TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Chemie

Lehrstuhl für Organische Chemie II

CLPXP PROTEASE NETWORKS STUDIED BY CHEMICAL CROSS-LINKING COMBINED WITH MASS SPECTROMETRY

A PROTEOMIC SURVEY ON HUMAN PYRIDOXAL 5'-PHOSPHATE-BINDING PROTEINS

DISSERTATION ZUR ERLANGUNG DES AKADEMISCHEN GRADES EINES
DOKTORS DER NATURWISSENSCHAFTEN VON

Anja Stefanie Fux

MÜNCHEN 2019

TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Chemie

Lehrstuhl für Organische Chemie II

CLPXP PROTEASE NETWORKS STUDIED BY CHEMICAL CROSS-LINKING COMBINED WITH MASS SPECTROMETRY

and

A PROTEOMIC SURVEY ON HUMAN PYRIDOXAL 5'-PHOSPHATE-BINDING PROTEINS

Anja Stefanie Fux

Vollständiger Abdruck der von der Fakultät für Chemie der Technische Universität München zur Erlangung des akadamischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende: Prof. Dr. Kathrin Lang

Prüfer der Dissertation: 1. Prof. Dr. Stephan A. Sieber

2. Prof. Dr. Matthias J. Feige

Die Dissertation wurde am 18.09.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 28.10.2019 angenommen.

Danksagung

An erster Stelle möchte ich mich bei Prof. Dr. Stephan Sieber bedanken, dass er mir die Möglichkeit gegeben hat, meine Doktorarbeit an seinem Lehrstuhl anfertigen zu können. Besonders geschätzt habe ich die Vielzahl an wissenschaftlichen Fragestellungen und die damit verbundenen erlernten Techniken. Ich möchte mich auch für die stete Begeisterung, Unterstützung, und Anregung, die mir bei meinen Projekten entgegengebracht wurden, bedanken. Zu schätzen gewusst habe ich vor allem auch die exzellente Ausstattung unseres Labors, die durch sein Engagement und Talent bezüglich der Erlangung von Fördermitteln möglich gewesen ist.

Ich möchte mich in diesem Zuge bei meinen Kooperationspartnern Markus Schneider und Iris Antes (TUM, Weihenstephan) für die angenehme Zusammenarbeit und deren Expertise im Feld des *Molecular Modeling* bedanken. Zudem bedanke ich mich bei Isabel Buchsbaum und Silvia Cappello (Max-Planck-Insitut für Psychatrie, München) für immunohistochemische Aufnahmen.

Mein besonderer Dank gilt allen derzeitigen als auch ehemaligen Mitgliedern des AK Sieber für die angenehme und inspirierende Atmosphäre, für zahlreiche fachlich hochwertige Gespräche, den Austausch von Ideen, aber auch für nette Abende zum Ausklang. Ich möchte mich insbesondere bei Dr. Vadim Korotkov und Martin Pfanzelt für die synthetische Unterstützung bedanken. Des Weiteren bedanke ich mich bei Volker Kirsch und Philipp Le für die ausgezeichneten und lustigen Unterhaltungen während zahlreicher Mensa-Sessions. Bei Volker als meinen Banknachbaren möchte ich mich besonders bedanken für die schönen Jahre nebeneinander. Zusätzlich geht mein Dank an alle Mitglieder von Lab D, Ines Hübner, Till Reinhardt, Robert Macsics, Volker Kirsch, und Katja Bäuml für die lockere aber auch äußerst produktive Arbeitsatmosphäre. Des Weiteren möchte ich mich noch bei Carolin Gleissner und Martin Pfanzelt für alle fachlichen als auch nicht so fachlichen, aber sehr unterhaltenden Gespräche bedanken. Zuletzt gilt mein Dank Mona Wolff, Katja Bäuml, und Katja Gliesche für die technische Unterstüzung die den Laboralltag um ein Vieles erleichtert haben.

Ich möchte mich zudem bei allen ehemaligen Studienkollegen bedanken, von denen einige inzwischen gute Freunde geworden sind und auch im Doktor meinen Weg weiter begleitet haben, namentlich Marina Schopf, Barabara Tremmel und Katharina Wißmiller.

Zuletzt gilt mein größter Dank meinen Freunden und meiner Familie, insbesondere meiner Mutter, für die Unterstützung und den immerwährenden Rückhalt über acht lange Jahre des

Studiums und der Promotion. Meiner besten Freundin Jennifer Albrecht danke ich für jedes aufmunternde Gespräch und jedes „Das wird schon alles!". Des Weiteren danke ich meinem langjährigen Mitbewohner und Freund Fabian Hörmann für die beste WG aller Zeiten, den starken Zusammenhalt, und einfach für alles! An letzter Stelle möchte ich mich bei Dr. Christoph Brenninger bedanken, den ich im Laufe der Doktorarbeit kennen gelernt habe und der seitdem stets unterstützend sowie aufbauend an meiner Seite war. Vielen Dank!

Abstract

The first part of this dissertation presents the analysis of the ClpXP protease by chemical cross-linking combined with mass spectrometry (XL-MS). The ClpXP protease represents a conserved machinery, built up by the association of heptameric ClpP peptidase rings with hexameric ClpX chaperones. Although the protease has been studied extensively over the last decades and several high-resolution structures of the ClpP part are available, structural information about the ClpX-ClpP interface, as well as the subset of cellular binding partners are lacking. XL-MS displays a versatile technology to address diverse structural questions and becomes more and more the method of choice for the study of multisubunit protein complexes like the ClpXP system. Moreover, advanced MS-technologies and cross-linker chemistries have guided the way to the analysis of intracellular protein-protein interactions by XL-MS in the background of whole proteomes. Here, we set-up an efficient experimental workflow for the analysis of recombinant proteins with the MS-cleavable cross-linker DSSO. We select the peptidase ClpP for proof-of-principle studies to evaluate our sample preparation, acquisition of MS spectra and data analysis strategies. With a validated and robust cross-linking workflow in hand, we examined the interface of the ClpX-ClpP system. We were able to confirm several interactions postulated by previous mutagenesis and docking studies, as well as to identifiy so far unknown contacts. A molecular dynamics simulation on the basis of the cross-links revealed a first model of the ClpXP interface guided by structural data suggesting high flexibility of the protease machinery in solution. Further, we applied *in situ* chemical cross-linking followed by co-immunoprecipitation (cross-link/co-IP) to study cellular binding partners of the human and *Escherichia coli* peptidase. Besides the confirmation of several previously identified interactors, our proteomic analysis resulted in a manifold of so far not recognized contacts. Moreover, we demonstrated the superior performance of cross-link/co-IP over conventional co-IPs and co-IPs combined with on-bead cross-linking. Finally, we explore here for the first time a combination of the activity-based protein-profiling (ABPP)-approach with *in situ* chemical cross-linking. Our goal was to study in parallel both, the cellular targets of active compounds or small molecules, as well as identify the interactors of the target proteins in one experiment.

The second part of this thesis deals with the application of previously developed vitamin B6-mimics to study the subset of human pyridoxal 5'-phosphate (**PLP**)-binding proteins. In a first step, we evaluated the metabolic activation of our probes by the human pyridoxal kinase, as this displays the crucial step to generate the biologically active form of the vitamin. Interestingly, we observed significant differences in the turnover of the probes compared to the previously

Abstract

evaluated *Staphylococcus aureus* system. However, we examined the probes for their capacity to bind human **PLP**-dependent proteins *in situ* by a workflow based on ABPP. A metabolic labeling approach with the minimal modified probe led to selective enrichment of a large fraction (one-third) of the known human **PLP**-dependent proteins, including the unacharacterized pyridoxal 5'-phosphate-binding protein (PLPBP). We subsequently validated our labeling results for two vitamin B6-dependent proteins. As an example of application, we proceeded with the evaluation of the *in situ* targets of two anti-vitamin B6-compounds *via* a competitive chemoproteomic profiling strategy. Moreover, we expanded our proteomic screening to different human cell lines which resulted in signature labeling profiles of **PLP**-dependent proteins varying in tested cell types. As we excluded different expression levels as reason for this observation, an altered cofactor loading of vitamin B6-dependent proteins might be a possible but so far not recognized mechanism of activity regulation within the cell. Further, we investigated the subset of binding partners to vitamin B6 salvage proteins in order to study intracellular trafficking of the cofactor more in detail. Finally, we evaluated structural features as well as cellular roles of PLPBP, revealing interesting insights into the function of this highly conserved protein.

Zusammenfassung

Der erste Teil dieser Doktorarbeit befasst sich mit der Untersuchung der konservierten Protease ClpXP mittels chemischem Cross-linking in Kombination mit Massenspektrometrie (XL-MS). XL-MS entwickelte sich im Laufe der Jahre zu einer populären Methodik um die Zusammensetzung von insbesondere hoch-molekularen Komplexen mit mehreren Untereinheiten aufzuklären. Eine fortlaufende Weiterentwicklung und Optimierung von Cross-linker-Chemie als auch massenspektrometrischen Messmethoden ermöglichte die Untersuchung von Interaktionen im Hintergrund ganzer Proteome. Die ClpXP Protease setzt sich aus zwei Heptameren der Peptidase ClpP, die zusammen ein Tetradekamer bilden, sowie aus ein bis zwei Hexameren der ATPase ClpX, die mit den jeweiligen ClpP Heptameren assoziiert sind, zusammen. Obwohl die Protease seit vielen Jahrzehnten im Fokus der Forschung ist und bereits zahlreiche hochauflösende Strukturen von ClpP Orthologen gelöst wurden, fehlt jedoch strukturelle Information über den gesamten ClpXP Komplex, als auch detaillierte Untersuchungen zu zellulären Bindungspartnern unter nativen Bedingungen. Ziel der Dissertation ist die Entwicklung eines effizienten Cross-linking Arbeitsablaufs zur Untersuchung von Kontakten innerhalb von ClpXP, zusammen mit einer robusten Messanalyse sowie nachfolgenden bioinformatischen Auswertung. Um die experimentelle Vorgehensweise zu validieren wurden zunächst ClpP Komplexe von unterschiedlichen Organismen mit dem DSSO Cross-linker behandelt und identifizierte Kontake anhand bekannter Kristallstrukturen verifiziert. In einem nächsten Schritt wurde der Arbeitsablauf auf ClpXP Orthologe übertragen. Im Zuge dessen wurden viele interessante Kontakte identifiziert. Die experimentell-ermittelten Cross-links dienten in der Folge als Input für eine molekulardynamische Simulation des ClpX-ClpP Interfaces. In einem nächsten Schritt wurde *in situ* chemisches Cross-linken von humanen als auch bakteriellen Zellen mit co-Immunopräzipitation (Cross-link/co-IP) und MS kombiniert um das Repertoir von Bindepartnern der Protease unter nativen Bedingungen zu studieren. Dabei wurden zunächst konventionelle co-IP Methoden mit co-IP in Kombination mit *on-bead* Cross-linken und Cross-link/co-IP verglichen. Dabei konnte eindeutig gezeigt werden, dass Cross-link/co-IP die konfidenteren Ergebnisse liefert. Neben bereits in früheren Arbeiten identifizierten Interaktionspartnern konnten viele bisher unbekannte Kontakte aufgedeckt werden. In einem letzten Schritt wurde chemisches Cross-linken mit der Methodik des aktivitätsbasierten Protein-Profiling (ABPP) kombiniert. Ziel war die parallele Identifikation von Zielproteinen der eingesetzten Sonden, sowie deren zellulären Interaktionspartnern in einem Experiment.

Zusammenfassung

Der zweite Teil der Arbeit hatte die Applikation von sich von Vitamin B6-ableitenden Sonden zur Untersuchung humaner Pyridoxal 5'-phosphat (**PLP**)-bindender Proteine zum Ziel. Diese bereits in einer vorherigen Arbeit entwickelten Sonden wurden zunächst auf ihre Bioaktivierung durch die humane Pyridoxalkinase untersucht. Dabei ergaben sich signifikante Unterschiede in der Affinität der Kinase zu den unterschiedlichen Sonden v.a. im Vergleich mit dem bereits untersuchten *Staphylococcus aureus* System. In einem nächsten Schritt wurden durch einen Arbeitsablauf basierend auf ABPP zelluläre Targets der Sonden ermittelt. Die minimal-modifizierte Sonde reicherte einen großen Teil (nahezu ein Drittel) der bekannten **PLP**-bindenden Proteine nach metabolischer Inkooperation an. In einem nächsten Schritt validierten wir unsere Ergebnisse anhand zweier Vitamin B6-abhängiger Proteine, mitunter anhand von einem bisher nahezu uncharakterisierten Kandidaten namens Pyridoxal 5'-phosphat-bindendes Protein (PLPBP). Als ein Anwendungsbeispiel ermittelten wir die *in situ* Zielproteine zweier bekannter Vitamin B6-Antagonisten. Dabei wurde ein unerwartet schmales Spektrum an **PLP**-bindenen Substraten identifiziert. Fortführend verglichen wir den Pool an **PLP**-bindenen Proteinen, die durch die minimal-modifizierte Sonde getroffen werden in verschiedenen Humanzelllinien. Dabei offenbarten sich Unterschiede in Anzahl und Art der Vitamin B6-abhängigen Proteine, sowie deren Beladung mit Kofaktor. Da wir alternierende Expressionsgrade sowohl von den **PLP**-bindenen Proteinen selbst, als auch von den Vitamin B6-metabolisierenden Enzymen wie der Pyridoxalkinase ausschließen konnten, suggerieren unsere Ergebnisse eine Zelltyp-spezifische Beladung mit Kofaktor, was ein Mittel zur Aktivitätsregulation darstellen könnte. Des Weiteren untersuchten wir die Bidungspartner der Vitamin B6-metabolisierenden Enzyme um Einblicke in den intrazellulären Transport von **PLP** zu bekommen. In einem letzten Schritt widmeten wir uns der strukturellen als auch funktionellen Charakterisierung von PLPBP und dessen Epilepsie-assoziierten Mutanten.

Publications

Journal Publications

A. Fux, V. S. Korotkov, M. Schneider, I. Antes, and S. A. Sieber. **Chemical cross-linking enables drafting ClpXP proximity maps and taking snapshots of *in situ* interaction networks.** Cell Chem Biol., 2019 Jan 17, 26 (1):48-59

A. Fux, M. Pfanzelt, V. C. Kirsch, A. Hoegl, and S. A. Sieber. **Customizing functionalized cofactor mimics to study the human pyridoxal 5'-phosphate-binding proteome.** Cell Chem Biol, 2019 Oct 17, 26 (1):1-8.

A. Fux and S. A. Sieber. **Biochemical and proteomic studies of human pyridoxal 5'-phosphate-binding protein (PLPBP).** Manuscript under revision.

