Meniscus 6.2491
☐ Bottom 7.1749
confidence level (F-ratio) 0.95000

$\bar{v}$ = 0.73875 [cm^{-2} g^{-1}]

10 mM Sodium phosphate, pH 7.4
1 % Glycerol

Density (ρ): 1.0022 [g/cm³]
Viscosity (η): 0.0102958 [P]

Figure 94: Overview of the sedimentation velocity run of GST-Asp1$^{365-920}$ reconstituted with 10-fold excess of iron and sulfur sources which was fitted using SEDFIT and plotted with GUSSI.

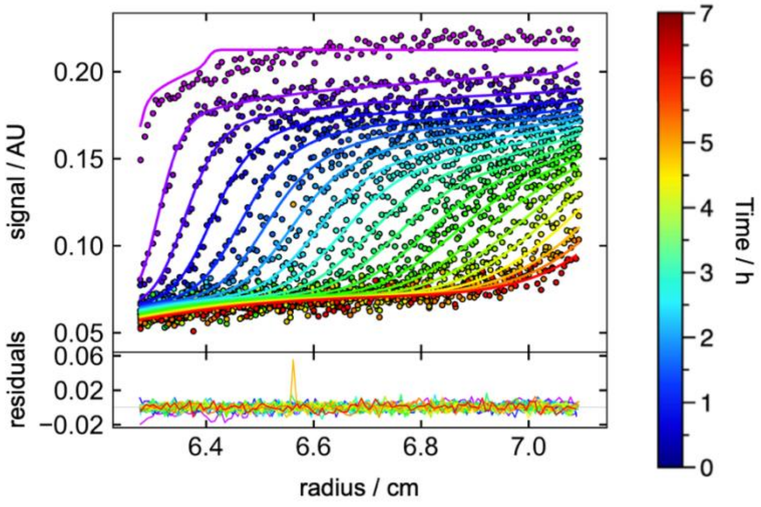


Figure 95: Sedimentation profile of GST-Asp1$^{365-920}$ including the color-coded time scale.

Supplementary table 3: Overview of the retention volumes and corresponding calculated molecular weights of the oligomeric states observed in the SEC chromatograms.

Variant	Oligomeric state	Retention volume	Molecular weight
GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$ apo	high-order oligomers	45.0 mL	650.3 kDa
GST-Asp1$^{383\text{-}918}$ apo	high-order oligomers	46.7 mL	798.2 kDa
GST-Asp1$^{365\text{-}920}$ as-isolated + 10% glycerol	high-order oligomers	46.5 mL	711.1 kDa
	trimer	58.5 mL	264.0 kDa
	dimer	65.4 mL	162.3 kDa
	monomer	75.2 mL	67.5 kDa
	unidentified species	80.0 mL	45.4 kDa
GST-Asp1$^{365\text{-}920}$ as-isolated	high-order oligomers	46.3 mL	711.1 kDa
	trimer	58.4 mL	264.0 kDa
	dimer	65.7 mL	162.3 kDa
	monomer	75.1 mL	67.5 kDa
	unidentified species	79.8 mL	45.4 kDa
GST-Asp1$^{365\text{-}920}$ reconstituted (5x excess)	high-order oligomer	47.5 mL	651.4 kDa
	dimer	62.9 mL	184.4 kDa
GST-Asp1$^{365\text{-}920}$ reconstituted (5x excess) DTT/GSH/TCEP	high-order oligomers	42.4 mL	940.0 kDa
	high-order oligomers	54.7 mL	414.0 kDa
	dimer	64.8 mL	169.6 kDa
	monomer	72.7 mL	85.2 kDa
GST-Asp1$^{365\text{-}920}$ reconstituted (10x excess)	high-order oligomers	41.6 mL	1200.0 kDa
	high-order oligomers	55.2 mL	350 kDa
	monomer	67.9 mL	102.9 kDa

Supplementary table 4: Overview of the detected oligomeric states during the AUC measurement of GST-Asp1$^{365\text{-}920}$ chemically reconstituted with 10-fold molar excess of iron and sulfur.

Variant	Oligomeric state	*s*-value	Molecular weight
GST-Asp1$^{365\text{-}920\ \mathrm{QM}}$ reconstituted (10x excess)	unidentified species	3.9 S	51.3 kDa
	monomer	6.1 S	100 kDa

6.4 Reductive chelation of the Asp1 Fe-S cluster

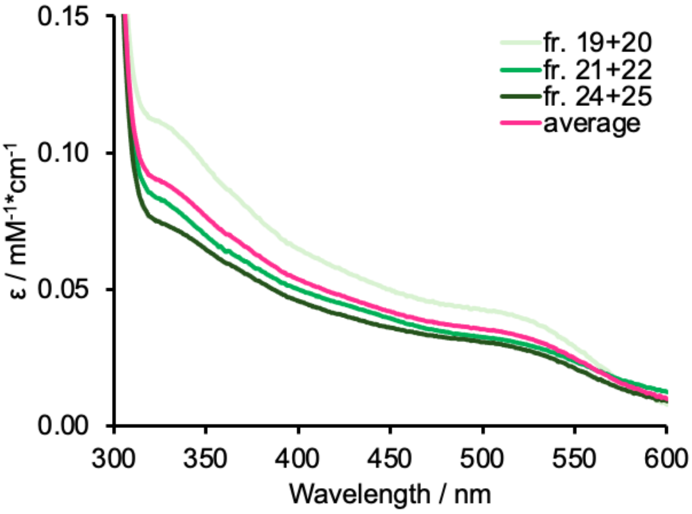


Figure 96: Overview of the recorded EA spectra of every protein-containing fraction from the reductive chelation assay with 2 mM EDTA.

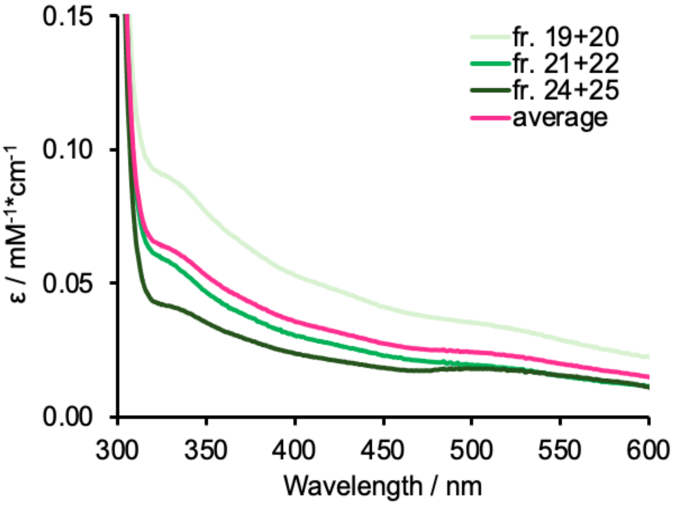


Figure 97: Overview of the recorded EA spectra of every protein containing fraction from the reductive chelation assay with 20 mM EDTA.

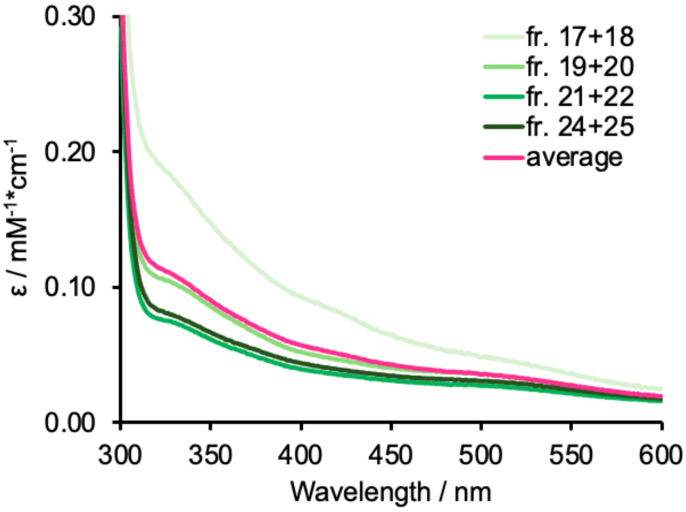


Figure 98: Overview of the recorded EA spectra of every protein containing fraction from the reductive chelation assay with 200 mM EDTA.