Conferences

A. Fux, V. S. Korotkov, and S. A. Sieber. **Protease Networks elucidated by Chemical Cross-Linking and Mass Spectrometry.** *SFB1035-Conference*, Venice, Italy, 01-06 May, 2018 (Poster presentation)

A. Fux, M. Pfanzelt, A. Hoegl, and S. A. Sieber. **Customizing vitamin B6 cofactor mimics to study human pyridoxal 5'-phosphate binding proteins.** *Gordon Research Conference (Bioorganic Chemistry)*, Andover, NH, US, 09-14 June, 2019 (Poster presentation)

Table of Contents

Table of Contents

Table of Contents

Table of Contents

I – CLPXP PROTEASE NETWORKS STUDIED BY CHEMICAL CROSS-LINKING COMBINED WITH MASS SPECTROMETRY

The following part of this thesis is based upon:

Contributions:

A.F. planned and conducted all experiments, performed MS analysis, analyzed MS data, performed protein modeling, and mapped cross-links. V.S.K. synthesized DSSO. M.S. performed protein modeling and molecular dynamics simulations. I.A. supervised computational modeling studies. S.A.S. supervised experiments.

1 Introduction

1.1 Chemical cross-linking combined with mass spectrometry

In order to fully understand the functions of proteins in a cell, a detailed knowledge of the three-dimensional protein structure and the interacting amino acid sequences in protein complexes are of outstanding importance. High-resolution techniques like X-ray crystallography, nuclear magnetic resonance (NMR), and electron-microscopy (EM) address the analysis of topological and spatial organization of proteins and protein complexes. However, these methods depend on the availability of comparable high sample amounts. Therefore, mass spectrometry (MS) offers an attractive alternative for high-throughput analysis of proteins, as required amounts are in the ng-region. Nowadays, MS applications enable the large-scale identification and quantification of proteins within whole cells or tissues.[1,2] However, a detailed knowledge about protein structure, as well as topologies of protein complexes is often missing. Chemical cross-linking in combination with MS (XL-MS) has developed in recent years into a robust technology offering a powerful tool to gain insight into structural information about proteins, composition of protein complexes, and protein-protein interactions (PPI) even on a cellular scale. A big advantage of XL-MS is its ability to provide a snapshot from an ensemble of protein conformations and interactions in solution representing their native states, whereas other structural techniques like EM require the fixation of the protein previous to analysis.

1.1.1 Cross-linker types and reactivities

The aim of the cross-linking reaction is the introduction of a covalent bond between two spatial proximate residues. Cross-linkers therefore consist of two reactive end groups which are connected by a spacer, typically an alkyl chain.[3] The reactive groups determine which amino

acids are targeted and therefore the amount of spatial information that can be obtained. A cross-linker can either be homobifunctional with two equal moieties at each end or heterobifunctional combining different reactivities. The length of the spacer arm ranges from zero-length, where no additional atoms are added to the cross-linked species, to 40 Å and therefore determines the maximal possible distance between two residues that can be cross-linked.[4]

Several different cross-linker reactivities have been described, with the most commonly applied ones displaying *N*-hydroxysuccinimide (NHS) or sulfosuccinimidyl esters (**Scheme I-1A**) targeting primary amines as in lysine residues.[5] Lysine cross-linking benefits from the high reaction rates of applied esters as well as from the fact that lysine residues have a high prevalence of about 6% within proteins and are mainly located on the protein surface.[3]

However, several proteins display only few or no lysine residues and are therefore difficult to characterize with amine-targeting cross-linkers. Therefore, application of other reactivities in chemical cross-linking received attraction. Cysteine residues are targeted with cross-linkers equipped for instance with a maleimide moiety, though these amino acids have only a low abundance of 2% within the proteome (**Scheme I-1B**).[3] Further, acidic residues as in aspartate and glutamate with a high prevalence of 12% are addressed by the application of hydrazides in combination with a coupling reagent like 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM) enabling reaction at physiological pH values (**Scheme I-1C**).[6] In addition, photo-reactive cross-linkers like aryl azides are applied (**Scheme I-1D**),[7,8] as well as linkers containing aryl sulfonyl fluoride moieties for proximity-enhanced sulfur-fluoride exchange (SuFEx) reactions with weak nucleophilic amino acid side-chains like in Ser, Thr, Tyr, or His (**Scheme I-1E**).[9]

To date, a manifold of cross-linker types and reactivities have been developed, enabling a wide-ranging application of XL-MS and the study of nearly any protein or protein complex of interest.

A

B

C

D

E

Nuc = Lys, His, Ser, Thr, Tyr
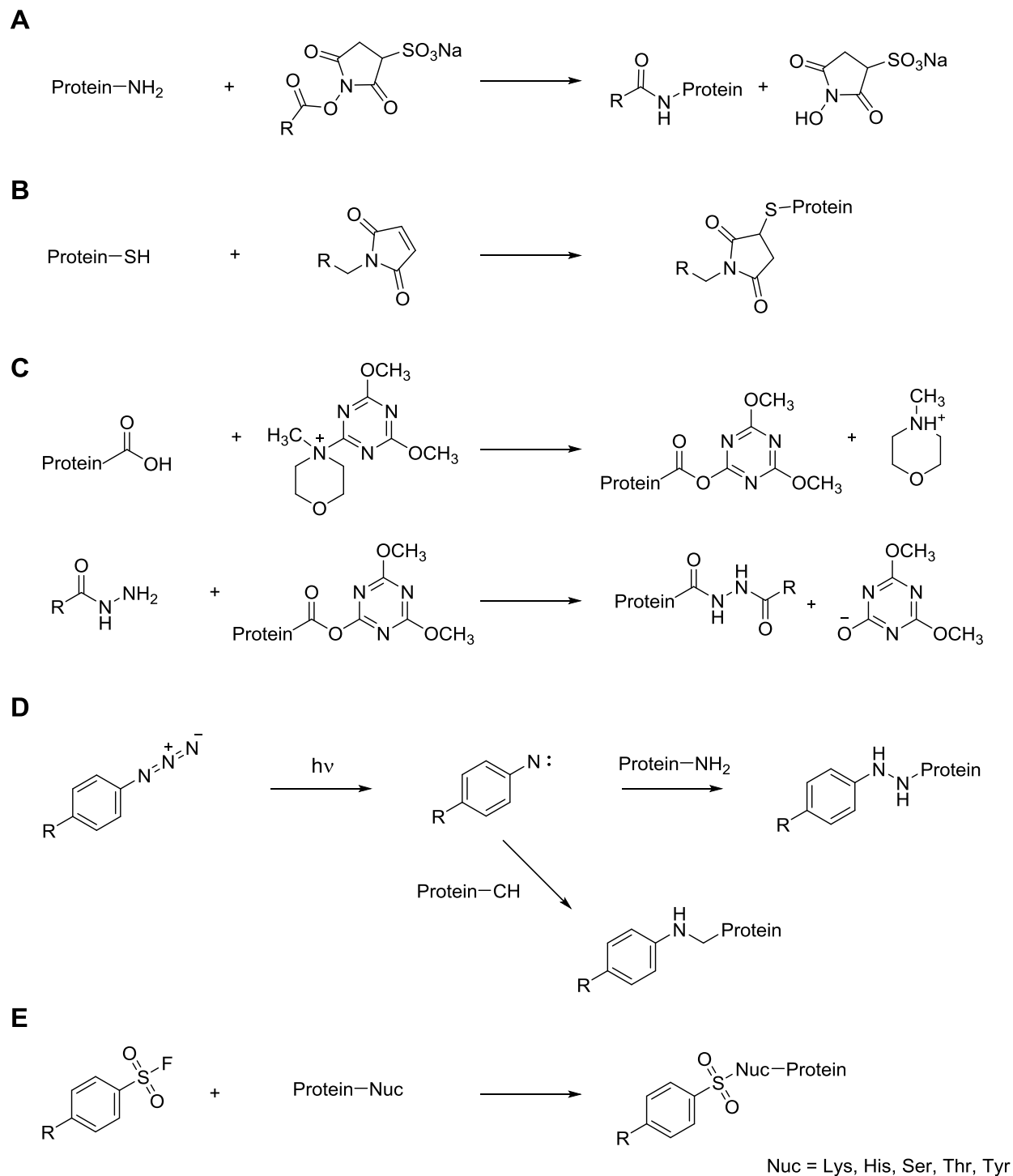
Scheme I-1 Reactivities of chemical cross-linkers. Typical cross-linking moieties are amine-reactive sulfosuccinidmyl esters (**A**) and maleimides reacting with sulfhydryl-groups as in cysteines (**B**). (**C**) Upon activation of carboxyl-containing amino acids with the coupling reagent DMTMM, glutamic and aspartic residues can react with hydrazides. (**D**) Photoreactive linkers like aryl azides insert unspecific into CH or NH bonds after irradiation with UV-light. (**E**) Cross-linkers displaying a sulfonyl fluoride moiety allow cross-linking of several nucleophilic amino acids *via* SuFEx reaction.

1.1.2 The general XL-MS workflow

A typical cross-linking workflow (**Figure I-1**) is initiated with the addition of the cross-linker or activation reagent previous to the addition of the linker. Cross-linking carries out for a certain time depending on the reactivity of the linker as well as the accessibility of residues to be connected. Afterwards, the reaction is terminated by either quenching or removal of the cross-linker for example by gel-filtration. Cross-linked samples are then either directly subjected to proteolysis (*e.g.* with trypsin) or can first be analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). A gel-based separation step can in addition facilitate the removal of non-cross-linked species and/or undesired oligomeric states. However, sample recovery is often hampered due to inefficient tryptic digestion within gels.[2]

The resulting peptide mixture consists of several different species.[10,11] Besides single, unmodified linear peptides, peptides that are attached to only one arm of the cross-linker while the other moiety remains free (dead-end peptides) and so called loop-linked peptides, which have two of their amino acids attached to either end of the same cross-linker, are detected. These kind of peptides do not bear any spatial information, but rather provide information on the solvent accessibility of the modified residues.[12] The desired fraction display cross-linked peptides, where two peptides are linked by a single cross-linker. These two peptides can either originate from the same protein (intra protein cross-links) or from different proteins (inter protein cross-links). Intra protein cross-links give hints on the positions of domains within a protein or its conformation and provide restraints for homology modeling or docking studies.[13,14] Inter protein cross-links on the other hand contain valuable information on spatial proximities of single proteins or proteins within a complex.[15]

After digestion of proteins, the peptide mixture can be subjected to an optional enrichment step for cross-linked peptides (cf. Part I, Chapter 1.1.4) and is finally analyzed by liquid chromatography combined with tandem mass spectrometry (LC-MS/MS). The resulting data are then fed into different software tools for identification of cross-links. Examples of currently available and widely applied programs for XL-MS data analysis are Xi, developed by the group of *Juri Rappsilber*,[16] StavroX, developed by the group of *Andrea Sinz*,[17] xQuest, developed by the group of *Ruedi Aebersold*,[5] or XlinkX, developed by the group of *Albert Heck*.[18] The type of software to choose depends largely on the type of cross-linker used, or on the type of experiment performed. Identified cross-links can be applied as distance restraints and therefore be converted into structure models for example by molecular docking with ROSETTA[14] or HADDOCK,[19] or *de novo* structure prediction.[20] Finally, several tools have been developed for

visualization of cross-linking data for instance by mapping identified contacts onto modeled or existing structures for validation (*e.g.* Xlink Analyzer[15]).
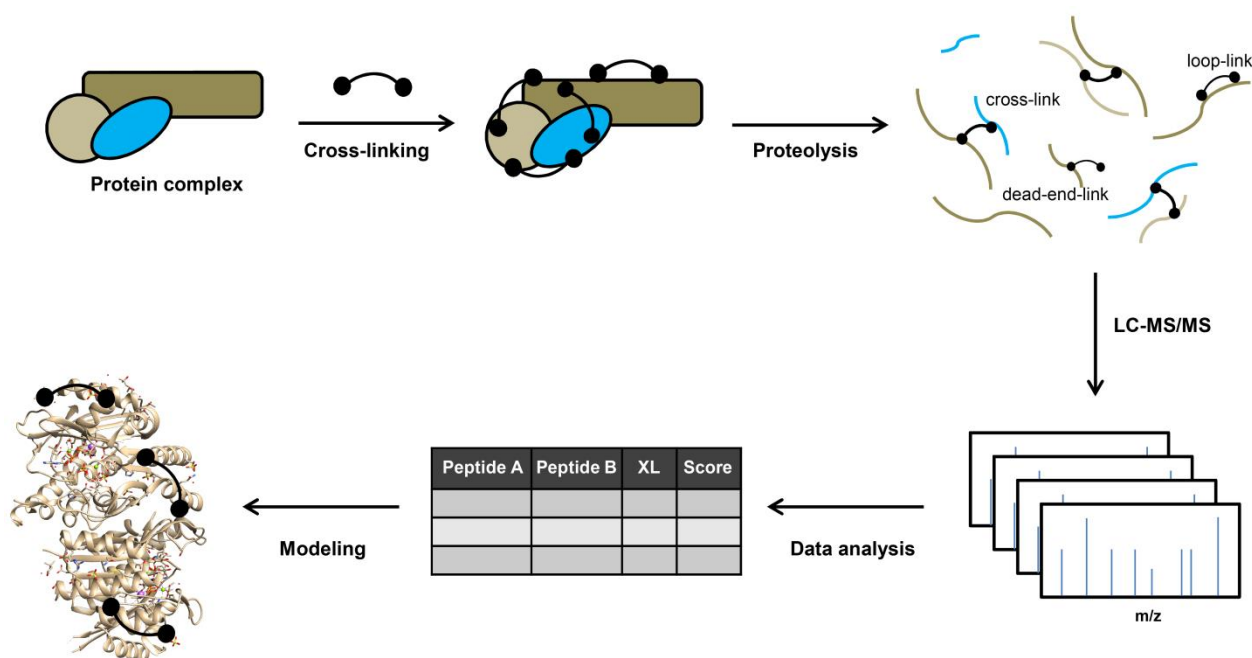
Figure I-1 Schematic representation of a general XL-MS workflow. A protein complex of choice is incubated with the cross-linking reagent to introduce covalent connections within the subunits. After proteolysis the peptide mixture consisting of several different species is analyzed by LC-MS/MS and data are further processed by suitable software tools. Finally, identified connections may serve as input for modeling or docking studies.

1.1.3 Applications of XL-MS

The most established application of XL-MS displays the investigation of enriched or purified protein complexes.[1] To date a manifold of data is published which focus on the elucidation of composition or interacting partners of protein complexes by *in vitro* chemical cross-linking. For instance, the architecture of multisubunit protein complexes like RNA polymerase II-TFIIF complex,[16] the human phosphatase 2A network,[20] or the eukaryotic chaperonin TRiC/CCT[21] have been (partly) solved by the cross-linking methodology.

In addition to purified protein systems, XL-MS can generate distance restraints in extremely complex biological samples like the entire proteome. For *in situ* studies of protein interaction partners, cross-linkers must be able to cross the cell membranes to fix intracellular proteins although generating highly complex mixtures in which the desired cross-linked peptides make up only a small fraction (presumably less than 0.1%).[22] Thereby, several studies have examined the

interactome of HeLa,[23] *Pseudomonas aeruginosa*,[24] *Acinetobacter baumannii*,[25] *Shewanella oneidensis*,[26] *Escherichia coli*[27] or *Bacillus subtilis*.[28] Another example is the elucidation of the interactome of purified mitochondria, revealing insights in the electron transport chain membrane complexes, the mitochondrial contact site and cristae organizing system, or supercomplex assembly of the oxidative phosphorylation complexes.[29,30] Further, host-microbe protein interactions during infection of lung epithelia with *A. baumanii* were monitored by XL-MS.[31] Thereby, interspecies contacts displayed that the key virulence factor is involved in host cell-to-cell adhesion. Recently, the first application of XL-MS to study interactions in tissue samples revealed for instance insights into multiple conformational states of sarcomere proteins.[32] However, a current limitation is that most detected cross-links from whole proteome XL-MS datasets are within the highly abundant protein complexes like the nucleosome, ribosome or proteasome.[33]

Determination of often transient and weak protein-protein contacts remains challenging, if not impossible. Two methods to identify interaction partners of proteins have been widely applied: co-immunoprecipitation (co-IP), which enriches the protein of interest and its surrounding using specific antibodies,[34–36] and affinity-enrichment were heterologously expressed proteins are fused in-frame to an affinity tag like a FLAG-tag enabling its selective pull-down.[37] Although those methods have been conducted widely and many of their results were confirmed with orthogonal experiments, their application requires cell lysis which disturbs the native environment of the cell and leads to loss or false identification of interaction partners.[26] Albeit combinations of affinity-based enrichment of protein complexes from intact cells with subsequent on-bead cross-linking have been applied, weak interactions will be lost during cell lysis as well as PPI reorganization can occur yielding in non-physiological interactions.[38] As *in situ* cross-linking previous to lysis preserves the native biological state and stabilizes transient interactions which would get lost upon cell lysis, a combination of co-IP or affinity-enrichment with *in situ* cross-linking seems beneficial. This method has been mostly used to study the proteoasome interactome.[39–41] Other studies combining *in situ* cellular cross-linking previous to lysis and co-IP focused on the outer membrane protein OmcA-MtrC in *Shewanella oneidensis*,[26] the telomere complex in embryonic kidney cells,[42] or the CD4 cell surface protein in hybridoma cells.[43]

XL-MS reveals information on structural composition and topologies of single proteins or protein complexes under various experimental conditions or biological states. The elucidation of such interactions in a quantitative manner would therefore confer an additional important layer of information. Quantitative XL-MS (QXL-MS) represents the next generation of XL-MS to enable the characterization of dynamic proteomes under different physiological states. Recent

application of QXL-MS involve application of isotope-labeled cross-linkers or labeling of the interacting proteins with heavy and light isotopes as in the *stable isotope labeling by amino acids in cell culture* (SILAC)-approach.[1,23,44] QXL-MS was successfully applied to study purified proteins and protein complexes,[45,46] as well as interactions within whole cells.[47–49]

1.1.4 Challenges and possibilities in XL-MS

Although XL-MS is nowadays widely applied and became for several questions the method of choice, the field is confronted with several challenges. First, the complexity of a cross-linked sample is markedly increased compared to a linear peptide mixture as several different types of peptides can occur (cf. Part I, Chapter 1.1.2).[10,11] As presumably less than 0.1% of all peptides within a sample display cross-linked ones,[22] there is an urgent need for enrichment. Therefore, several strategies have been developed, including the enrichment of cross-linked peptides *via* peptide size-exclusion chromatography (SEC) as they typically have a higher mass than unmodified ones.[50] Further, if the protease applied for digestion is trypsin, cross-linked peptides can be enriched *via* strong cation exchange (SCX) as they usually display a doubled charge state compared to a linear 2-fold positively charged tryptic peptide.[3,51] Another possibility is the application of trifunctional cross-linkers displaying an additional moiety for enrichment in the spacer. For example, the Protein Interaction Reporter (PIR) cross-linker developed by the *Bruce* group contains an additional biotin molecule at the spacer arm to allow for direct enrichment of cross-linked peptides (**Figure I-2B**, right).[47] However, cross-linkers directly carrying the affinity handle are more bulky which may affect their reactivity because of steric hindrance. Therefore, cross-linkers with a smaller alkyne or azide tag in the spacer region were developed allowing for subsequent bioorthogonal ligation using copper-catalyzed azide-alkyne 1,3-dipolar cycloaddition (CuAAC, click chemistry) to attach biotin after the cross-linking step.[52–55]

Further, the confident identification of cross-linked peptides is often hampered due to insufficient fragmentation and ion peak intensity of both peptides connected by the cross-linker. This comes especially into play when two peptides which differ markedly in length are cross-linked. Usually, the longer peptide is fragmented sufficiently, but the shorter partner is fragmented less, impeding its confident identification.[11,56,57]

Finally, the field of XL-MS is faced with an enormously increased search space when analyzing cross-linked peptide samples as each peptide can be in theory linked to any other peptide resulting in a squared database compared to the analysis of linear peptides.[10] In addition to the development and application of improved MS-methodologies for enhanced precursor

fragmentation as well as downstream bioinformatics and statistical analysis,[38] issues concerning sufficient fragmentation and increased search space can be addressed by the application of cleavable cross-linkers, displaying MS-labile bonds (cf. Part I, Chapter 1.1.5).[47,58]

1.1.5 MS-cleavable cross-linking reagents

MS-cleavable reagents have received considerable attention as the resulting cross-linked products can be identified due to their characteristic fragmentation behavior upon activation in the mass spectrometer.
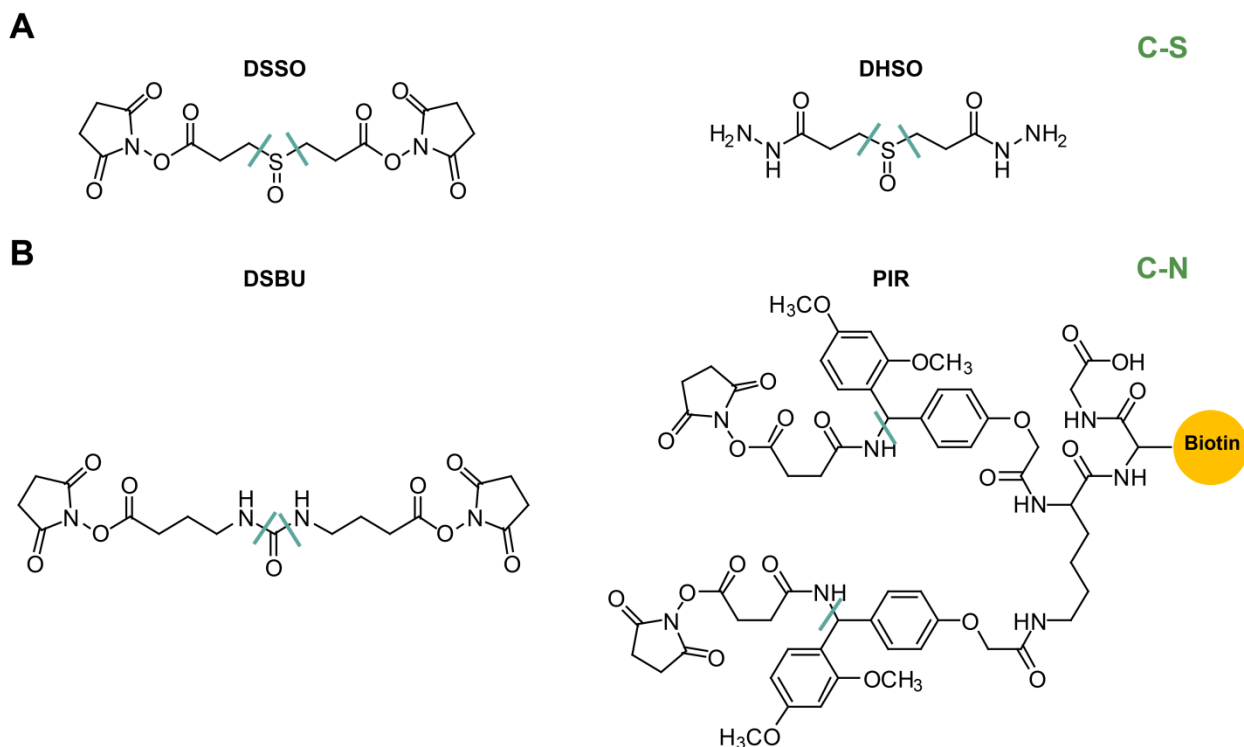
Figure I-2 Types of collision-induced dissociation (CID)-cleavable cross-linking reagents. (**A**) Applied are C-S bonds adjacent to sulfoxide moieties as in disuccinimidyl sulfoxide (DSSO)[58] and dihydrazide sulfoxide (DHSO)[59] and (**B**) C-N bonds as in disuccinimidyl dibutyric urea (DSBU)[60] and the PIR cross-linker.[48] Fragmentation of MS-labile bonds is highlighted with cyan lines.

Gas-phase cleavage sites can result in reporter ions,[61] single peptide chain fragment ions as the cleavable bonds are localized within the spacer region of the linker,[58,62] or both reporter and fragment ions.[47] MS-cleavable cross-linkers enable an easy correlation between separated cross-link peptide constituents and their respective parent ions. The generation of linear peptides upon cleavage in the mass spectrometer results in a database reduction from n^2 to $2n$ and therefore to a simplified data processing with accurate identification of cross-links. CID-

cleavable bonds are the most popular and widely applied MS-cleavable cross-linkers. One of the two commonly used cleavable bonds are C-S bonds adjacent to for instance sulfoxides as in the DSSO cross-linker,[58] a homobifunctional linker connecting two NHS-ester moieties or in DHSO connecting two hydrazides (**Figure I-2A**).[59] The second are C-N bonds for instance associated with urea like in DSBU[60] or associated with Rink, a compound commonly used for amidation in peptide synthesis, as applied in the PIR cross-linker (**Figure I-2B**).[48]

The ideal cross-linkers for multistage MS/MS (MS^n) analysis should possess MS-cleavable bonds that dissociate at lower collision energies where peptide bounds remain mainly intact. Such a property displays for instance DSSO with two symmetric CID-labile C-S bonds, which are cleaved during low-energy CID previous to MS^2 scanning, resulting in the separation of the cross-linked peptides yielding unique fragment pairs (α_A/β_S or α_S/β_A) with a defined mass relationship.[58] The resulting α and β peptide fragments are modified with complementary cross-linker remnant moieties, either an alkene (A) or a sulfenic acid (S), which often undergoes dehydration resulting in the more stable thiol moiety (T) (**Figure I-3**).

Full-MS mass fingerprinting / MS² cross-link fragmentation

DSSO cross-linked peptides
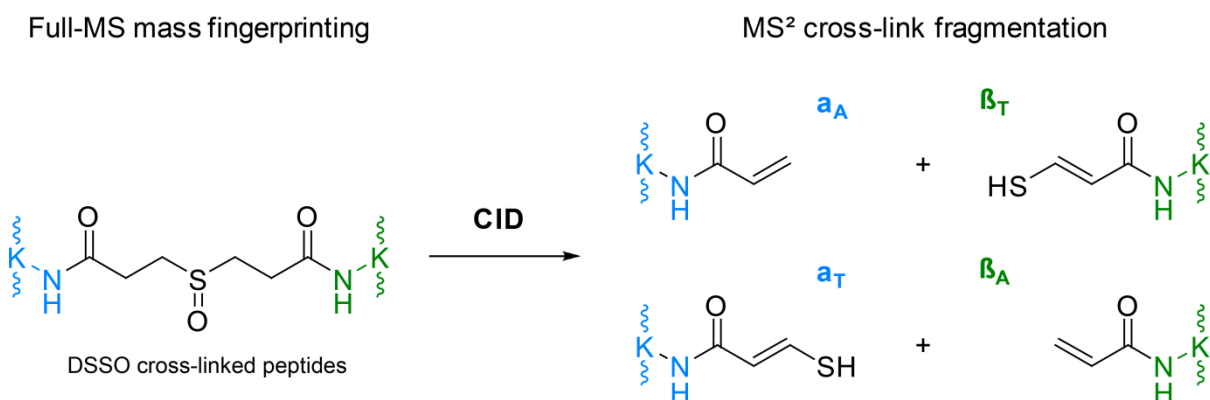
Figure I-3 Fragmentation of DSSO cross-linked peptides. Upon CID activation of the precursor, four characteristic signature ions α_A/β_T or α_T/β_A are detected in the MS² scan, consisting of the alkene and thiol modified version of each peptide.

These predictable MS² fragment ion pairs are subsequently fragmented in a next MS³ analysis, resulting in unambiguous identification of both cross-linked peptides. The false discovery rate (FDR) of cross-links is decreased due to incooperation of three lines of evidence: MS^1 scan for identification of the cross-linked precursor, MS² for fingerprint fragmentation, and MS³ for peptide sequencing. In addition, a targeted acquisition can effectively increase the number of identified cross-links. Thereby the mass difference of thiol and alkene modified peptides (31.97 Da) can be implemented in the mass spectrometer for selection for MS³ fragmentation. With the availability