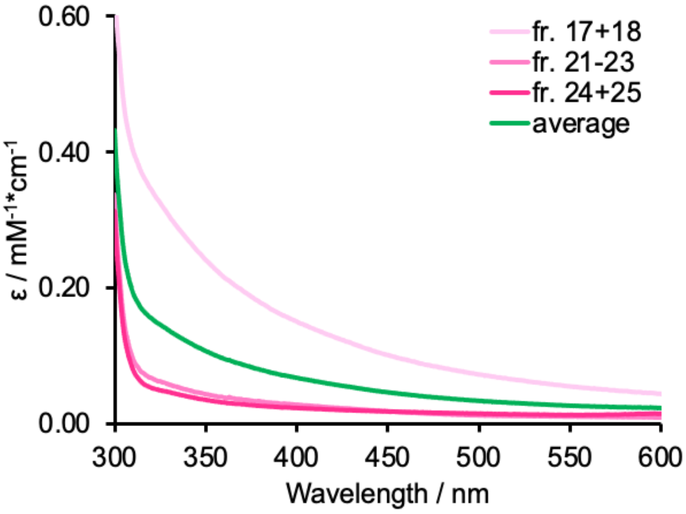


Figure 99: Overview of the recorded EA spectra of every protein containing fraction from the reductive chelation assay with 20 mM BIPY.

Supplementary table 5: Overview of the retention volumes and corresponding calculated molecular weights of each observed oligomeric state.

Assay	Oligomeric state	Retention volume	Molecular weight
2 mM EDTA	high-order oligomer	40 mL	1202.3 kDa
	trimer	59.2 mL	249.3 kDa
	dimer/monomer	67.4 mL	127.3 kDa
	unidentified species	84.0 mL	32.7 kDa
20 mM EDTA	high-order oligomer	48.5 mL	599.2 kDa
	trimer	58.5 mL	264.1 kDa
	dimer/monomer	67.0 mL	131.6 kDa
	unidentified species	87.5 mL	24.5 kDa
200 mM EDTA	high-order oligomer	39.1 mL	639.8 kDa
	trimer	58.4 mL	266.2 kDa
	dimer/monomer	67.4 mL	127.3 kDa
	unidentified species	80.0 mL	45.4 kDa
20 mM BIPY	high-order oligomer	41.2 mL	1089.7 kDa
	trimer	52.6 mL	243.3 kDa
	monomer	69.1 mL	110.8 kDa
	unidentified species	82.1 mL	38.2 kDa

6.5 Analysis of secondary structure of the Asp1$^{365\text{-}920}$ pyrophosphatase domain

Supplementary table 6: Distribution of secondary structure elements among different iron-loaded states of GST-Asp1$^{365\text{-}920}$. The categories helix and antiparallel are divided in subclasses, the percentual amount is depicted per class as well as in total. The terms parallel and antiparallel are associated with β-sheets.

Sample	**Secondary structure elements**		**Amount per element**	**total**
0 Fe/monomer (QM)	Helix	Regular	15.1%	26.2%
		Distorted	11.1%	
	Antiparallel	Left-twisted	1.2%	15.1%
		Relaxed	8.1%	
		Right-twisted	5.8%	
	Parallel	Parallel	8.5%	8.5%
	Turn	Turn	11.3%	11.3%
	Random coil	Random coil	38.9%	38.9%
0 Fe/monomer	Helix	Regular	17.3%	28.8%
		Distorted	11.5%	
	Antiparallel	Left-twisted	0.0%	10.5%
		Relaxed	8.1%	
		Right-twisted	2.4%	
	Parallel	Parallel	8.5%	8.5%
	Turn	Turn	11.6%	11.6%
	Random coil	Random coil	40.7%	40.7%
1 Fe/monomer	Helix	Regular	4.2%	9.5%
		Distorted	5.3%	
	Antiparallel	Left-twisted	0.5%	30.6%
		Relaxed	13.8%	
		Right-twisted	16.3%	
	Parallel	Parallel	0.4%	0.4%
	Turn	Turn	15.0%	15.0%
	Random coil	Random coil	44.4%	44.4%
2 Fe/monomer	Helix	Regular	18.2%	32.2%
		Distorted	14.0%	
	Antiparallel	Left-twisted	0.0%	4.3%
		Relaxed	4.3%	
		Right-twisted	0.0%	
	Parallel	Parallel	7.3%	7.3%
	Turn	Turn	12.4%	12.4%
	Random coil	Random coil	43.8%	43.8%

Sample	Secondary structure elements		Amount	
			per element	total
🟥 4 Fe/monomer	Helix	Regular	8.4%	16.2%
		Distorted	7.8%	
	Antiparallel	Left-twisted	0.8%	20.3%
		Relaxed	12.5%	
		Right-twisted	7.0%	
	Parallel	Parallel	8.4%	8.4%
	Turn	Turn	12.0%	12.0%
	Random coil	Random coil	43.1%	43.1%
🟩 4 Fe/monomer + 10-fold IP_8	Helix	Regular	4.9%	11.7%
		Distorted	6.8%	
	Antiparallel	Left-twisted	0.0%	21.1%
		Relaxed	11.9%	
		Right-twisted	9.2%	
	Parallel	Parallel	8.7%	8.7%
	Turn	Turn	11.8%	11.8%
	Random coil	Random coil	46.8%	46.8%

Supplementary table 7: Distribution of secondary structure elements among different iron-loaded states of His-Asp1$^{365\ 920}$.

Sample	Secondary structure elements		Amount	
			per element	total
⬛ His-Asp1$^{365-920}$ apo	Helix	Regular	20.0%	34.8%
		Distorted	14.8%	
	Antiparallel	Left-twisted	0.0%	7.5%
		Relaxed	1.7%	
		Right-twisted	5.7%	
	Parallel	Parallel	1.6%	1.6%
	Turn	Turn	11.3%	11.3%
	Random coil	Random coil	44.8%	44.8%
🟧 His-Asp1$^{365-920}$ holo	Helix	Regular	25.2%	41.5%
		Distorted	16.4%	
	Antiparallel	Left-twisted	0.0%	3.7%
		Relaxed	3.7%	
		Right-twisted	0.0%	
	Parallel	Parallel	3.4%	3.4%
	Turn	Turn	11.2%	11.2%
	Random coil	Random coil	40.2%	40.2%

Supplementary table 8: Distribution of secondary structure elements among different iron loaded states of His-Asp1$^{383-918}$.

Sample	Secondary structure elements		Amount per element	total
His-Asp1$^{383-918}$ apo	Helix	Regular	18.9%	32.2%
		Distorted	13.4%	
	Antiparallel	Left-twisted	0.0%	11.5%
		Relaxed	3.9%	
		Right-twisted	7.7%	
	Parallel	Parallel	4.3%	4.3%
	Turn	Turn	11.7%	11.7%
	Random coil	Random coil	40.3%	40.3%
His-Asp1$^{383-918}$ holo	Helix	Regular	21.8%	38.0%
		Distorted	16.2%	
	Antiparallel	Left-twisted	0.0%	6.9%
		Relaxed	2.2%	
		Right-twisted	4.6%	
	Parallel	Parallel	2.9%	2.9%
	Turn	Turn	11.2%	11.2%
	Random coil	Random coil	41.0%	41.0%

Supplementary table 9: Distribution of secondary structure elements among different iron-loaded states of His Asp1$^{365-920}$.

Sample	Secondary structure elements		Amount per element	total
Asp1$^{365-920}$ 1 Fe/monomer	Helix	Regular	10.7%	19.8%
		Distorted	9.1%	
	Antiparallel	Left-twisted	1.3%	22.2%
		Relaxed	9.9%	
		Right-twisted	11.0%	
	Parallel	Parallel	4.1%	4.1%
	Turn	Turn	13.3%	13.3%
	Random coil	Random coil	40.6%	40.6%