of other fragmentation strategies, DSSO cross-linked peptides were additionally analyzed applying sequential CID-electron-transfer-dissociation (ETD). A combination of dual CID-ETD MS² fragmentation with CID-MS² and subsequent CID-MS³ acquisition was demonstrated to result in superior cross-link identification.[63] The XlinkX software was specifically developed for the computational analysis of XL-MS data resulting from DSSO cross-linked peptide mixtures fragmented with both CID-MS²-MS³ and CID-ETD-MS² (**Figure I-4**).[18]
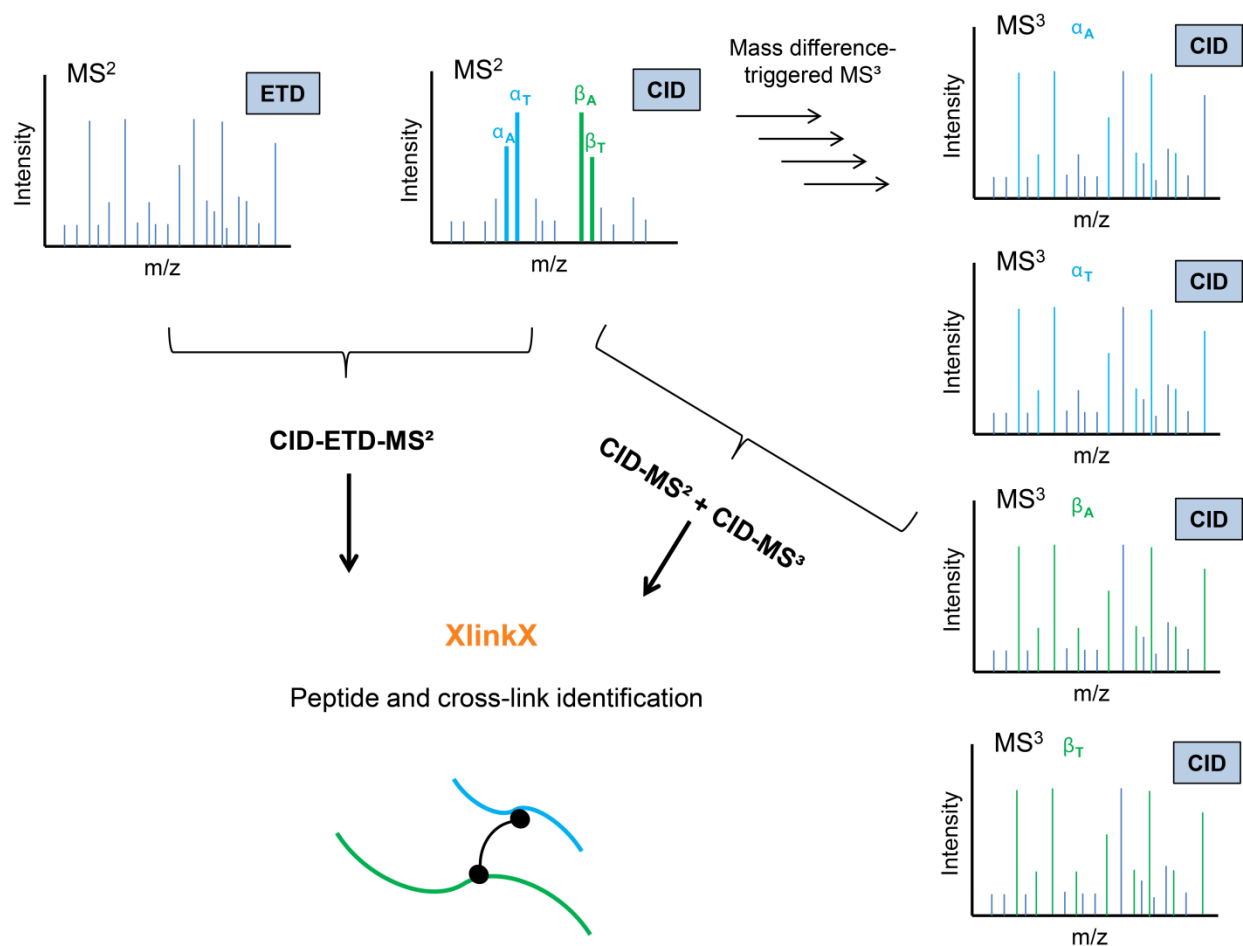
Figure I-4 MS analysis of DSSO cross-linked peptides. The precursor ion is fragmented either with both ETD and CID in the MS² stage, or only with CID-MS² resulting in the characteristic MS² peptide fragments α_A/β_T and α_T/β_A. A subsequent mass difference-triggered MS³ fragmentation allows confident peptide identification upon analysis with XlinkX.[18,58,63]

Collectively, MS^n-based workflows with MS-cleavable reagents are simple, robust, and effective for mapping PPI even in complex samples like the global proteome.[18]

1.2 The multicomponent protease complex ClpXP

In order to tightly regulate protein levels as well as their activity, the cell expresses a set of proteases playing important roles in protein quality control and homeostasis. Proteases are essential for the removal of short-lived regulatory, as well as misfolded or damaged proteins. A certain proportion of protein degradation is carried out by oligomeric, cylindrical proteases. One member of those proteases is caseinolytic protease P (ClpP), a highly conserved serine protease which is expressed throughout bacteria and eukaryote.[64] ClpP was first described in *E. coli* in 1987.[65] Since then, several crystal structures of the protease have been solved which all display the protease to be comprised of 14 identical subunits arranged into two heptameric rings stacked back-to-back to form the cylindrical structure enclosing a chamber containing the Ser-His-Asp catalytic triad.[66] ClpP typically forms compelxes with ATPases associated with various cellular activities (AAA+), which function as chaperones that recognize and denature substrates under ATP hydrolysis to guide them into the ClpP catalytic chamber for degradation.[64] The most prominent AAA+ protein interacting with ClpP is ClpX, which assembles into a hexameric ring in its active state that is associated with ClpP either on one or both ends to build the active protease complex ClpXP (**Figure I-5**).[67] Without its ATPase components only small peptides can enter the ClpP chamber and be degraded.[68–70]
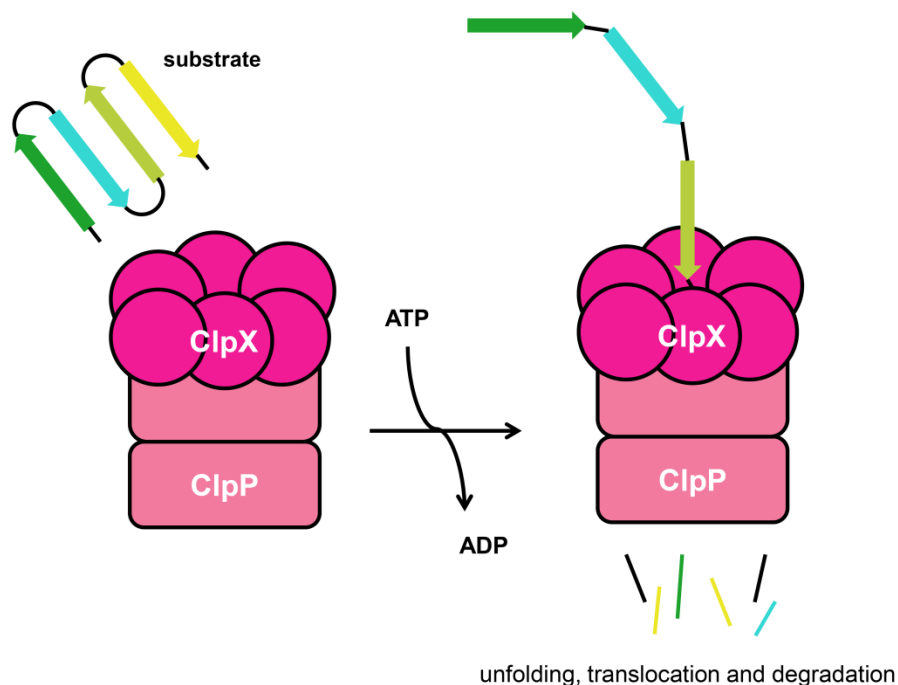
Figure I-5 Schematic representation of the ClpXP protease complex. The tetradecameric ClpP peptidase is associated with the hexameric ATPase ClpX at either one or both ends. Substrates are

recognized, unfolded and translocated into the ClpP proteolytic chamber by ClpX under the consumption of ATP.

1.2.1 The human ClpP ortholog

In the cytosol and nucleus of eukaryotes, the major protein-processing enzyme is the 26S proteasome. However, this complex is absent in organelles as the mitochondria, where other proteases contribute to protein housekeeping. One of those proteases is the human ClpP (hClpP) ortholog, which is solely present in the mitochondrial matrix, where it contributes to mitochondrial protein homeostasis together with the two other AAA+ superfamily proteases Lon and matrix-oriented AAA (*m*-AAA) proteases.[71] Each of these proteases are nuclear encoded and are imported into mitochondria *via* a targeting sequence. Unlike most bacterial orthologs that exist as double-ring tetradecamers, hClpP exists as heptamer under physiological conditions that assembles into the tetradecameric form in the presence of ClpX.[72] Although the ClpXP complex is one of the least characterized mitochondrial proteases and there are no confirmed endogenous mammalian ClpXP substrates, the complex is hypothesized to be involved in the mitochondrial unfolded protein response (UPR^{mt}).[73] Further, genetic mutations in hClpP cause diseases like Friedreich's Ataxia[74] and Perrault syndrome[75]. Moreover, hClpP was shown to play an important role in cell growth[76] as well as tumor cell proliferation,[77,78] invasion and metastasis.[79] It is overexpressed in different tumor types, *e.g.* acute myeloid leukemia, prostate cancer, sarcomas and breast adenocarcinoma.[77,79] Furthermore, genetic disruption or chemical inhibition of hClpP reduces the viability of leukemia cells.[77] However, the exact role of ClpP in the human cell remains unknown.

1.2.2 Bacterial ClpP orthologs

ClpP orthologs are present in all bacteria sequenced to data except for *Mollicutes*.[80] The major part of bacteria contains one copy of ClpP, but there are some strains that possess two forms of ClpP, for instance ClpP1 and ClpP2 expressed in *Listeria monocytogenes* and *Mycobacterium tuberculosis*.[81] While *L. monocytogenes* ClpP2 (LmClpP2) resembles the relative to other bacterial orthologs and can form homo-tetradecameric assemblies, LmClpP1 shares only 44% sequence identity with EcClpP and predominantly forms inactive heptamers. Further, LmClpP1 lacks an N-terminal ATPase binding motif and exhibits a truncated catalytic triad in which the Asp is replaced by an Asn residue.[82] Negative stain EM displayed the existence of hetero-tetradecameric complexes formed by LmClpP1 and LmClpP2 heptamers (LmClpP1/2).[81] Moreover, it was found that LmClpP1 is only active when complexed with LmClpP2.[82] Although a

crystal structure of LmClpP1/2 is available, it is unclear if both parts of the heterocomplex can bind LmClpX or if the interaction is only mediated *via* one side.[83] Contrarily to hClpP which was only found to be associated with ClpX, bacterial ClpPs do not only form complexes with ClpX, but also with other AAA+ proteins as for instance *E. coli* ClpP (EcClpP) with ClpA[64] or *Staphylococcus aureus* ClpP (SaClpP) with ClpC.[84] An additional association with other chaperones may allow for a broader substrate scope and recognition as well as further regulation mechanisms.[85,86] Interestingly, several bacterial ClpP orthologs like the food-borne pathogen *L. monocytogenes*[87] and *S. aureus*[88–90] were demonstrated to be virulence-associated. Bacterial ClpPs display therefore substrates of current research including the development of specific inhibitors.[91–94]

1.2.3 Structural features of the protease

To date, several structures of ClpP orthologs have been solved including the ones from *E. coli*,[95] *S. aureus*,[96] *L. monocytogenes*,[82,83] *M. tuberculosis*[97] and human.[98] The overall ClpP structures display a high similarity indicating that the protease conformation is conserved throughout evolution. Although ClpP is structural studied extensively, only two crystal structures of ClpX orthologs have been determined, one displaying a continuous spiral formed by ClpX subunits of *Helicobacter pylori* in the lattice,[99] the other displaying covalently linked subunits of *E. coli* ClpX (EcClpX).[100] The association of ClpX, that forms a hexamer in solution, with a ClpP heptameric ring leads to a symmetry mismatch. As no high-resolution structure of the overall ClpXP protease is present, the exact binding mode remains a riddle. However, several studies have focused on the elucidation of regions or loops important for ClpXP assembly. Mutations in a tripeptide motif [LIV]-G-[FL] located in ClpX flexible loops disrupted the ClpXP complex formation.[101] Docking and mutagenesis studies suggested that this tripeptide motif interacts with a hydrophobic depression on the ClpP surface.[66,99] Further, mutational studies revealed ClpP N-termini to play important roles in complex formation. Disulfide cross-linking, as well as insertion of point mutations revealed that the ClpP N-termini are in contact with loops at the bottom of the central ClpX channel (pore-2 loops) (**Figure I-6**).[66] Contacts between ClpX pore-2 loops and ClpP N-termini are declared as more dynamic, whether interactions of ClpX IGF-loops with ClpP hydrophobic pockets are rather static.[102]
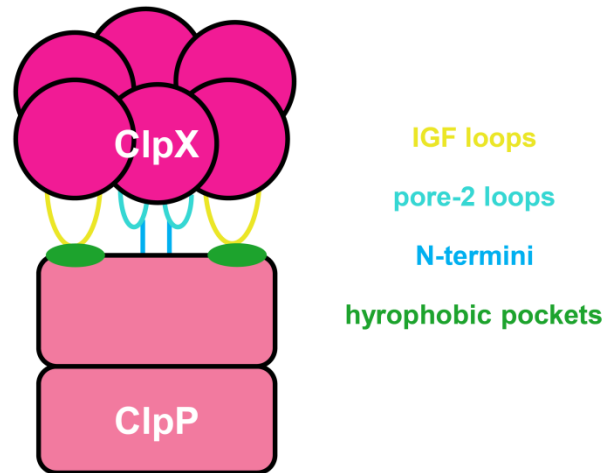
Figure I-6 Essential interactions for assembly of the ClpXP complex. IGF- and pore-2 loops in ClpX are in contact with hydrophobic pockets and N-termini of ClpP, respectively.

1.2.4 Cellular interaction network and substrate scope of ClpP orthologs

Though structural features of ClpP are well-studied, less is known about its cellular binding partners and substrates. Analysis of potential cellular substrates, as well as expression changes associated with ClpP perturbations have been mainly investigated by MS. For instance, a global proteome analysis upon inhibition[89] or deletion of ClpP from *S. aureus*[103] revealed perturbations of expression levels of major virulence regulators like RNAIII. The main strategy for identification of cellular ClpP substrates is the application of a proteolytic inactive protease, which serves as a molecular trap that will retain but not degrade substrates translocated into its protoelytic chamber. Thereby, affinity-tagged inactive ClpP orthologs were heterologously expressed and subsequently co-purified interactors identified by MS. Several potential substrates were discovered for EcClpP under normal physiological conditions[104] as well as under DNA damage.[105] Further, an inactive ClpP trap revealed potential substrates of *Caulobacter crescentus* ClpP[106] and SaClpP.[107] Pull-downs of inactive mammalian ClpP resulted in co-purification of for instance mitochondrial respiratory chain proteins, proteins involved in translation, or chaperones.[77,108,109] Taken together, these studies provide first insights into cellular protease interaction networks as well as suggest potential *in situ* substrates. However, as all strategies require genetic manipulation of the studied cells, as well as artificial working conditions including inactive proteases, an investigation of cellular binding partners under native conditions is of urgent need.