6.6 Characterization of reconstituted wild-type Asp1$^{365-920}$ by Mössbauer spectroscopy

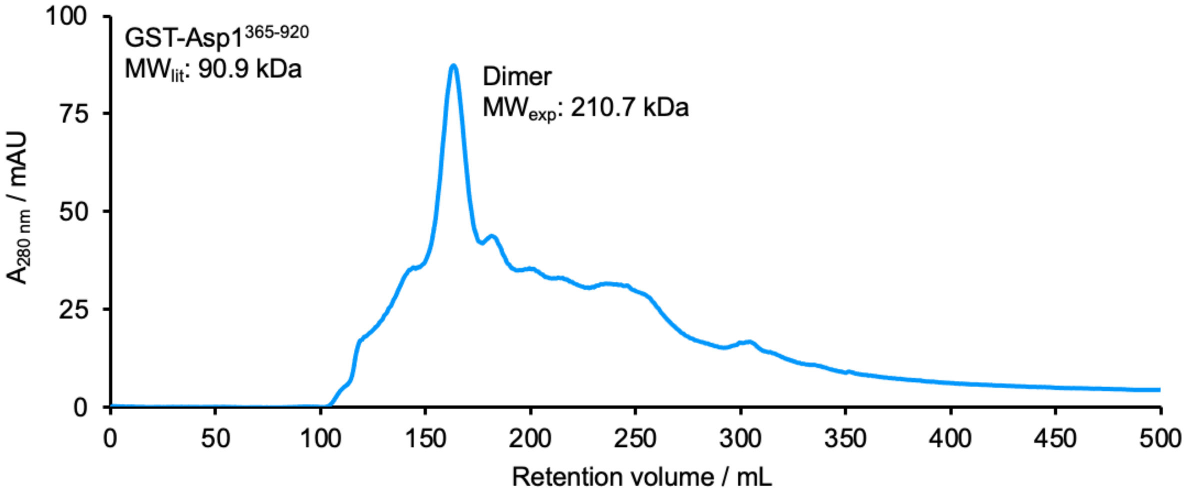


Figure 100: SEC chromatogram from the desalting of the reconstituted and ^{57}Fe-labeled GST-Asp1$^{365-920}$ using a HiLoad 26/600 Superdex 200 pg column. The absorbance at 280 nm in mAU was recorded during the chromatography and plotted as a function of the volume on the primary y-axis.

6.7 Spectroscopic characterization of Asp1$^{365\text{-}920}$ mutant variants

Supplementary table 10: Overview of the optimized purification conditions used for the anaerobic isolation of the two pyrophosphatase mutant variants.

Condition	Buffer composition	Tested mutant variants
1	1x PBS, 10% glycerol, 1 mM DTT, pH 7.3	GST-Asp1$^{383\text{-}918\ C868S}$ GST-Asp1$^{383\text{-}918\ C643S\ C644S}$
2	1x PBS, 20% glycerol, 1 mM DTT, pH 7.3	GST-Asp1$^{383\text{-}918\ C868S}$ GST-Asp1$^{383\text{-}918\ C643S\ C644S}$
3	1x PBS, 20% glycerol, 1 M urea, 1 mM DTT, pH 7.3	GST-Asp1$^{383\text{-}918\ C868S}$ GST-Asp1$^{383\text{-}918\ C643S\ C644S}$
4	50 mM Tris-HCl, 500 mM NaCl, 10% glycerol, 1 mM DTT, pH 8.0	GST-Asp1$^{383\text{-}918\ C868S}$ GST-Asp1$^{383\text{-}918\ C643S\ C644S}$
5	50 mM Tris-HCl, 500 mM NaCl, 20% glycerol, 1 mM DTT, pH 8.0	GST-Asp1$^{383\text{-}918\ C868S}$ GST-Asp1$^{383\text{-}918\ C643S\ C644S}$
6	50 mM Tris-HCl, 500 mM NaCl, 20% glycerol, 1 M urea, 1 mM DTT, pH 8.0	GST-Asp1$^{383\text{-}918\ C868S}$ GST-Asp1$^{383\text{-}918\ C643S\ C644S}$

6.8 Crystallization of Asp1$^{365\text{-}920}$ and Asp1$^{383\text{-}918}$ variants

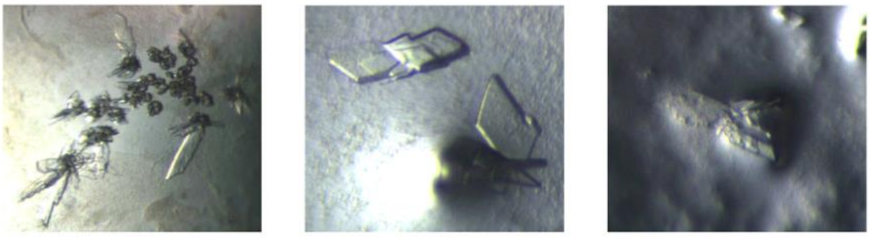
Supplementary table 11: Overview of the chosen crystallization conditions. The colour of the protein variant is indicative for the different cofactor-bound states: grey = 0 Fe/monomer (apo), yellow = 1 Fe/monomer (as-isolated), orange = 2 Fe/monomer (reconstituted with 5-fold excess), red = 4 Fe/monomer (reconstituted with 10-fold excess).

Protein variant	Concentration	Buffer composition	Sparse matrix kit
GST-Asp1$^{365\text{-}920}$	7 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	7 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	MemGold 1-2
GST-Asp1$^{365\text{-}920}$	5 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	10 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	20 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	30 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	5 mg/mL	100 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	10 mg/mL	100 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	20 mg/mL	100 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	30 mg/mL	100 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	5 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	10 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	20 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	30 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	5 mg/mL	10 mM NaPi, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365\text{-}920}$	10 mg/mL	10 mM NaPi, 1 mM DTT, pH 7.4	JBScreen Classic 1-8

Protein variant	**Concentration**	**Buffer composition**	**Sparse matrix kit**
GST-Asp1$^{365-920}$	10 mg/mL	50 mM Tris-HCl, 100 mM NaCl, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	5 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	10 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	20 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	30 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	5 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	10 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	20 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	30 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	3.5 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
Asp1$^{365-920}$	5 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
Asp1$^{365-920}$	10 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
Asp1$^{365-920}$	5 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
Asp1$^{365-920}$	10 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
Asp1$^{365-920}$	2 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
Asp1$^{365-920}$	4 mg/mL	10 mM NaPi, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	5 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	10 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	5 mg/mL	50 mM Tris-HCl, 100 mM NaCl, pH 7.4	JBScreen Classic 1-8

Protein variant	**Concentration**	**Buffer composition**	**Sparse matrix kit**
GST-Asp1$^{383-918}$	20 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	30 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	5 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	10 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	20 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	30 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	5 mg/mL	10 mM NaPi, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	10 mg/mL	10 mM NaPi, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	20 mg/mL	10 mM NaPi, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{383-918}$	30 mg/mL	10 mM NaPi, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
Asp1$^{383-918}$	2 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
Asp1$^{383-918}$	3.7 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	5 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	10 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	20 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	30 mg/mL	10 mM NaPi, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	5 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	10 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	20 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
GST-Asp1$^{365-920}$	30 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8

Protein variant	**Concentration**	**Buffer composition**	**Sparse matrix kit**
Asp1$^{365-920}$	2.8 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
His-Asp1$^{365-920}$	5 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
His-Asp1$^{365-920}$	10 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
His-Asp1$^{365-920}$	20 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
His-Asp1$^{365-920}$	30 mg/mL	50 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 8.0	JBScreen Classic 1-8
His-Asp1$^{383-918}$	5 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
His-Asp1$^{383-918}$	10 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
His-Asp1$^{383-918}$	20 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
His-Asp1$^{383-918}$	30 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
His-Asp1$^{383-918}$	5 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
His-Asp1$^{383-918}$	10 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
His-Asp1$^{383-918}$	20 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8
His-Asp1$^{383-918}$	30 mg/mL	10 mM NaPi, 10% glycerol, 1 mM DTT, pH 7.4	JBScreen Classic 1-8



A8 GST-PP$^{365-920}$ 5 mg/mL
Buffer: 100 mM NaPi, 10% glycerol, pH 7.4
30% PEG 400, 100 mM sodium acetate, pH 4.6, 100 mM calcium chloride

E3 GST-PP$^{365-920}$ 5 mg/mL
Buffer: 100 mM NaPi, 10% glycerol, pH 7.4
15% w/v 2-Methyl-2,4-pentanediol, 100 mM sodium acetate, pH 4.6, 20 mM calcium chloride

E11 GST-PP$^{365-920}$ 5 mg/mL
Buffer: 100 mM NaPi, 10% glycerol, pH 7.4
30% w/v 2-Methyl-2,4-pentanediol, 100 mM sodium acetate, pH 4.6, 20 mM calcium chloride

Figure 101: Overview of the inorganic phosphate crystals obtained during crystallization attempts of GST-Asp1$^{365-920}$.

7 Acknowledgements

I would like to thank first and foremost Prof. Dr. Ingrid Span for the possibility of working on such a captivating and highly challenging project during my doctoral studies. This has been a very instructive time, and I am grateful for the many opportunities like conferences and group seminars allowing me to exit my academic comfort zone and polish my interdisciplinary skills. Also, I wish to give my thanks for the warm welcome in the group and support throughout this period even if from afar via videocalls. Next, I would like to thank apl.-Prof. Dr. Ursula Fleig for allowing me to join the Asp1 team. Her passion for this topic extended on to me and motivated me to make the most out of this project and go beyond my limits, physically as well as mentally. I appreciate her efforts to train her doctoral students to elucidate the mystery behind the Asp1 protein of *S. pombe*. Moreover, I am grateful for welcoming me in the Fleig group and for many fruitful discussions. I would like to thank Prof. Dr. Dieter Willbold for providing the laboratories and consumables and for allowing me to complete my doctoral studies in the IPB.