1.3 Strategies and Objectives

Although the multi-component ClpXP protease is subject of research for decades, several questions remain unanswered. Albeit the ClpP peptidase compartment has been crystallized from several organisms, a high resolution structure of the whole ClpXP complex is still missing. Knowledge about structural regions essential for interaction of ATPase and peptidase stems exclusively from mutational and modeling studies. XL-MS became one of the methods of choice for investigation of multimeric complexes and protein assemblies like the proteasome.[39] Our goal was the elucidation of the enigmatic ClpXP interface by chemical cross-linking which has not been performed so far. We aimed at the development of a robust and effective cross-linking workflow applying the MS-cleavable cross-linker DSSO. For this purpose we chose ClpP model systems from major bacterial classes as *S. aureus*, *E. coli*, and *L. monocytogenes*, as well as the human ortholog. After protein cross-linking and removal of excess cross-linker by SEC, proteins were digested by trypsin and cross-linked peptides were enriched by SCX. After LC-MS/MS and a confident downstream data analysis to identify cross-linked residues, we aimed at mapping them onto existing crystal structures for validation (**Figure I-7**).
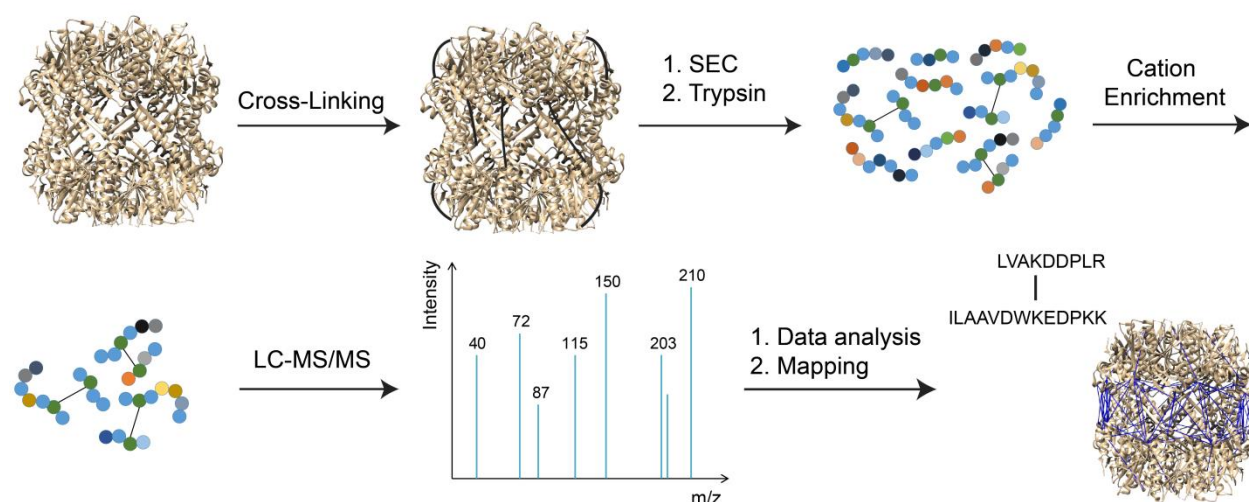
Figure I-7 Overview of the chemical cross-linking strategy for elucidation of ClpP composition. After ClpP cross-linking, excess of linker was removed by SEC and proteins were digested with trypsin. Following a cation enrichment step and LC-MS/MS measurement, data were analyzed and cross-links subsequently mapped onto ClpP crystal structures for validation.[110]

With a robust cross-linking strategy in hand, our next goal was the analysis of the ClpXP interface to identify residues in close proximity. Information from cross-linking experiments could

further serve as restraints for docking studies, to build a first model of the ClpXP system derived from structural data.

Moreover, we aimed at the investigation of the cellular binding partners of ClpP. Therefore, we combined *in situ* chemical cross-linking with co-IP (cross-link/co-IP) applying cell-permeable DSSO to stabilize weak and transient PPI that would get lost upon cell lysis (**Figure I-8**).[26] Compared to previous ClpP pull-down studies,[77,104,107] we aimed at examining the interaction partners under native conditions without any genetic manipulation of the cells. We planned to compare cross-link/co-IP with co-IPs from non-cross-linked cells, as well as one-bead cross-linking to examine stabilities of certain interactions under different experimental conditions.
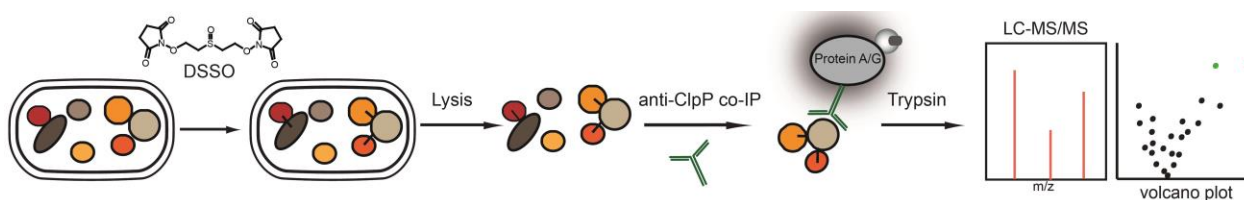
Figure I-8 *In situ* chemical cross-linking combined with co-IP (cross-link/co-IP). The DSSO cross-linker is applied to intact cells, followed by lysis and incubation with the anti-ClpP antibody for pull-down of cross-linked complexes. Tryptic peptides are measured by LC-MS/MS and evaluated by downstream statistical analysis.[110]

In a last step we aimed at the combination of activity-based protein-profiling (ABPP, cf. Part II, Chapter 1.4),[111,112] where functionalized small molecules are applied onto the proteome to enrich for the specific interaction partners or targets of the probe, with cross-linking. Such a combination has never been reported so far, but nevertheless displays a promising strategy to enrich not only the probe targets, but also their subset of cellular interactors in one experiment. Therefore, we chose *β*-lactones as selective inhibitors of bacterial ClpP[88,90,94] to test this strategy by the application of varying experimental conditions.

2 Validation of the cross-linking method with different ClpP orthologs

Our first goal was the set-up of an appropriate cross-linking workflow that efficiently captures the oligomeric structures of ClpP complexes for subsequent MS analysis.

For proof of principle studies, we selected the homo-tetradecameric ClpP complexes of *S. aureus* due to its role in virulence regulation[89] and *E. coli* as one of the best characterized ortholog for reference.[86] As mentioned previously, *L. monocytogenes* encode two isoforms of ClpP.[81,82] As biological reasons for that are not clear, we further chose the hetero-tetradecameric ClpP1/2 complex of *L. monocytogenes* for validation of our workflow, as well as the cancer associated human ortholog,[77] which forms homo-heptamers *in vitro.*[72] As all orthologs have been crystalized previously they display excellent candidates for validation studies.[66,83,96,98]
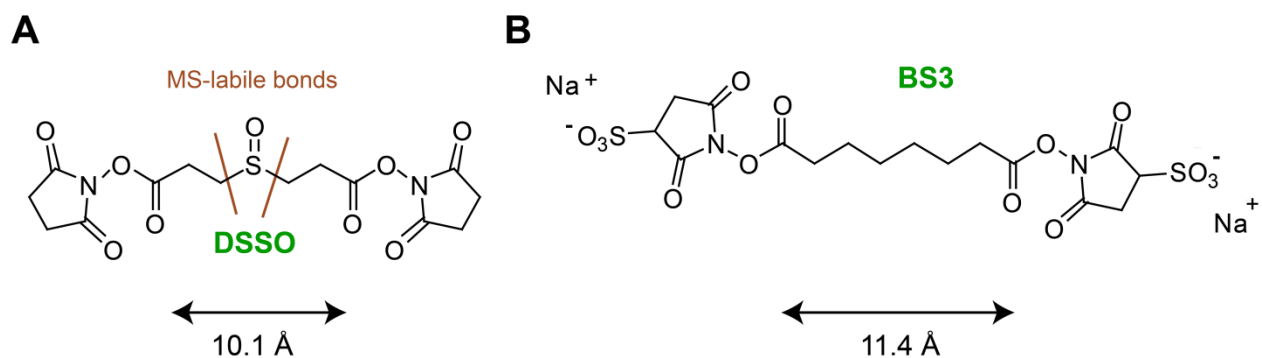
Figure I-9 Chemical structures of DSSO (A) and BS3 (B) in comparison. For both cross-linkers the spacer arm lengths are indicated.

For our cross-linking studies we chose DSSO with a 10.1 Å spacer arm length targeting primary amines (**Figure I-9A**), as well as BS3 (**Figure I-9B**). As mentioned above, DSSO contains an MS-labile C-S bond in the spacer arm which is preferentially cleaved upon low collision-energy in the mass spectrometer to generate linear peptides.[58] This feature ensures equal fragmentation efficiency of the cross-linked peptides and simplifies bioinformatic analysis.[18] The cross-linker was synthesized following literature methods with slight modifications.[58]

2.1 Gel-based analysis of cross-linked ClpP orthologs

First, cross-linking efficiency was optimized for all four ClpPs by variation of DSSO concentration, reaction time and temperature.
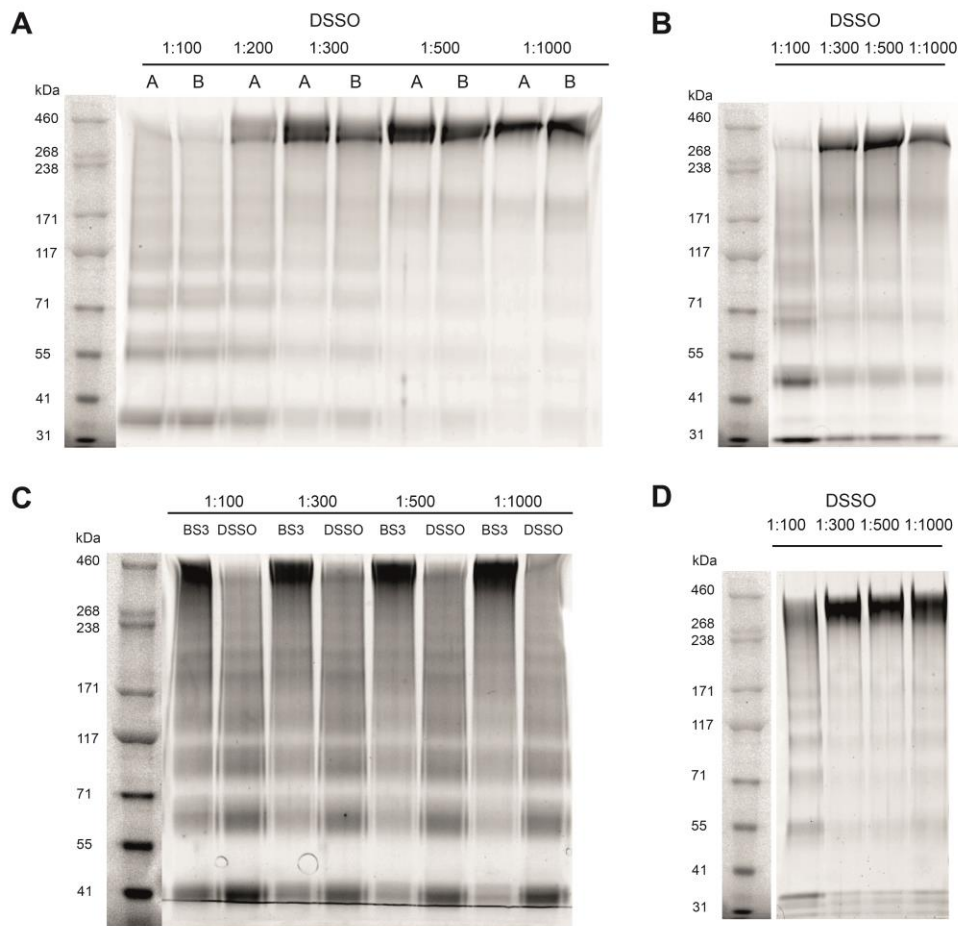
Figure I-10 Analysis of cross-linked ClpP orthologs by SDS-PAGE. Bacterial tetradecamers have a molecular weight of approx. 320 kDa, the human ortholog of approx. 360 kDa. (**A**) SaClpP cross-linked with 100- to 1000-fold excess of DSSO cross-linker over ClpP monomer. Reaction carried out at 37 °C for 2 h [A] or overnight at 4 °C [B], respectively. (**B**) LmClpP1/2 cross-linked with 100- to 1000-fold excess of

DSSO cross-linker over ClpP monomer at 37 °C for 2 h. (**C**) EcClpP cross-linked with 100- to 1000-fold excess of DSSO or BS3 cross-linker over ClpP monomer. Reaction carried out at 37 °C for 2 h. (**D**) hClpP cross-linked with 100- to 1000-fold excess of DSSO cross-linker over ClpP monomer at 37 °C for 2 h. In all cases, cross-linkers shift the equilibrium to tetradecameric species.[110]

Separation of treated proteins by SDS-PAGE revealed that a 300-fold molecular excess of DSSO over ClpP monomer and a reaction time of 2 h at 37 °C efficiently cross-linked the tetradecamers of SaClpP and LmClpP1/2 (**Figure I-10A** and **B**). In the case of hClpP, which predominantly constitutes a heptamer *in vitro*, the equilibrium was shifted to the tetradecameric state upon DSSO addition (**Figure I-10D**). For EcClpP, on the other hand, the tetradecameric state could not be efficiently fixed by DSSO, as lower molecular weight oligomeric species were predominantly observed (**Figure I-10C**). Nonetheless, use of the longer, non-cleavable bis(sulfosuccinimidyl)suberate (BS3, 11.4 Å spacer arm length, **Figure I-9B**) cross-linker, which targets primary amines like DSSO, led to the appearance of the tetradecameric EcClpP complex (**Figure I-10C**). Comparison of lysine distributions in the crystal structures of EcClpP, SaClpP and LmClpP1/2 revealed an overall lower lysine density at the heptamer-heptamer interface of EcClpP (for example, no lysine residues in F- or G-helices) resulting in a lack of shorter cross-link distances (**Figure I-11**). Therefore, a longer cross-linker like BS3 could have been able to bridge the longer lysine-lysine distances in the range from 17 to 18 Å (end-to-end distance) and thereby to covalently fix the tetradecameric state.

| LmClpP1/2 | SaClpP | EcClpP |
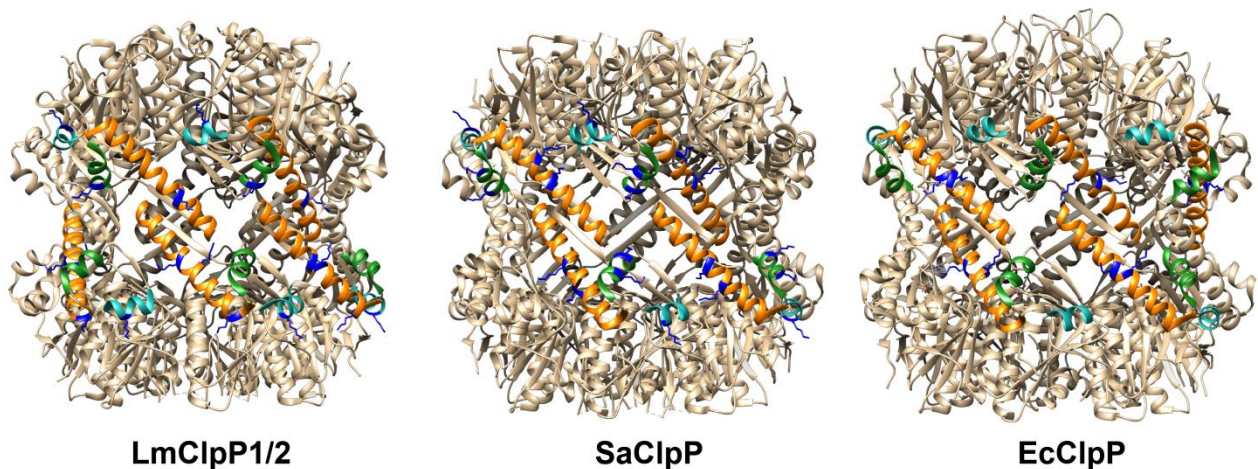
Figure I-11 Comparison of ClpP heptamer-heptamer interfaces. A comparison of LmClpP1/2 (PDB 4RYF), SaClpP (PDB 3V5E), and EcClpP (PDB 1YG6) reveals a lower lysine density (blue) at the heptamer-heptamer interface for EcClpP as lysines are only present in the E-helix (orange), but are missing in F- (green) and G-helices (cyan).