Furthermore, I wish to extend my thanks to one of our collaborators Dr. James Birrell for the possibility of a research internship in his group at the Max-Planck-Institute for Chemical Energy Conversion and for the great support in the preparation of the Mössbauer samples. I could learn a lot from our discussions on Mössbauer theory and iron-sulfur proteins and came back to the IPB more skilled in working at the glovebox. Next, I would like to especially thank two postdocs that shared their expert knowledge in their respective research areas with me along the years: Dr. Hamed Shaykhalishahi from the IPB and Dr. Abel Alcazar-Roman from the IFGR. Additionally, I wish to thank my numerous and excellent collaboration partners, without whom this work would have not been possible: Dr. Sabine Metzger for ICP-MS measurents, Eva Walla for assistance during biological experiments and Prof. Dr. Dorothea Fliedler for providing substrates for the Asp1 pyrophosphatase domain.

Additionally, I would like to thank my colleagues from the two groups (IPB: Nina Kirchgässler, Melissa Jansing, Konstantin Bikbaev; IFGR: Lisa Juhran, Sarune Daunaraviciute) for the nice work atmosphere and being supportive in the lab. A special thank you goes to Lora Denson, Katharina Vormann, Dr. Najoua Boulakhrif and Dr. Thomas Pauly, who went from being my lab buddies to some of my closest friends. Thank you for always cheering me up and supporting me. Life in the lab is so much more enjoyable with you around. I am very happy to be able to call them friends by now as well. I wish to thank my great office mates Luis and Tina for always having a laugh with me. A great thank you goes out to Robin, our technical staff, that has always prioritized my work during my most stressful times and accompanies me to our dance classes. I would like to extend some of my deepest thanks to Florian Tucholski, for showing me the ups and downs of life and for leaving and resurfacing in my life always at the wrong time.

On another note, I wish to thank all my friends either from school (TFG: Özgül, Iza, Sepehr, Katja, Backtash, Tanja) and university (Helena, Stella, Nadiia) for the great mental support and for sticking with me for so long. In time, these connections turned so special that I cannot picture life without you anymore. Finally, I would like to truly thank my parents Anca and Dorin and my recently departed furry companion Dalooa for the undying and unconditional love, for their efforts, their sacrifices, and their support throughout the years. Achieving this would have not been possible without you. Thank you for accepting me the way I am and continuously being there for me through all the good and bad. Words cannot express the love I have for you. I hope that being the first doctor in our family brings joy to you and other departed family.

8 Bibliography

1. Feoktistova, A., McCollum, D., Ohi, R. & Gould, K. L. Identification and characterization of Schizosaccharomyces pombe asp1+, a gene that interacts with mutations in the Arp2/3 complex and actin. *Genetics* **152**, 895–908 (1999).
2. Pöhlmann, J. & Fleig, U. Asp1, a conserved 1/3 inositol polyphosphate kinase, regulates the dimorphic switch in Schizosaccharomyces pombe. *Mol Cell Biol* (2010).
3. Rosenbach, H. *et al.* The Asp1 pyrophosphatase from S. pombe hosts a [2Fe-2S] 2+ cluster in vivo. *JBIC Journal of Biological Inorganic Chemistry* **26**, 93–108 (2021).
4. Benjamin, B. *et al.* Activities and structure-function analysis of fission yeast inositol pyrophosphate (IPP) kinase-pyrophosphatase Asp1 and its impact on regulation of pho1 gene expression. *mBio* **13**, e01034-22 (2022).
5. Wang, H. *et al.* Asp1 from Schizosaccharomyces pombe binds a [2Fe-2S] 2+ cluster which inhibits inositol pyrophosphate 1-phosphatase activity. *Biochemistry* **54**, 6462–6474 (2015).
6. Mulugu, S. *et al.* A conserved family of enzymes that phosphorylate inositol hexakisphosphate. *Science (1979)* **316**, 106–109 (2007).
7. Lin, H. *et al.* Structural analysis and detection of biological inositol pyrophosphates reveal that the family of VIP/diphosphoinositol pentakisphosphate kinases are 1/3-kinases. *Journal of Biological Chemistry* **284**, 1863–1872 (2009).
8. Steidle, E. A. *et al.* A novel inositol pyrophosphate phosphatase in Saccharomyces cerevisiae: Siw14 protein selectively cleaves the β-phosphate from 5-diphosphoinositol pentakisphosphate (5PP-IP5). *Journal of Biological Chemistry* **291**, 6772–6783 (2016).
9. Weaver, J. D., Wang, H. & Shears, S. B. The kinetic properties of a human PPIP5K reveal that its kinase activities are protected against the consequences of a deteriorating cellular bioenergetic environment. *Biosci Rep* **33**, e00022 (2013).
10. Thomas, M. P. & Potter, B. V. L. The enzymes of human diphosphoinositol polyphosphate metabolism. *FEBS J* **281**, 14–33 (2014).
11. Fridy, P. C., Otto, J. C., Dollins, D. E. & York, J. D. Cloning and characterization of two human VIP1-like inositol hexakisphosphate and diphosphoinositol pentakisphosphate kinases. *Journal of Biological Chemistry* **282**, 30754–30762 (2007).
12. Laha, D. *et al.* VIH2 regulates the synthesis of inositol pyrophosphate InsP8 and jasmonate-dependent defenses in Arabidopsis. *Plant Cell* **27**, 1082–1097 (2015).
13. Williams, S. P., Gillaspy, G. E. & Perera, I. Y. Biosynthesis and possible functions of inositol pyrophosphates in plants. *Front Plant Sci* **6**, 67 (2015).
14. Irvine, R. F. & Schell, M. J. Back in the water: the return of the inositol phosphates. *Nat Rev Mol Cell Biol* **2**, 327–338 (2001).

15. Pattni, K. & Banting, G. Ins (1, 4, 5) P3 metabolism and the family of IP3-3Kinases. *Cell Signal* **16**, 643–654 (2004).

16. Berridge, M. J. Inositol trisphosphate and calcium signalling. *Nature* **361**, 315–325 (1993).

17. Raboy, V. & Bowen, D. Genetics of inositol polyphosphates. *Biology of Inositols and Phosphoinositides: Subcellular Biochemistry* 71–101 (2006).

18. Leyman, A. *et al.* The absence of expression of the three isoenzymes of the inositol 1, 4, 5-trisphosphate 3-kinase does not prevent the formation of inositol pentakisphosphate and hexakisphosphate in mouse embryonic fibroblasts. *Cell Signal* **19**, 1497–1504 (2007).

19. Odom, A. R., Stahlberg, A., Wente, S. R. & York, J. D. A role for nuclear inositol 1, 4, 5-trisphosphate kinase in transcriptional control. *Science (1979)* **287**, 2026–2029 (2000).

20. Saiardi, A. *et al.* Mammalian inositol polyphosphate multikinase synthesizes inositol 1, 4, 5-trisphosphate and an inositol pyrophosphate. *Proceedings of the National Academy of Sciences* **98**, 2306–2311 (2001).

21. Gunesekera, B., Torabinejad, J., Robinson, J. & Gillaspy, G. E. Inositol polyphosphate 5-phosphatases 1 and 2 are required for regulating seedling growth. *Plant Physiol* **143**, 1408–1417 (2007).

22. Ivetac, I. *et al.* The type Iα inositol polyphosphate 4-phosphatase generates and terminates phosphoinositide 3-kinase signals on endosomes and the plasma membrane. *Mol Biol Cell* **16**, 2218–2233 (2005).

23. York, J. D., Odom, A. R., Murphy, R., Ives, E. B. & Wente, S. R. A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. *Science (1979)* **285**, 96–100 (1999).

24. Shears, S. B., Baughman, B. M., Gu, C., Nair, V. S. & Wang, H. The significance of the 1-kinase/1-phosphatase activities of the PPIP5K family. *Adv Biol Regul* **63**, 98–106 (2017).

25. Shears, S. B. *et al.* Defining signal transduction by inositol phosphates. *Phosphoinositides II: The Diverse Biological Functions* 389–412 (2012).

26. Shears, S. B., Gokhale, N. A., Wang, H. & Zaremba, A. Diphosphoinositol polyphosphates: what are the mechanisms? *Adv Enzyme Regul* **51**, 13–25 (2011).