2.2 MS-based analysis of cross-linked ClpP orthologs

Next, we purified DSSO- and BS3-treated ClpP complexes by SEC to remove excess of cross-linker and to ensure correct integrity after cross-linking. Satisfyingly, we observed similar shaped peaks after cross-linking comparable to non-cross-linked ClpP orthologs (**Figure I-12**). Notably, the retention time of cross-linked complexes was shifted about 1 mL. Earlier elution volumes could be caused by the additional mass of the cross-linking molecules or altered surface properties influencing the interactions with the SEC column material. In addition, non-cross-linked hClpP moved at comparable elution volumes as bacterial orthologs, though it is present as heptamer. The unusual high mobility of hClpP during gel filtration was reported previously, probably caused by a more asymmetric shape and larger hydrodynamic radius.[72] However, although cross-linked hClpP was demonstrated to be a tetradecamer by SDS-PAGE analysis, the SEC retention volume was only slightly shifted compared to the non-cross-linked control as observed for the bacterial orthologs.
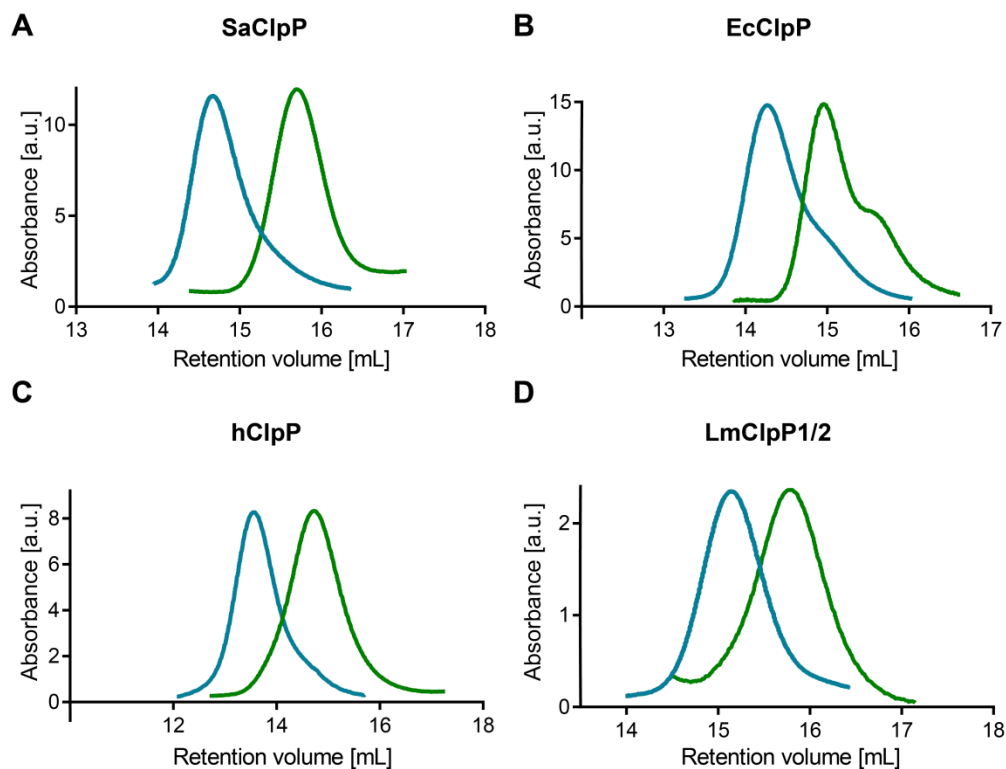
Figure I-12 SEC analysis of ClpP orthologs. Comparison of non-cross-linked ClpP complexes (green) with cross-linked ClpP complexes (blue) reveals a similar elution profile. Retention times of cross-linked orthologs were shifted to earlier elution volumes.

Protein containing fractions after separation by SEC were combined and evaporated *in vacuo*. Afterwards, proteins were re-dissolved, reduced, alkylated and subsequently digested in solution using trypsin. Cross-linked peptides were enriched *via* SCX, on the basis of their higher charge after tryptic digest compared to not cross-linked peptides since they make up only a small fraction of the total peptide pool (presumably less than 0.1%).[22]
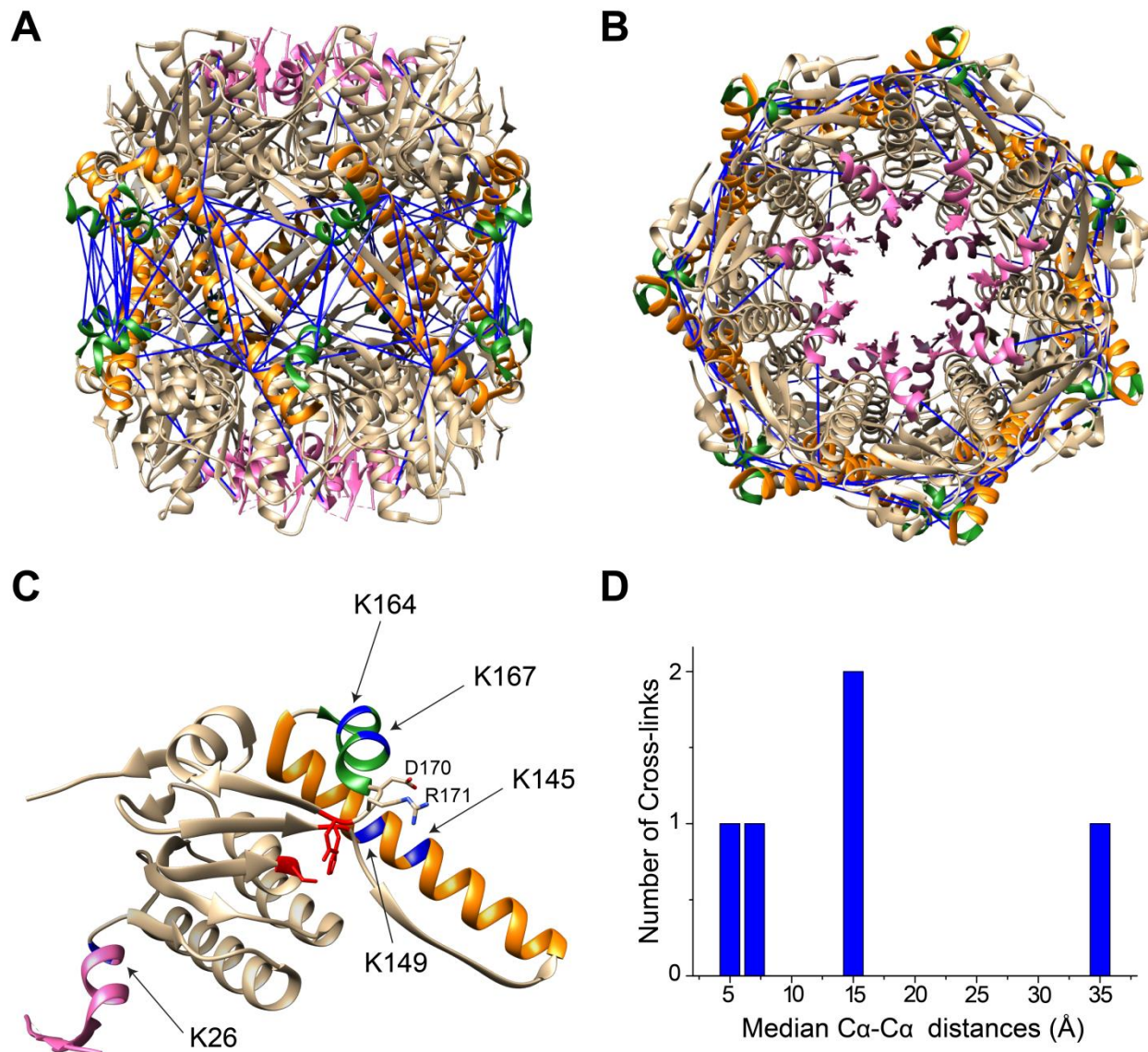
Figure I-13 SaClpP was cross-linked with the MS-cleavable linker DSSO. Identified cross-links (FDR <0.01) were mapped onto the crystal structure of SaClpP in the active, extended conformation (PDB: 3V5E) represented in the side-view (**A**) and the top-view (**B**). (**C**) Residues involved in cross-linking contacts are highlighted in blue in one SaClpP monomer. Residues of the catalytic triad are colored red. Further, oligomerization sensors D170 and R171 are labeled. E-helices are colored orange, F-helices

green, and N-terminal regions pink. (**D**) The histogram of median Cα-Cα cross-link distances reveals that all contacts are within the maximum distance span of DSSO (35 Å).[110]

Each sample was measured twice by LC-MS/MS, once performing electron-transfer/higher-energy collision dissociation (EThcD)-CID dual MS[2] fragmentation and once performing CID-MS[2] followed by higher-energy collision induced dissociation (HCD)-MS[3] fragmentation for MS-cleavable cross-linkers (**Figure I-4**). The use of these complementary MS fragmentation strategies revealed superior cross-link coverage (34% more unique cross-links when combined) compared to samples measured only with one type of fragmentation.[63] We could also confirm the benefit of using both fragmentation methods, as several cross-links would have otherwise escaped detection. For example, cross-link between K26 and K193 of LmClpP2 was only detected upon MS[2]-MS[3] fragmentation, but was missing upon EThcD fragmentation (for more examples see also **Tables S1**-**7**). Additionally, we prepared MS samples using different cross-linker concentrations (300-, 500-, and 1000-fold excess over ClpP monomer) in order to further enhance the number of detected cross-links. DSSO cross-linked sites were identified using the XlinkX software,[18] and BS3 cross-linked sites with the Kojak software (**Table S1-4**).[113] Residues identified as cross-linked by DSSO or BS3 were mapped onto corresponding ClpP crystal structures published previously as described in the experimental section (**Figure I-13** and **S1-4**).[83,95,98,114] Considering the length of the cross-linkers (10.1 Å for DSSO and 11.4 Å for BS3) and of the lysine side chains as well as backbone dynamics, we assumed that lysine residues within a Cα-Cα distance of up to 35 Å for DSSO and 36 Å for BS3 would be preferentially cross-linked.[39] All detected cross-links of the orthologs fit this Cα-Cα distance restraint.

For all cross-linked bacterial orthologs the N-terminal lysine (K26 in SaClpP and LmClpP2, K25 in EcClpP, and K24 in LmClpP1) was identified to be particularly reactive to the linker. As this residue is localized near the entrance pore of the tetradecamer the good accessibility points to a distinct role. Indeed, N-terminal residues of the protease were shown to be important for the ClpX interaction[115] and to contribute to the narrow axial pores, which control access to the catalytic chamber.[64] The protein sequence of the human ortholog lacks a corresponding N-terminal lysine residue and therefore no cross-links in this region could be identified. In general, hClpP displayed only about 50% sequence identity compared to the bacterial orthologs. Nevertheless, several residues of the E-helical handle region near the heptameric interface were involved in cross-linking contacts in all ClpPs. These helical regions are proposed to build pores for substrate release.[116] The flexibility of the N-terminal loops and the E-helices was also demonstrated by comparison of crystal structures of SaClpP in the extended and compressed conformations. Here, the N-terminal loops change from a straight to a completely unstructured

orientation and the handle helices also display substantial differences in configuration (**Figure I-13** and **S4**).[96] Moreover, two lysines (K164 and K167) from SaClpP formed cross-linking contacts within the F-helix and to the adjacent E-helices. The F-helix directly precedes the oligomerization sensors D170 and R171 which form a salt bridge network necessary for tetradecamer formation.[114] As compressed and extended states are believed to be in an equilibrium, this region might be more accessible to DSSO. These results emphasize that structurally flexible regions are efficiently targeted by the cross-linker.

2.3 Conclusion

Taken together, we were able to set up an efficient and robust chemical cross-linking workflow to study different ClpP orthologs. We applied the MS-cleavable DSSO cross-linker as it enables a superior fragmentation and therefore a more confident identification of both cross-linked peptides.[58] First, we examined cross-linking products by separation with SDS-PAGE. Satisfyingly, SaClpP, LmClpP1/2, and hClpP were covalently fixed in their tetradecameric states upon DSSO addition. In contrast, EcClpP could not be captured in its tetradecameric state probably caused by a lack of lysine residues in the heptamer-heptamer interface when compared with the other orthologs. However, the longer BS3 cross-linker could efficiently fix the whole EcClpP complex and was therefore selected for downstream MS-analysis. Digested peptides were enriched by SCX and subsequently measured by LC-MS/MS analysis using both fragmentation strategies for DSSO (EThcD-CID dual MS^2 fragmentation and CID-MS^2 followed by HCD-MS^3 fragmentation).[63] Data analysis with the XlinkX[18] or Kojak[113] software, respectively, revealed several cross-linking contacts for all orthologs. Subsequent mapping of cross-links onto corresponding crystal structures revealed all contacts to be within the maximum distance span of the cross-linkers.

3 Insights into the enigmatic ClpX-ClpP interface

Although the ClpXP complex was studied for decades, no high resolution structure of the holo-complex is available to date. With validated cross-linking tools in hand, we therefore utilized the methodology for identifying residues involved in the binding between ClpX and ClpP.