27. Safrany, S. T. *et al.* A novel context for the 'MutT'module, a guardian of cell integrity, in a diphosphoinositol polyphosphate phosphohydrolase. *EMBO J* (1998).

28. Caffrey, J. J., Safrany, S. T., Yang, X. & Shears, S. B. Discovery of molecular and catalytic diversity among human diphosphoinositol-polyphosphate phosphohydrolases: an expanding NUDT family. *Journal of Biological Chemistry* **275**, 12730–12736 (2000).

29. Luo, H. R. *et al.* Inositol pyrophosphates mediate chemotaxis in Dictyostelium via pleckstrin homology domain-PtdIns (3, 4, 5) P3 interactions. *Cell* **114**, 559–572 (2003).

30. Morrison, B. H., Bauer, J. A., Kalvakolanu, D. V & Lindner, D. J. Inositol hexakisphosphate kinase 2 mediates growth suppressive and apoptotic effects of interferon-β in ovarian carcinoma cells. *Journal of Biological Chemistry* **276**, 24965–24970 (2001).
31. Saiardi, A., Sciambi, C., McCaffery, J. M., Wendland, B. & Snyder, S. H. Inositol pyrophosphates regulate endocytic trafficking. *Proceedings of the National Academy of Sciences* **99**, 14206–14211 (2002).
32. Saiardi, A. Protein pyrophosphorylation: moving forward. *Biochemical Journal* **473**, 3765–3768 (2016).
33. Chatree, S., Thongmaen, N., Tantivejkul, K., Sitticharoon, C. & Vucenik, I. Role of inositols and inositol phosphates in energy metabolism. *Molecules* **25**, 5079 (2020).
34. Shah, A., Ganguli, S., Sen, J. & Bhandari, R. Inositol pyrophosphates: energetic, omnipresent and versatile signalling molecules. *J Indian Inst Sci* **97**, 23–40 (2017).
35. Westheimer, F. H. Why nature chose phosphates. *Science (1979)* **235**, 1173–1178 (1987).
36. Topolski, B., Jakopec, V., Künzel, N. A. & Fleig, U. Inositol pyrophosphate kinase Asp1 modulates chromosome segregation fidelity and spindle function in Schizosaccharomyces pombe. *Mol Cell Biol* (2016).
37. Laurent, F. *et al.* Inositol pyrophosphate catabolism by three families of phosphatases regulates plant growth and development. *PLoS Genet* **20**, e1011468 (2024).
38. Benjamin, B. *et al.* Structures of fission yeast inositol pyrophosphate kinase Asp1 in ligand-free, substrate-bound, and product-bound states. *mBio* **13**, e03087-22 (2022).
39. York, J. D. & Lew, D. J. IP7 guards the CDK gate. *Nat Chem Biol* **4**, 16–17 (2008).
40. Desmarini, D. *et al.* IP7-SPX domain interaction controls fungal virulence by stabilizing phosphate signaling machinery. *mBio* **11**, 10–1128 (2020).
41. Chabert, V. *et al.* Inositol pyrophosphate dynamics in yeast reveals control of the PHO starvation program through 1, 5-IP8 and the SPX domain of the CDK inhibitor Pho81. *bioRxiv* (2023).
42. Lonetti, A. *et al.* Identification of an evolutionarily conserved family of inorganic polyphosphate endopolyphosphatases. *Journal of Biological Chemistry* **286**, 31966–31974 (2011).
43. Gu, C. *et al.* The significance of the bifunctional kinase/phosphatase activities of diphosphoinositol pentakisphosphate kinases (PPIP5Ks) for coupling inositol pyrophosphate cell signaling to cellular phosphate homeostasis. *Journal of Biological Chemistry* **292**, 4544–4555 (2017).
44. Randall, T. A., Gu, C., Li, X., Wang, H. & Shears, S. B. A two-way switch for inositol pyrophosphate signaling: Evolutionary history and biological significance of a unique, bifunctional kinase/phosphatase. *Adv Biol Regul* **75**, 100674 (2020).
45. Rigden, D. J. The histidine phosphatase superfamily: structure and function. *Biochemical Journal* **409**, 333–348 (2008).

46. Gokhale, N. A., Zaremba, A. & Shears, S. B. Receptor-dependent compartmentalization of PPIP5K1, a kinase with a cryptic polyphosphoinositide binding domain. *Biochemical Journal* **434**, 415–426 (2011).

47. Pöhlmann, J. *et al.* The Vip1 inositol polyphosphate kinase family regulates polarized growth and modulates the microtubule cytoskeleton in fungi. *PLoS Genet* **10**, e1004586 (2014).

48. Isbrandt, L. R. & Oertel, R. P. Conformational states of myo-inositol hexakis (phosphate) in aqueous solution. A carbon-13 NMR, phosphorus-31 NMR, and Raman spectroscopic investigation. *J Am Chem Soc* **102**, 3144–3148 (1980).

49. Hager, A. *et al.* Cellular cations control conformational switching of inositol pyrophosphate analogues. *Chemistry–A European Journal* **22**, 12406–12414 (2016).

50. Kurz, L., Schmieder, P., Veiga, N. & Fiedler, D. One scaffold, two conformations: the ring-flip of the messenger InsP8 occurs under cytosolic conditions. *Biomolecules* **13**, 645 (2023).

51. Müller, O. *et al.* The Vtc proteins in vacuole fusion: coupling NSF activity to V0 trans-complex formation. *EMBO J* (2002).

52. Cohen, A., Perzov, N., Nelson, H. & Nelson, N. A novel family of yeast chaperons involved in the distribution of V-ATPase and other membrane proteins. *Journal of Biological Chemistry* **274**, 26885–26893 (1999).

53. Hothorn, M. *et al.* Catalytic core of a membrane-associated eukaryotic polyphosphate polymerase. *Science (1979)* **324**, 513–516 (2009).

54. Wild, R. *et al.* Control of eukaryotic phosphate homeostasis by inositol polyphosphate sensor domains. *Science (1979)* **352**, 986–990 (2016).

55. Gerasimaite, R. *et al.* Inositol pyrophosphate specificity of the SPX-dependent polyphosphate polymerase VTC. *ACS Chem Biol* **12**, 648–653 (2017).

56. Zhu, Q., Yaggi, M. F., Jork, N., Jessen, H. J. & Diver, M. M. Transport and InsP8 gating mechanisms of the human inorganic phosphate exporter XPR1. *Nat Commun* **16**, 2770 (2025).

57. Liu, W. *et al.* Cryo-EM structure of the polyphosphate polymerase VTC reveals coupling of polymer synthesis to membrane transit. *EMBO J* **42**, e113320 (2023).

58. Pascual-Ortiz, M., Walla, E., Fleig, U. & Saiardi, A. The PPIP5K family member Asp1 controls inorganic polyphosphate metabolism in S. pombe. *Journal of Fungi* **7**, 626 (2021).

59. Dong, J. *et al.* Inositol pyrophosphate InsP8 acts as an intracellular phosphate signal in Arabidopsis. *Mol Plant* **12**, 1463–1473 (2019).

60. Ried, M. K. *et al.* Inositol pyrophosphates promote the interaction of SPX domains with the coiled-coil motif of PHR transcription factors to regulate plant phosphate homeostasis. *Nat Commun* **12**, 384 (2021).

61. Zhou, J. *et al.* Mechanism of phosphate sensing and signaling revealed by rice SPX1-PHR2 complex structure. *Nat Commun* **12**, 7040 (2021).

62. Desmarini, D. *et al.* IP7-SPX domain interaction controls fungal virulence by stabilizing phosphate signaling machinery. *mBio* **11**, 10–1128 (2020).

63. Chabert, V. *et al.* Inositol pyrophosphate dynamics in yeast reveals control of the PHO starvation program through 1, 5-IP8 and the SPX domain of the CDK inhibitor Pho81. *bioRxiv* (2023).

64. Lee, Y.-S., Huang, K., Quiocho, F. A. & O'Shea, E. K. Molecular basis of cyclin-CDK-CKI regulation by reversible binding of an inositol pyrophosphate. *Nat Chem Biol* **4**, 25–32 (2008).

65. Lonetti, A. *et al.* Identification of an evolutionarily conserved family of inorganic polyphosphate endopolyphosphatases. *Journal of Biological Chemistry* **286**, 31966–31974 (2011).

66. Saiardi, A. *et al.* Inositol pyrophosphate: physiologic phosphorylation of proteins. *Science (1979)* **306**, 2101–2105 (2004).

67. Bhandari, R. *et al.* Protein pyrophosphorylation by inositol pyrophosphates is a posttranslational event. *Proceedings of the National Academy of Sciences* **104**, 15305–15310 (2007).

68. Morgan, J. A. M. *et al.* Pyrophosphoproteomics: extensive protein pyrophosphorylation revealed in human cell lines. *bioRxiv* 2011–2022 (2022).

69. Ganguli, S. *et al.* A high energy phosphate jump-From pyrophospho-inositol to pyrophospho-serine. *Adv Biol Regul* **75**, 100662 (2020).

70. Richter, A. *et al.* Proteome-wide quantification of inositol pyrophosphate-protein interactions. *bioRxiv* 2024–2027 (2024).

71. Saiardi, A., Resnick, A. C., Snowman, A. M., Wendland, B. & Snyder, S. H. Inositol pyrophosphates regulate cell death and telomere length through phosphoinositide 3-kinase-related protein kinases. *Proceedings of the National Academy of Sciences* **102**, 1911–1914 (2005).

72. Bhandari, R. *et al.* Protein pyrophosphorylation by inositol pyrophosphates is a posttranslational event. *Proceedings of the National Academy of Sciences* **104**, 15305–15310 (2007).

73. Stephens, L. *et al.* The detection, purification, structural characterization, and metabolism of diphosphoinositol pentakisphosphate (s) and bisdiphosphoinositol tetrakisphosphate (s). *Journal of Biological Chemistry* **268**, 4009–4015 (1993).

74. Voglmaier, S. M. *et al.* Purified inositol hexakisphosphate kinase is an ATP synthase: diphosphoinositol pentakisphosphate as a high-energy phosphate donor. *Proceedings of the National Academy of Sciences* **93**, 4305–4310 (1996).

75. LAUSSMANN, T., EUJEN, R., Michael WEISSHUHN, C., THIEL, U. & VOGEL, G. Structures of diphospho-myo-inositol pentakisphosphate and bisdiphospho-myo-inositol tetrakisphosphate from Dictyostelium resolved by NMR analysis. *Biochemical Journal* **315**, 715–720 (1996).

76. Morgan, J. A. M. *et al.* Extensive protein pyrophosphorylation revealed in human cell lines. *Nat Chem Biol* 1–12 (2024).

77. Kamiguchi, H. The role of cell adhesion molecules in axon growth and guidance. *Axon growth and guidance* 95–102 (2007).

78. Rohrschneider, M. R. & Nance, J. Polarity and cell fate specification in the control of Caenorhabditis elegans gastrulation. *Developmental Dynamics* **238**, 789–796 (2009).

79. Wang, Y. CDKs and the yeast-hyphal decision. *Curr Opin Microbiol* **12**, 644–649 (2009).

80. Smukalla, S. *et al.* FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* **135**, 726–737 (2008).

81. Pascual-Ortiz, M. *et al.* Asp1 bifunctional activity modulates spindle function via controlling cellular inositol pyrophosphate levels in Schizosaccharomyces pombe. *Mol Cell Biol* **38**, e00047-18 (2018).

82. Kuenzel, N. A. *et al.* Inositol pyrophosphate-controlled kinetochore architecture and mitotic entry in S. pombe. *Journal of Fungi* **8**, 933 (2022).

83. Sanchez, A. M., Garg, A., Shuman, S. & Schwer, B. Inositol pyrophosphates impact phosphate homeostasis via modulation of RNA 3′ processing and transcription termination. *Nucleic Acids Res* **47**, 8452–8469 (2019).

84. Weiss, M. C. *et al.* The physiology and habitat of the last universal common ancestor. *Nat Microbiol* **1**, 1–8 (2016).

85. Arndt, N. T. & Nisbet, E. G. Processes on the young Earth and the habitats of early life. *Annu Rev Earth Planet Sci* **40**, 521–549 (2012).

86. Vallières, C. *et al.* Iron-sulfur protein odyssey: exploring their cluster functional versatility and challenging identification. *Metallomics* **16**, mfae025 (2024).

87. Mortenson, L. E., Valentine, R. C. & Carnahan, J. E. An electron transport factor from Clostridium pasteurianum. *Biochem Biophys Res Commun* **7**, 448–452 (1962).

88. Schulz, V. *et al.* Mitochondrial [2Fe-2S] ferredoxins: new functions for old dogs. *FEBS Lett* **597**, 102–121 (2023).

89. Crack, J. C. & Le Brun, N. E. Redox-sensing iron–sulfur cluster regulators. *Antioxid Redox Signal* **29**, 1809–1829 (2018).

90. Broderick, W. E. & Broderick, J. B. Radical SAM enzymes: surprises along the path to understanding mechanism. *JBIC Journal of Biological Inorganic Chemistry* **24**, 769–776 (2019).

91. Bai, Y., Chen, T., Happe, T., Lu, Y. & Sawyer, A. Iron–sulphur cluster biogenesis via the SUF pathway. *Metallomics* **10**, 1038–1052 (2018).

92. Lovenberg, W. & Sobel, B. E. Rubredoxin: a new electron transfer protein from Clostridium pasteurianum. *Proceedings of the National Academy of Sciences* **54**, 193–199 (1965).

93. Perrin, B. S. & Ichiye, T. Identifying sequence determinants of reduction potentials of metalloproteins. *JBIC Journal of Biological Inorganic Chemistry* **18**, 599–608 (2013).

94. Belmonte, L. & Mansy, S. S. Patterns of ligands coordinated to metallocofactors extracted from the Protein Data Bank. *J Chem Inf Model* **57**, 3162–3171 (2017).

95. Meyer, J. Iron–sulfur protein folds, iron–sulfur chemistry, and evolution. *Journal of Biological Inorganic Chemistry* **13**, 157–170 (2008).

96. Wiley, S. E. *et al.* The outer mitochondrial membrane protein mitoNEET contains a novel redox-active 2Fe-2S cluster. *Journal of Biological Chemistry* **282**, 23745–23749 (2007).

97. Fee, J. A. *et al.* Purification and characterization of the Rieske iron-sulfur protein from Thermus thermophilus. Evidence for a [2Fe-2S] cluster having non-cysteine ligands. *Journal of Biological Chemistry* **259**, 124–133 (1984).

98. Liu, F. *et al.* An iron reservoir to the catalytic metal: the rubredoxin iron in an extradiol dioxygenase. *Journal of Biological Chemistry* **290**, 15621–15634 (2015).

99. Colbert, C. L., Couture, M. M.-J., Eltis, L. D. & Bolin, J. T. A cluster exposed: structure of the Rieske ferredoxin from biphenyl dioxygenase and the redox properties of Rieske Fe-S proteins. *Structure* **8**, 1267–1278 (2000).

100. Netz, D. J. A., Pierik, A. J., Stümpfig, M., Mühlenhoff, U. & Lill, R. The Cfd1–Nbp35 complex acts as a scaffold for iron-sulfur protein assembly in the yeast cytosol. *Nat Chem Biol* **3**, 278–286 (2007).

101. Lill, R. *et al.* The role of mitochondria and the CIA machinery in the maturation of cytosolic and nuclear iron–sulfur proteins. *Eur J Cell Biol* **94**, 280–291 (2015).

102. Zheng, L., Cash, V. L., Flint, D. H. & Dean, D. R. Assembly of iron-sulfur clusters: identification of an iscSUA-hscBA-fdx gene cluster from Azotobacter vinelandii. *Journal of Biological Chemistry* **273**, 13264–13272 (1998).

103. Cupp-Vickery, J. R., Urbina, H. & Vickery, L. E. Crystal structure of IscS, a cysteine desulfurase from Escherichia coli. *J Mol Biol* **330**, 1049–1059 (2003).

104. Smith, A. D. *et al.* Sulfur transfer from IscS to IscU: the first step in iron− sulfur cluster biosynthesis. *J Am Chem Soc* **123**, 11103–11104 (2001).

105. Agar, J. N. *et al.* IscU as a scaffold for iron− sulfur cluster biosynthesis: sequential assembly of [2Fe-2S] and [4Fe-4S] clusters in IscU. *Biochemistry* **39**, 7856–7862 (2000).

106. Agar, J. N., Zheng, L., Cash, V. L., Dean, D. R. & Johnson, M. K. Role of the IscU protein in iron− sulfur cluster biosynthesis: IscS-mediated assembly of a [Fe2S2] cluster in IscU. *J Am Chem Soc* **122**, 2136–2137 (2000).

107. Braymer, J. J., Freibert, S. A., Rakwalska-Bange, M. & Lill, R. Mechanistic concepts of iron-sulfur protein biogenesis in Biology. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1868**, 118863 (2021).