3.1 Optimization of cross-linking conditions for ClpXP orthologs

ClpP and the Walker-B mutant of ClpX, which can bind but not hydrolyze ATP,[117] were incubated to generate the ClpXP complex. Since recombinant EcClpP turned out to be not predominantly fixed as tetradecamer upon DSSO addition, cross-linking of ClpXP was only performed with the *S. aureus*, *L. monocytogenes* and human orthologs. As SDS-PAGE analysis was not suitable for resolving the high molecular weight ClpXP complexes due to the lack of standards (for example, SaClpXP = 870 kDa), cross-linked reaction products were separated by SEC and elution profiles were compared to not cross-linked ClpXP complexes to verify correct integrity. Not cross-linked ClpXP complexes require ATP to stay intact (**Figure I-14A** and **B**). We therefore separated the cross-linked products in the absence of ATP to ensure a covalent fixation. The amount of DSSO as well as the reaction time was exemplarily optimized for SaClpXP. An 80- to 150-fold excess of cross-linker over ClpP monomer and an incubation time of 2 h at 37 °C turned out to cross-link the SaClpXP complex very efficiently (**Figure I-14B** and **C**). Also an overnight incubation with 150-fold excess of DSSO trapped the complex (**Figure I-14D**). In contrast, 40-fold excess was not sufficient to fix SaClpXP (**Figure I-14E**), whereas a 2 h incubation with 800-fold excess of DSSO led to complete aggregation of the protein (**Figure I-**

14F). As already observed for the ClpP orthologs, retention times of cross-linked complexes compared to not cross-linked ClpXP shift to earlier elution volumes. Again, this may be caused by the mass addition due to DSSO cross-linking, or changes in the shape and surface of the proteins which might influence their SEC running behavior. Protein containing fractions (80- and 150-fold excess, 2 h at 37 °C reaction) were combined and evaporated *in vacuo* for subsequent LC-MS/MS preparation.
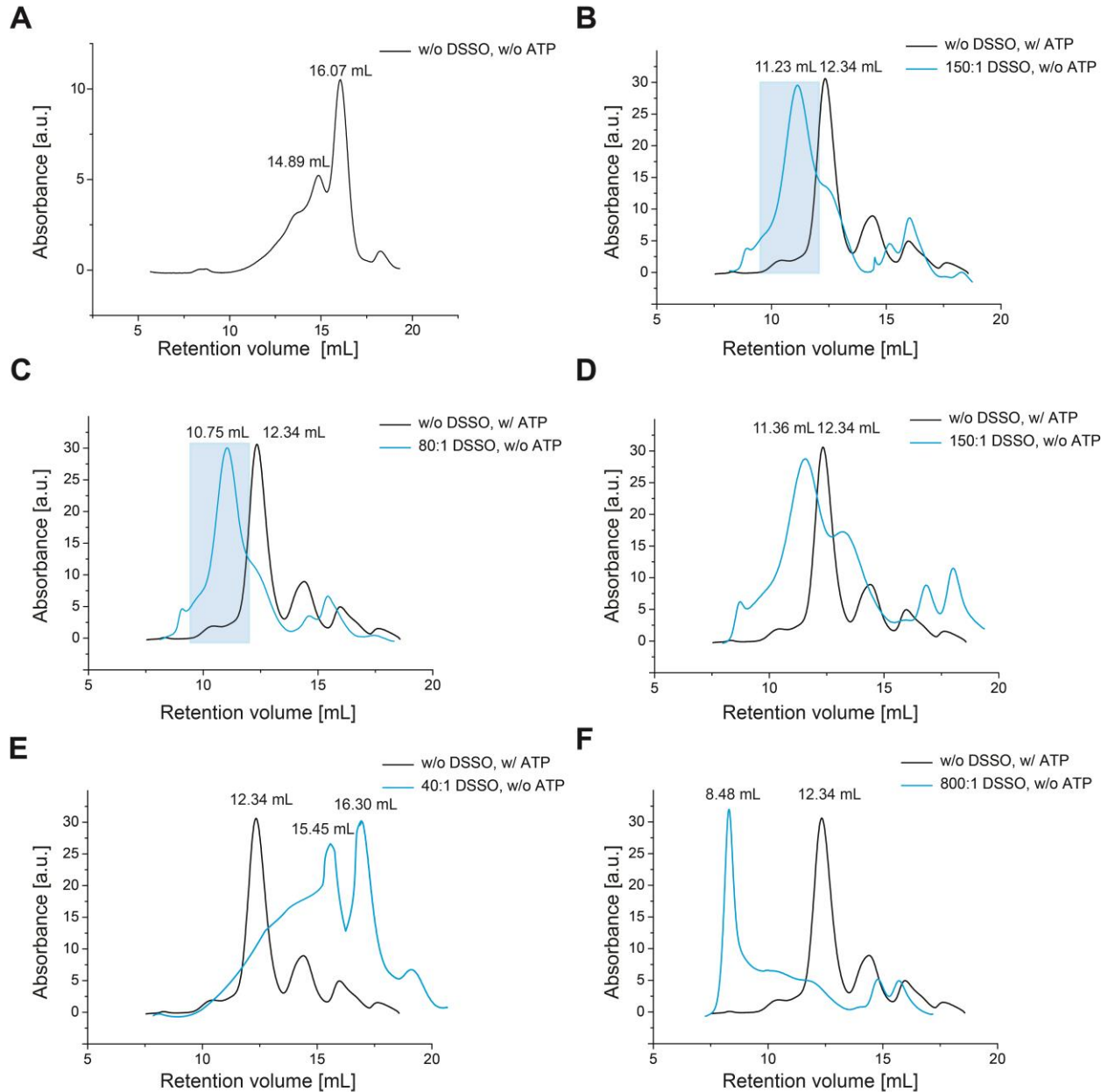
Figure I-14 Cross-linking of SaClpXP and separation by SEC. (**A**) Not cross-linked ClpXP complexes require ATP in the SEC running buffer to stay intact. Therefore, cross-linked proteins (blue lines) were

separated in a SEC running buffer lacking ATP to demonstrate a covalent fixation upon DSSO addition. In order to verify complex integrity, cross-linked ClpXP elution profiles are compared to a not cross-linked ClpXP sample separated in SEC running buffer containing ATP (black lines). A 150- (**B**) or 80-fold (**C**) excess of DSSO over SaClpP monomer and a reaction time of 2 h at 37 °C are sufficient to trap the SaClpXP complex (SaClpXP elutes at 12.34, SaClpP at 15.50 and SaClpX at 14.60 mL as determined by comparison to a molecular weight (MW) standard). (**D**) Incubation with 150-fold excess of DSSO at 4 °C overnight. (**E**) Lower amounts of DSSO were not sufficient to covalently fix the complex. (**F**) A higher molecular excess of DSSO as 800-fold led to complete aggregation of the proteins. Highlighted regions were pooled for MS sample preparation.[110]

The optimized reaction conditions were successfully transferred to hClpXP and LmClpXP1/2 as determined by SEC (**Figure I-15A** and **B**). Compared to other ClpXP complexes the yield of cross-linked LmClpXP1/2 was comparatively low. The different retention volumes of ClpXPs already indicated altered complex compositions. SaClpXP, as well as hClpXP eluted at around 12 mL, whereas LmClpXP1/2 eluted at around 13.9 mL. The shift in retention volumes points to a lower molecular weight of the LmClpXP1/2 complex, which may be caused by the presence of only one hexameric ClpX, whereas two hexamers may be associated with SaClpP and hClpP.
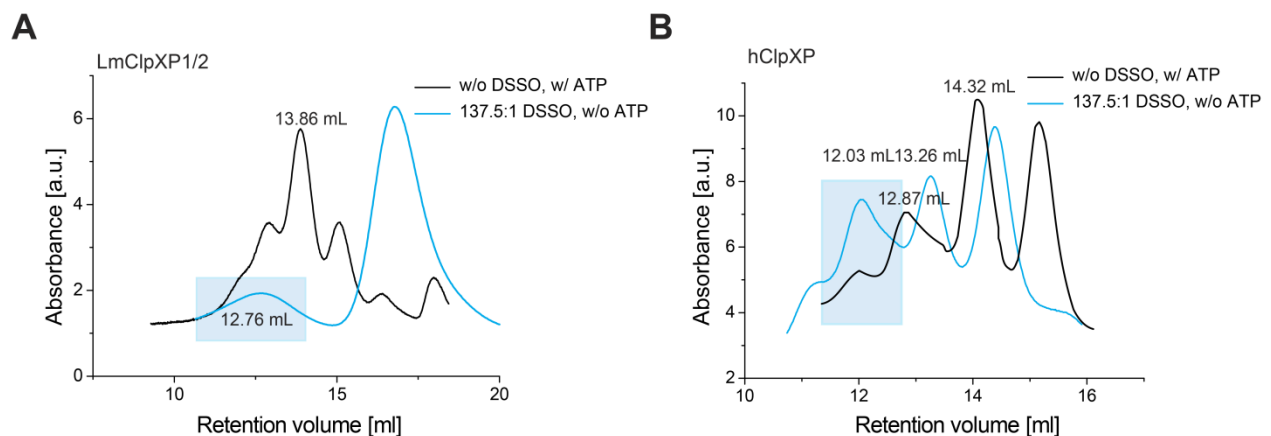
Figure I-15 Cross-linking of LmClpXP1/2 (A) as well as hClpXP (B) complexes and separation by SEC. Both complexes were cross-linked with 137.5-fold excess DSSO over ClpP monomers. As the peak at around 12.03 mL was considered the human $(ClpX_6)_2ClpP_{14}$ complex by comparison to the SaClpXP complex elution volume, we chose corresponding fractions for MS sample preparation. Highlighted regions were pooled for MS sample preparation.[110]

3.2 MS-based analysis of cross-linked ClpXP orthologs

Each cross-linked complex was reduced, alkylated and digested with trypsin. Peptides were enriched by SCX as performed for the ClpP orthologs and subsequently measured by LC-MS/MS using both of the fragmentation strategies described above.

A complete list of all identified cross-linked peptide pairs is reported in the appendix (**Table S5-7**). Within the ClpXP samples, three different types of cross-links were identified: intra-ClpP, intra-ClpX, and inter-ClpX-ClpP cross-links. Intra-ClpP cross-links, which were already detected in the previous experiments, represented the smallest proportion of all identified cross-links. Almost 80% of detected cross-links for bacterial orthologs were intra-ClpX. This number was even higher for hClpXP. This indicates an extraordinarily good accessibility of ClpX towards DSSO, probably caused by high flexibility along with a high lysine density in the chaperone sequence (**Figure I-17A**, **S5A** and **S6A**). While structural data for the different ClpPs are available, models of the ClpX hexamers were derived by matching corresponding orthologs onto the EcClpX crystal structure (PDB: 4I81; **Figure I-16**).[100,118]
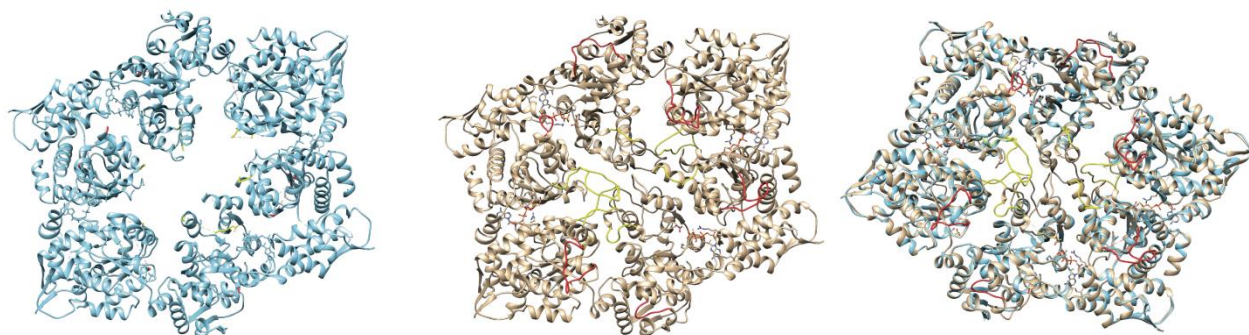
Figure I-16 SaClpX model crystal structure generated from the *E. coli* ortholog. Flexible loops, like the IGF-loops (red) and pore-2 loops (yellow) were almost completely disordered in the structure of EcClpX (PDB: 4I81) (left). Modeled structure of SaClpX. IGF-loops are colored in red, pore-2-loops in yellow (center). A superposition of both structures shows a sufficient fit without major deviations (right). SaClpX displays around 61% sequence identity with EcClpX. The human and LmClpX hexamers were modeled alike and displayed 52% and 62% sequence identity, respectively.[110]

In order to verify our models, we mapped all intra-ClpX cross-links onto corresponding hexamers. The resulting histogram of median Cα-Cα distances displayed most cross-links within the maximum distance constraint of DSSO (**Figure I-17B**, **S5B** and **S6B**). Remarkably, in the case of LmClpXP only inter-cross-links between LmClpP2 and LmClpX were identified (**Figure I-17C**, left part), which provides evidence that only the LmClpP2 but not the LmClpP1 ortholog is able to interact with its chaperone. These XL-MS results display the first structural data on the

Part I-3 Insights into the enigmatic ClpX-ClpP interface

LmClpXP1/2 complex, consolidating previous assumptions from activity and mutational studies.[83] The cross-links within LmClpXP2 matched equivalent inter-facial connections detected in the SaClpXP samples as expected due to the evolutionary conservation of the corresponding lysines (**Figure S5C**). For hClpXP we did not detect any inter-ClpX-ClpP cross-links. This is most likely caused by the fact that the hClpXP sequences lack analogous lysine residues at the ClpX-ClpP interface (for instance no lysine within N-terminal regions) compared to SaClpXP and LmClpXP1/2, which prevents the formation of DSSO cross-links.

Overall, signature connections in LmClpXP and SaClpXP were identified between lysines in the ClpX IGF-loops and the C- and N-terminal ClpP regions, as well as to the ClpP G-helix. As the C-terminal K193 of ClpP is located near the hydrophobic depression, a region with a high density of hydrophobic amino acids on top of the tetradecamer, our data confirms that the IGF-motifs are located in close proximity to this area.[99,119] Another hallmark of the ClpXP interaction is the connection of a lysine located in the ClpX pore-2-loops with the N-terminal region of ClpP[102] and to the ClpP E-helical handle-regions. Altogether, our cross-linking data indicate that the ClpX-ClpP interface is very flexible in solution. For ClpX, almost all cross-linking contacts to ClpP were formed *via* the IGF- and pore-2-loops, thereby confirming previously assumed interactions based on molecular docking as well as chemical and mutational studies.[99,102,119] Moreover, these ClpX loops appear to be in proximity to more than one structural region of the peptidase, suggesting a highly dynamic interaction.