108. Outten, F. W., Wood, M. J., Munoz, F. M. & Storz, G. The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in Escherichia coli. *Journal of Biological Chemistry* **278**, 45713–45719 (2003).

109. Loiseau, L., Ollagnier-de-Choudens, S., Nachin, L., Fontecave, M. & Barras, F. Biogenesis of Fe-S cluster by the bacterial Suf system: SufS and SufE form a new type of cysteine desulfurase. *Journal of Biological Chemistry* **278**, 38352–38359 (2003).

110. Layer, G. *et al.* SufE transfers sulfur from SufS to SufB for iron-sulfur cluster assembly. *Journal of biological chemistry* **282**, 13342–13350 (2007).

111. Sendra, M., de Choudens, S. O., Lascoux, D., Sanakis, Y. & Fontecave, M. The SUF iron–sulfur cluster biosynthetic machinery: sulfur transfer from the SUFS–SUFE complex to SUFA. *FEBS Lett* **581**, 1362–1368 (2007).

112. Liu, G. *et al.* High-quality homology models derived from NMR and X-ray structures of E. coli proteins YgdK and Suf E suggest that all members of the YgdK/Suf E protein family are enhancers of cysteine desulfurases. *Protein science* **14**, 1597–1608 (2005).

113. Riboldi, G. P., Verli, H. & Frazzon, J. Structural studies of the Enterococcus faecalis SufU [Fe-S] cluster protein. *BMC Biochem* **10**, 1–10 (2009).

114. Nachin, L., Loiseau, L., Expert, D. & Barras, F. SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe—S] biogenesis under oxidative stress. *EMBO J* (2003).

115. Tan, G., Lu, J., Bitoun, J. P., Huang, H. & Ding, H. IscA/SufA paralogues are required for the [4Fe-4S] cluster assembly in enzymes of multiple physiological pathways in Escherichia coli under aerobic growth conditions. *Biochemical Journal* **420**, 463–472 (2009).

116. Roy, A., Solodovnikova, N., Nicholson, T., Antholine, W. & Walden, W. E. A novel eukaryotic factor for cytosolic Fe–S cluster assembly. *EMBO J* (2003).

117. Hausmann, A. *et al.* The eukaryotic P loop NTPase Nbp35: an essential component of the cytosolic and nuclear iron–sulfur protein assembly machinery. *Proceedings of the National Academy of Sciences* **102**, 3266–3271 (2005).

118. Srinivasan, V. *et al.* Structure of the yeast WD40 domain protein Cia1, a component acting late in iron-sulfur protein biogenesis. *Structure* **15**, 1246–1257 (2007).

119. Benjamin, B. *et al.* Structures of fission yeast Inositol pyrophosphate kinase Asp1 in ligand-free, substrate-bound, and product-bound states. *mBio* **13**, e03087-22 (2022).

120. Sun, F. *et al.* AirSR, a [2Fe-2S] cluster-containing two-component system, mediates global oxygen sensing and redox signaling in Staphylococcus aureus. *J Am Chem Soc* **134**, 305–314 (2012).

121. Müllner, M. *et al.* A PAS domain with an oxygen labile [4Fe-4S] 2+ cluster in the oxygen sensor kinase NreB of Staphylococcus carnosus. *Biochemistry* **47**, 13921–13932 (2008).

122. Belén Moreno, M., Durán, A. & Carlos Ribas, J. A family of multifunctional thiamine-repressible expression vectors for fission yeast. *Yeast* **16**, 861–872 (2000).

123. Jakopec, V., Walla, E. & Fleig, U. Versatile use of S chizosaccharomyces pombe plasmids in S accharomyces cerevisiae. *FEMS Yeast Res* **11**, 653–655 (2011).

124. Micsonai, A. *et al.* BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**, W315–W322 (2018).

125. Micsonai, A. *et al.* BeStSel: webserver for secondary structure and fold prediction for protein CD spectroscopy. *Nucleic Acids Res* **50**, W90–W98 (2022).

126. Micsonai, A., Bulyáki, É. & Kardos, J. BeStSel: from secondary structure analysis to protein fold prediction by circular dichroism spectroscopy. *Structural genomics: general applications* 175–189 (2021).

127. Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. & Schubert, D. Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys J* **82**, 1096–1111 (2002).

128. Yang, Z., Zeng, X., Zhao, Y. & Chen, R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct Target Ther* **8**, 115 (2023).

129. Lensink, M. F. *et al.* Impact of AlphaFold on structure prediction of protein complexes: the CASP15-CAPRI experiment. *Proteins: Structure, Function, and Bioinformatics* **91**, 1658–1683 (2023).

130. Varadi, M. *et al.* AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res* **52**, D368–D375 (2024).

131. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

132. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).

133. Ursula Fleig. Key motifs of the Asp1 pyrophosphatase domain. (2024).

134. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389–W394 (2015).

135. Niedziela-Majka, A., Rymarczyk, G., Kochman, M. & Ozyhar, A. Pure, bacterially expressed DNA binding domains of the functional ecdysteroid receptor capable of interacting synergistically with the Hsp 27 20-hydro xyecdysone response element: GST induced dimerization of DNA binding domains alters characteristics of their interaction with DNA. *Protein Expr Purif* **14**, 208–220 (1998).

136. Parker, M. W., Bello, M. Lo & Federici, G. Crystallization of glutathione S-transferase from human placenta. *J Mol Biol* **213**, 221–222 (1990).

137. Robichon, C., Luo, J., Causey, T. B., Benner, J. S. & Samuelson, J. C. Engineering Escherichia coli BL21 (DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography. *Appl Environ Microbiol* **77**, 4634–4646 (2011).

138. Akhtar, M. K. & Jones, P. R. Deletion of iscR stimulates recombinant clostridial Fe–Fe hydrogenase activity and H2-accumulation in Escherichia coli BL21(DE3). *Appl Microbiol Biotechnol* **78**, 853–862 (2008).

139. Schwartz, C. J. *et al.* IscR, an Fe-S cluster-containing transcription factor, represses expression of Escherichia coli genes encoding Fe-S cluster assembly proteins. *Proceedings of the National Academy of Sciences* **98**, 14895–14900 (2001).

140. Takamura, K., Fischer, H. & Morrow, N. R. Physical properties of aqueous glycerol solutions. *J Pet Sci Eng* **98**, 50–60 (2012).

141. Vagenende, V., Yap, M. G. S. & Trout, B. L. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry* **48**, 11084–11096 (2009).

142. Waugh, D. S. An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr Purif* **80**, 283–293 (2011).

143. Leong, L. E. C. The use of recombinant fusion proteases in the affinity purification of recombinant proteins. *Mol Biotechnol* **12**, 269–274 (1999).

144. Achebach, S., Selmer, T. & Unden, G. Properties and significance of apoFNR as a second form of air-inactivated [4Fe-4S]· FNR of Escherichia coli. *FEBS J* **272**, 4260–4269 (2005).

145. Xu, F. F. & Imlay, J. A. Silver (I), mercury (II), cadmium (II), and zinc (II) target exposed enzymic iron-sulfur clusters when they toxify Escherichia coli. *Appl Environ Microbiol* **78**, 3614–3621 (2012).

146. Tan, G. *et al.* Competition of zinc ion for the [2Fe–2S] cluster binding site in the diabetes drug target protein mitoNEET. *Biometals* **25**, 1177–1184 (2012).

147. Iannuzzi, C. *et al.* The role of zinc in the stability of the marginally stable IscU scaffold protein. *Protein Science* **23**, 1208–1219 (2014).

148. Boyland, E. & Chasseaud, L. F. The role of glutathione and glutathione S-transferases in mercapturic acid biosynthesis. *Adv Enzymol Relat Areas Mol Biol* **32**, 173–219 (1969).

149. Daniel, T., Faruq, H. M., Laura Magdalena, J., Manuela, G. & Christopher Horst, L. Role of GSH and iron-sulfur glutaredoxins in iron metabolism. *Molecules* **25**, 3860 (2020).

150. Daniel, T., Faruq, H. M., Laura Magdalena, J., Manuela, G. & Christopher Horst, L. Role of GSH and iron-sulfur glutaredoxins in iron metabolism. *Molecules* **25**, 3860 (2020).

151. Smith, G. L. *et al.* Complexation of ferrous ions by ferrozine, 2, 2′-bipyridine and 1, 10-phenanthroline: Implication for the quantification of iron in biological systems. *J Inorg Biochem* **220**, 111460 (2021).

152. Lopata, A. *et al.* Beyond chelation: EDTA tightly binds taq DNA polymerase, MutT and dUTPase and directly inhibits dNTPase activity. *Biomolecules* **9**, 621 (2019).

153. Lovenberg, W., Buchanan, B. B. & Rabinowitz, J. C. Studies on the chemical nature of clostridial ferredoxin. *Journal of Biological Chemistry* **238**, 3899–3913 (1963).

154. Mernagh, D. R., Reynolds, L. A. & Kneale, G. G. DNA binding and subunit interactions in the type I methyltransferase M. Eco R124I. *Nucleic Acids Res* **25**, 987–991 (1997).

155. Ji, X., Zhang, P., Armstrong, R. N. & Gilliland, G. L. The three-dimensional structure of a glutathione S-transferase from the Mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2-. ANG. resolution. *Biochemistry* **31**, 10169–10184 (1992).

156. Haldeman, M. T., Xia, G., Kasperek, E. M. & Pickart, C. M. Structure and function of ubiquitin conjugating enzyme E2-25K: the tail is a core-dependent activity element. *Biochemistry* **36**, 10526–10537 (1997).

157. Nemoto, T., Ohara-Nemoto, Y., Shimazaki, S. & Ota, M. Dimerization characteristics of the DNA-and steroid-binding domains of the androgen receptor. *J Steroid Biochem Mol Biol* **50**, 225–233 (1994).

158. Maru, Y., Afar, D. E., Witte, O. N. & Shibuya, M. The dimerization property of glutathione S-transferase partially reactivates Bcr-Abl lacking the oligomerization domain. *Journal of Biological Chemistry* **271**, 15353–15357 (1996).

159. Mayhew, S. G. The redox potential of dithionite and SO− 2 from equilibrium reactions with flavodoxins, methyl viologen and hydrogen plus hydrogenase. *Eur J Biochem* **85**, 535–547 (1978).

160. O'Neil, M. J. Editor. The Merck Index. Preprint at (2001).

161. Alliegro, M. C. Effects of dithiothreitol on protein activity unrelated to thiol–disulfide exchange: for consideration in the analysis of protein function with Cleland's reagent. *Anal Biochem* **282**, 102–106 (2000).

162. Khazim, K. *et al.* Glutathione redox potential is low and glutathionylated and cysteinylated hemoglobin levels are elevated in maintenance hemodialysis patients. *Translational Research* **162**, 16–25 (2013).

163. Marmelstein, A. M., Yates, L. M., Conway, J. H. & Fiedler, D. Chemical pyrophosphorylation of functionally diverse peptides. *J Am Chem Soc* **136**, 108–111 (2014).

164. Pandelia, M.-E., Lanz, N. D., Booker, S. J. & Krebs, C. Mössbauer spectroscopy of Fe/S proteins. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1853**, 1395–1405 (2015).

165. Rouault, T. A. *et al. Iron-Sulfur Clusters in Chemistry and Biology*. (De Gruyter Berlin, 2014).

166. Jansing, M., Mielenbrink, S., Rosenbach, H., Metzger, S. & Span, I. Maturation strategy influences expression levels and cofactor occupancy in Fe–S proteins. *JBIC Journal of Biological Inorganic Chemistry* **28**, 187–204 (2023).

167. Münck, E., Debrunner, P. G., Tsibris, J. C. M. & Gunsalus, I. C. Mössbauer parameters of putidaredoxin and its selenium analog. *Biochemistry* **11**, 855–863 (1972).

168. Chandramouli, K. *et al.* Formation and properties of [4Fe-4S] clusters on the IscU scaffold protein. *Biochemistry* **46**, 6804–6811 (2007).

169. Arsın, H. *et al.* Exploring codon adjustment strategies towards escherichia coli-based production of viral proteins encoded by hth1, a novel prophage of the marine bacterium hypnocyclicus thermotrophus. *Viruses* **13**, 1215 (2021).

170. Atherton, N. M. *et al.* Spectroscopic investigation of rubredoxin and ferredoxin. *Nature* **212**, 590–593 (1966).

171. Jung, Y.-S., Gao-Sheridan, H. S., Christiansen, J., Dean, D. R. & Burgess, B. K. Purification and biophysical characterization of a new [2Fe-2S] ferredoxin from Azotobacter vinelandii, a putative [Fe-S] cluster assembly/repair protein. *Journal of Biological Chemistry* **274**, 32402–32410 (1999).

172. Zeng, J., Huang, X., Liu, Y., Liu, J. & Qiu, G. Expression, purification, and characterization of a [Fe 2 S 2] cluster containing ferredoxin from Acidithiobacillus ferrooxidans. *Curr Microbiol* **55**, 518–523 (2007).

173. Gräwert, T. *et al.* IspH protein of Escherichia c oli: studies on Iron− Sulfur cluster implementation and catalysis. *J Am Chem Soc* **126**, 12847–12855 (2004).

174. Glebov, L. B. & Boulos, E. N. Absorption of iron and water in the Na2O–CaO–MgO–SiO2 glasses. II. Selection of intrinsic, ferric, and ferrous spectra in the visible and UV regions. *J Non Cryst Solids* **242**, 49–62 (1998).

175. Ollagnier-de Choudens, S., Loiseau, L., Sanakis, Y., Barras, F. & Fontecave, M. Quinolinate synthetase, an iron–sulfur enzyme in NAD biosynthesis. *FEBS Lett* **579**, 3737–3743 (2005).

176. Krewulak, K. D. & Vogel, H. J. Structural biology of bacterial iron uptake. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1778**, 1781–1804 (2008).

177. Pandey, A., Bringel, F. & Meyer, J.-M. Iron requirement and search for siderophores in lactic acid bacteria. *Appl Microbiol Biotechnol* **40**, 735–739 (1994).

178. Imbert, M. & Blondeau, R. On the iron requirement of lactobacilli grown in chemically defined medium. *Curr Microbiol* **37**, 64–66 (1998).

179. Bruyneel, B., Vande Woestyne, M. & Verstraete, W. Lactic acid bacteria: micro-organisms able to grow in the absence of available iron and copper. *Biotechnol Lett* **11**, 401–406 (1989).

180. Archibald, F. Lactobacillus plantarum, an organism not requiring iron. *FEMS Microbiol Lett* **19**, 29–32 (1983).

181. Hansen, J. M. & Harris, C. Glutathione during embryonic development. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1850**, 1527–1542 (2015).

182. Mailloux, R. J. & Willmore, W. G. S-glutathionylation reactions in mitochondrial function and disease. *Front Cell Dev Biol* **2**, 68 (2014).

183. Van, V. *et al.* Iron-sulfur clusters are involved in post-translational arginylation. *Nat Commun* **14**, 458 (2023).

184. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proceedings of the National Academy of Sciences* **97**, 11984–11989 (2000).

185. Bhopatkar, A. A., Uversky, V. N. & Rangachari, V. Disorder and cysteines in proteins: A design for orchestration of conformational see-saw and modulatory functions. *Prog Mol Biol Transl Sci* **174**, 331–373 (2020).

186. McCarthy, E. L. & Booker, S. J. Bridging a gap in iron-sulfur cluster assembly. *Elife* **4**, e10479 (2015).

187. Camponeschi, F. & Banci, L. Metal cofactors trafficking and assembly in the cell: a molecular view. *Pure and Applied Chemistry* **91**, 231–245 (2019).

188. Van, V. & Smith, A. T. Reconstitution of the Arginyltransferase (ATE1) Iron-Sulfur Cluster. in *Protein Arginylation: Methods and Protocols* 209–217 (Springer, 2023).

189. Hausmann, A., Samans, B., Lill, R. & Mühlenhoff, U. Cellular and mitochondrial remodeling upon defects in iron-sulfur protein biogenesis. *Journal of Biological Chemistry* **283**, 8318–8330 (2008).

190. Sharma, A. K., Pallesen, L. J., Spang, R. J. & Walden, W. E. Cytosolic iron-sulfur cluster assembly (CIA) system: factors, mechanism, and relevance to cellular iron regulation. *Journal of Biological Chemistry* **285**, 26745–26751 (2010).

Declaration

I declare under oath that I have compiled this dissertation independently and without any undue assistance by third parties under consideration of the "Fundamental principles for safeguarding good scientific practice at Heinrich-Heine-Universität Düsseldorf".

Furthermore, neither this dissertation, nor a similar work, has been submitted to another faculty. I have not made any unsuccessful attempt to obtain a doctorate.

Düsseldorf, 16.07.2025
Ort, Datum

Bianca Axinte
Bianca Georgiana Axinte