As the ClpX-ClpP interface is asymmetric, we further investigated whether one LmClpP2 subunit forms substantial contacts not only with the closest LmClpX monomer, but also towards its two neighboring LmClpX subunits. We created an initial model of the LmClpX-ClpP2 complex in its full heptamer/hexamer form and extracted a sub-structure consisting of two adjacent trimers of LmClpX and LmClpP2. This sub-structure represents the smallest unit allowing a realistic computational investigation of all three relevant interfaces: LmClpP2-ClpP2, LmClpX-ClpX, and LmClpX-ClpP2. The complex was subjected to several short molecular dynamics simulations (11 • 40 ns) to allow for potential conformational adaptations of the subunits, especially of the loop structures, upon complex formation. Based on these simulation data, we calculated the average distances of all potential cross-link positions of the respective central subunits. For each subunit combination, the shortest average cross-linking distance is given in **Table S8**. In agreement with our previous observations, most cross-link positions are within the range of 35 Å. Interestingly, several of the shortest cross-links do not solely occur between the two central subunits, but also from the central LmClpP2 to the distant LmClpX monomers (left and right) (**Figure I-17C**, right part). These include cross-links from LmClpP2 N- and C-terminal

regions to LmClpX IGF-loops. The simulations further show that the LmClpP2 N-termini and the LmClpX pore-2-loops are in close proximity across the heterocomplex. Overall, the model explains the cross-linking diversity of both the pore-2- and IGF-loops due to their flexibility and proximity to key locations in the binding interface, emphasizing their important role in ClpX-ClpP binding.[99,102,119]
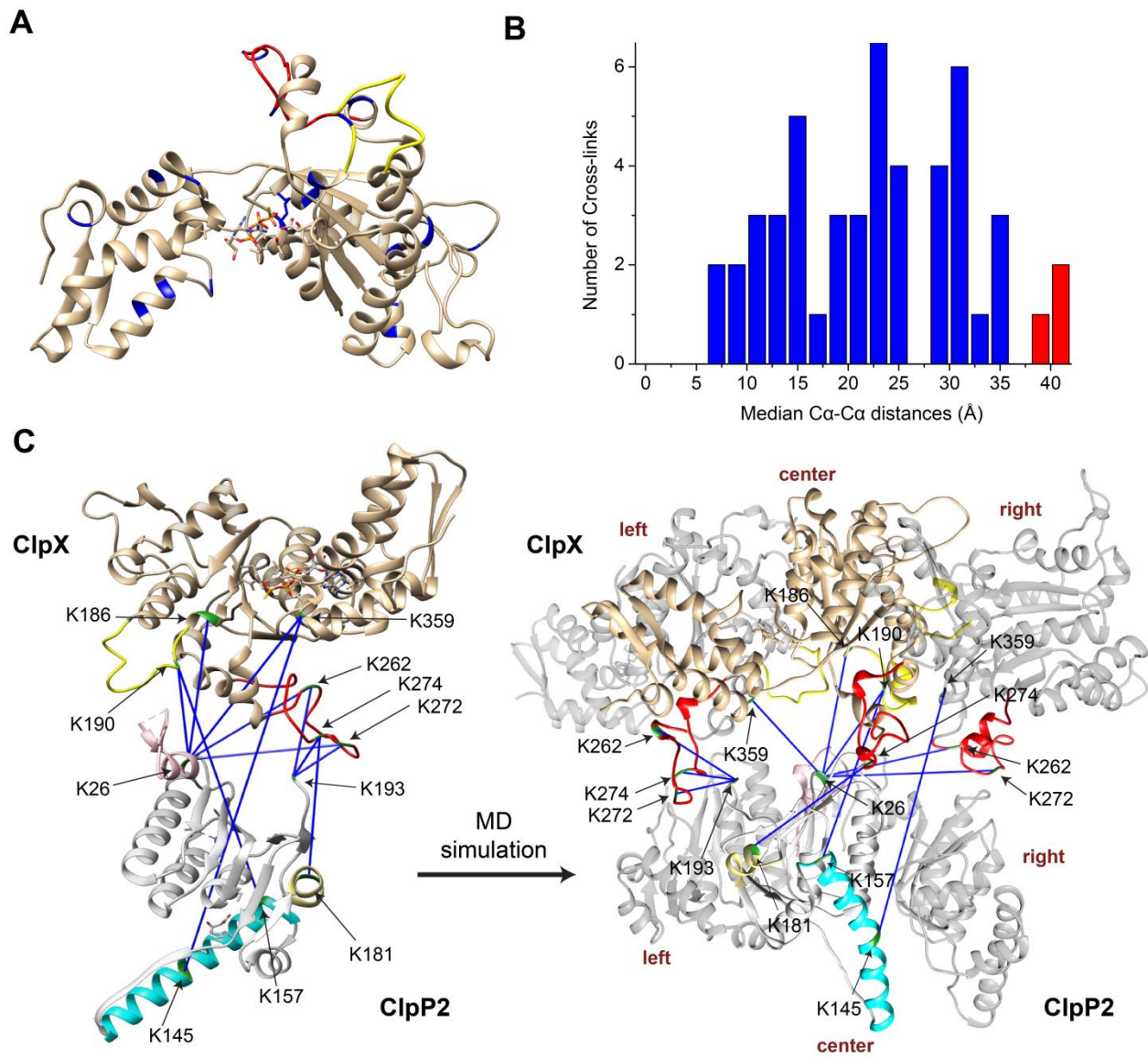
Figure I-17 Analysis of the ClpXP interface. (**A**) The largest proportion of cross-links (FDR <0.01) detected in the ClpXP samples are intra-ClpX links distributed all over the sequence (involved lysines are highlighted in blue in one monomer). The IGF-loop is colored red, the pore-2-loop yellow. (**B**) Cross-links are projected onto a ClpX hexameric model of *L. monocytogenes* derived from an *E. coli* crystal structure (PDB: 4I81). The histogram of median Cα-Cα distances displays most of the cross-links within the

maximum distance constraint of 35 Å. Cross-links exceeding the maximum distance constraint are highlighted in red. (**C**) LmClpX-ClpP2 cross-linking contacts: One subunit of LmClpP2 (PDB 4RYF, colored grey) and one subunit of LmClpX (colored beige) are positioned next to each other and inter-facial cross-links (involved lysine residues highlighted in green) are displayed from LmClpP2 to LmClpX. Notably, only intermolecular links between LmClpP2 and LmClpX, but not between LmClpP1 and LmClpX, were detected. The ClpX IGF-loop is colored red and the pore-2-loop yellow. ClpP N-terminus is colored pink, E-helix cyan and G-helix beige. Further, a representative molecular dynamics (MD) snapshot of an oligomeric LmClpX-ClpP2 model (two trimers) is presented. Only cross-links with the shortest distances from LmClpX subunits to the central LmClpP2 subunit are shown.[110]

3.3 Conclusion

We successfully applied our established chemical cross-linking workflow to study the ClpXP interface. We analyzed complexes from *S. aureus* and human, as well as the ClpXP1/2 hetero-complex from *L. monocytogenes*. SEC revealed a covalent fixation of all complexes upon cross-linking with DSSO. LC-MS/MS analysis resulted in three different types of cross-links: intra-ClpP, intra-ClpX, and inter-ClpP-ClpX. The major fraction displayed intra-ClpX cross-links as the chaperone shows a high lysine density as well as overall higher flexibility compared to ClpP. As no high resolution crystal structures of SaClpX, hClpX, and LmClpX are available, we modeled hexameric structures according to the existing structure from EcClpX.[100,118] Mapping of intra-ClpX cross-links onto hexameric models revealed most of the cross-links within the maximum distance threshold of DSSO. The highest number of inter-ClpP-ClpX cross-links was formed within the LmClpXP1/2 complex. Contacts were mainly mediated by IGF- and pore-2-loops of ClpX to several regions of ClpP. For instance, we could confirm contacts from IGF-loops to hydrophobic depressions, previously postulated by mutational and docking studies.[66,99] Remarkably, cross-links to ClpX were solely formed by LmClpP2, confirming for the first time that the interaction with the ATPase is mediated *via* the P2 part. Overall, our cross-linking data revealed the ClpXP interface to be very dynamic.

4 An optimized pull-down set-up to study cellular binding partners of hClpP

The two main methods of choice for studying cellular PPI are co-IP and affinity-enrichment of tagged proteins.[34–37] Although both strategies are widely applied, a loss of weak interactions or the generation of artificial contacts cannot be ruled out.[26] To overcome these limitations, XL-MS has been combined with both, co-IP and affinity-enrichment, to stabilize cellular PPI.[26,120] Thereby, several experimental set-ups were developed, including on-bead cross-linking after cell lysis and pull-down,[121] as well as *in situ* cross-linking prior to lysis and pull-down.[42]

Interaction networks of mammalian ClpP have been mainly studied by the use of inactive proteases, functioning as a molecular trap for substrates that cannot be degraded anymore. Subsequent ClpP pull-downs revealed several potential interactors and substrates identified by MS.[77,108,109] For instance, Cole *et al.* applied the BioID method to study hClpP interactors.[77,122] Thereby, the catalytically inactive ClpP mutant carrying a BirA tag, a promiscuous biotin ligase to biotinylate proteins based on proximity, was expressed in human embryonic kidney cells under the control of a tetracycline promotor. A subsequent anti-biotin pull-down enabled identification of 48 ClpP interactors/substrates by MS, mainly metabolic enzymes and proteins involved in the mitochondrial respiratory chain. Another study transiently overexpressed murine inactive ClpP fused to a FLAG-tag in mouse embryonic fibroblasts lacking any natural ClpP for subsequent pull-down.[108] A total of 66 possible substrates and interactors of murine ClpXP were detected, most of them involved in translational regulation, oxidative phosphorylation, and a number of metabolic pathways. Finally, interactors were studied by expression of active and inactive, FLAG-tagged hClpP in the fungi *Podospora anserina*, which was depleted of endogenous ClpP,

followed by pull-down and MS.[109] In total, 46 interacting partners and 20 potential substrates were identified, most of them associated with metabolism and the electron transport chain. Taken all studies together, several interactors and potential substrates of mammalian ClpP have been identified. However, heterologous expression of affinity-tagged proteins and genetic manipulation of host cells results in artificial working conditions.

Therefore, our goal was the application of the co-IP method to study hClpP interactors in their native state, as an antibody-based pull-down of the protein of interest does not require any manipulation of the cells. We aimed at comparing standard co-IP with co-IP combined with on-bead or *in situ* cross-linking to examine the stability of hClpP networks in a quantitative manner.

4.1 hClpP interaction network studied by conventional co-IP

First, we investigated cellular binding partners of hClpP by conventional co-IP from the liver-cancer derived cell line HepG2. Therefore, cells were lysed with a non-ionic lysis buffer containing 1% Nonidet P-40 (NP-40), a non-denaturing detergent. After pull-down with an anti-hClpP antibody at 4 °C, proteins were reduced, alkylated and digested on-bead using trypsin. Samples were analyzed by LC-MS/MS and subsequent label-free quantification (LFQ)[123] using the MaxQuant software (**Figure I-18A**).[124] To eliminate false-positive results caused by unspecific binding to the antibody and the beads, control experiments were included, in which lysates were incubated with an isotype control immobilized to the beads. Downstream statistical analysis and comparison to the control samples with the Perseus platform[125] revealed hClpP to be the most enriched protein (**Figure I-18B**). We further investigated whether the detected proteins are localized in the mitochondria (gene ontology (GO) term *cellular compartment* = mitochondrion; ID 0005739).[126,127]

Although the hClpP pull-down worked, its main interacting partner ClpX was enriched with a lower significance. As hClpP is localized solely in mitochondria,[72] it would be expected that the majority of interactors are mitochondrial proteins as well. However, only 25% of the identified ClpP binding proteins are localized in the mitochondria, suggesting that interactions with cytosolic proteins may be false-positives caused by cell disruption. Further, only two out of the 55 significantly enriched proteins were reported to be substrates or interactors in the previous studies.[77,108,109]
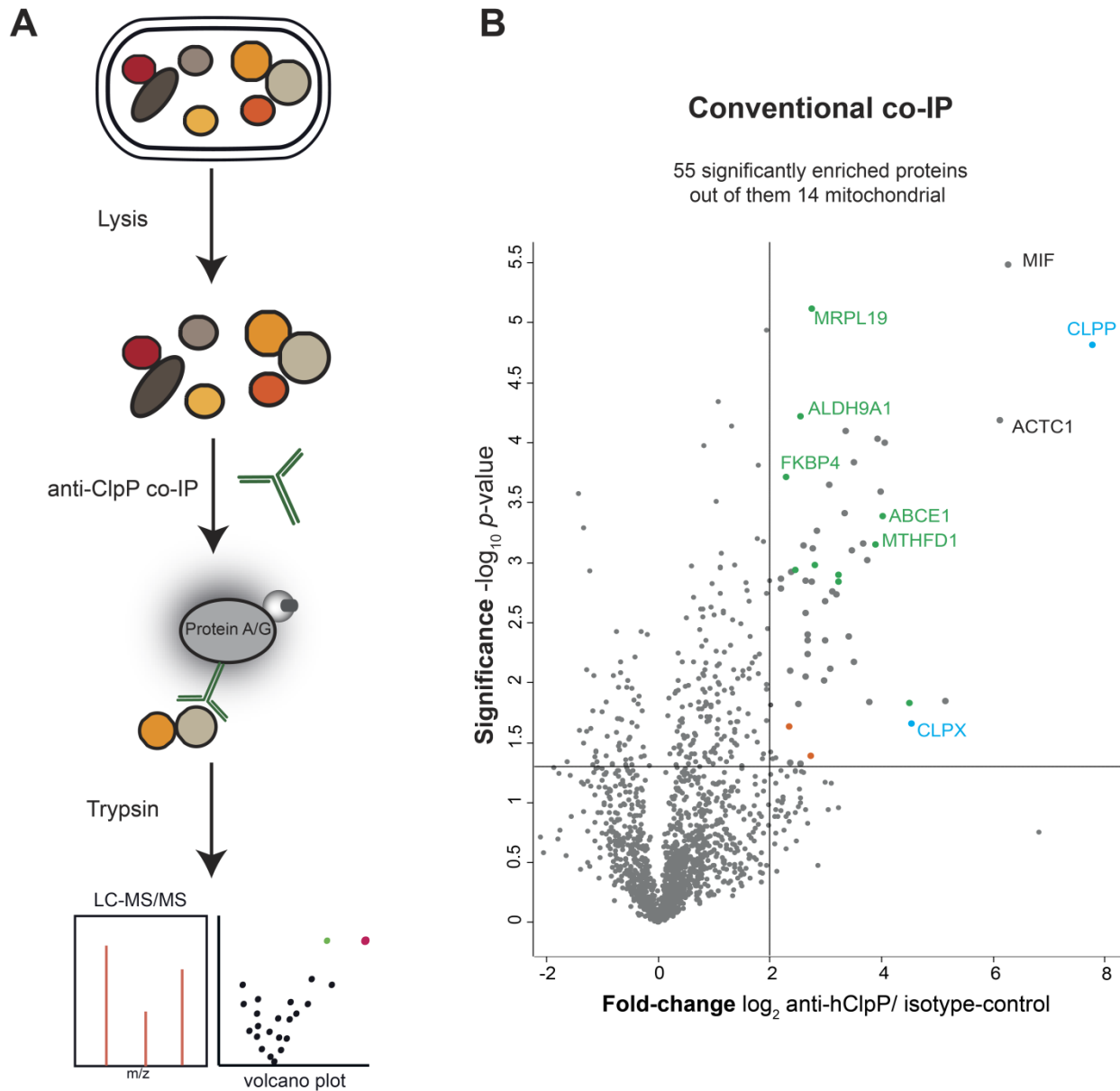
Figure I-18 Conventional co-IP against hClpP from HepG2. (**A**) Overview of the conventional co-IP workflow, including pull-down from cell lysate and subsequent LC-MS/MS analysis. (**B**) Volcano plot representing *t*-test results of anti-hClpP co-IP compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $log_2 = 2$ (4-fold enrichment) and $-log_{10}$ (*p*-value) of 1.3 (solid lines). The major target ClpP and its interacting chaperone ClpX are colored in blue. Mitochondrial proteins are highlighted in green, literature reported ClpP interacting proteins or substrates are colored orange. In the case that a mitochondrial protein was reported as an interacting protein or substrate in literature[77,108,109] it was colored orange.

4.2 hClpP interaction network studied by co-IP and on-bead cross-linking

To investigate if interactions especially with mitochondrial proteins get lost during the wash procedure following the hClpP pull-down, we applied the DSSO cross-linker at 2 mM for on-bead cross-linking for 2 h at 4 °C after incubation with the anti-hClpP antibody (**Figure I-19A**).
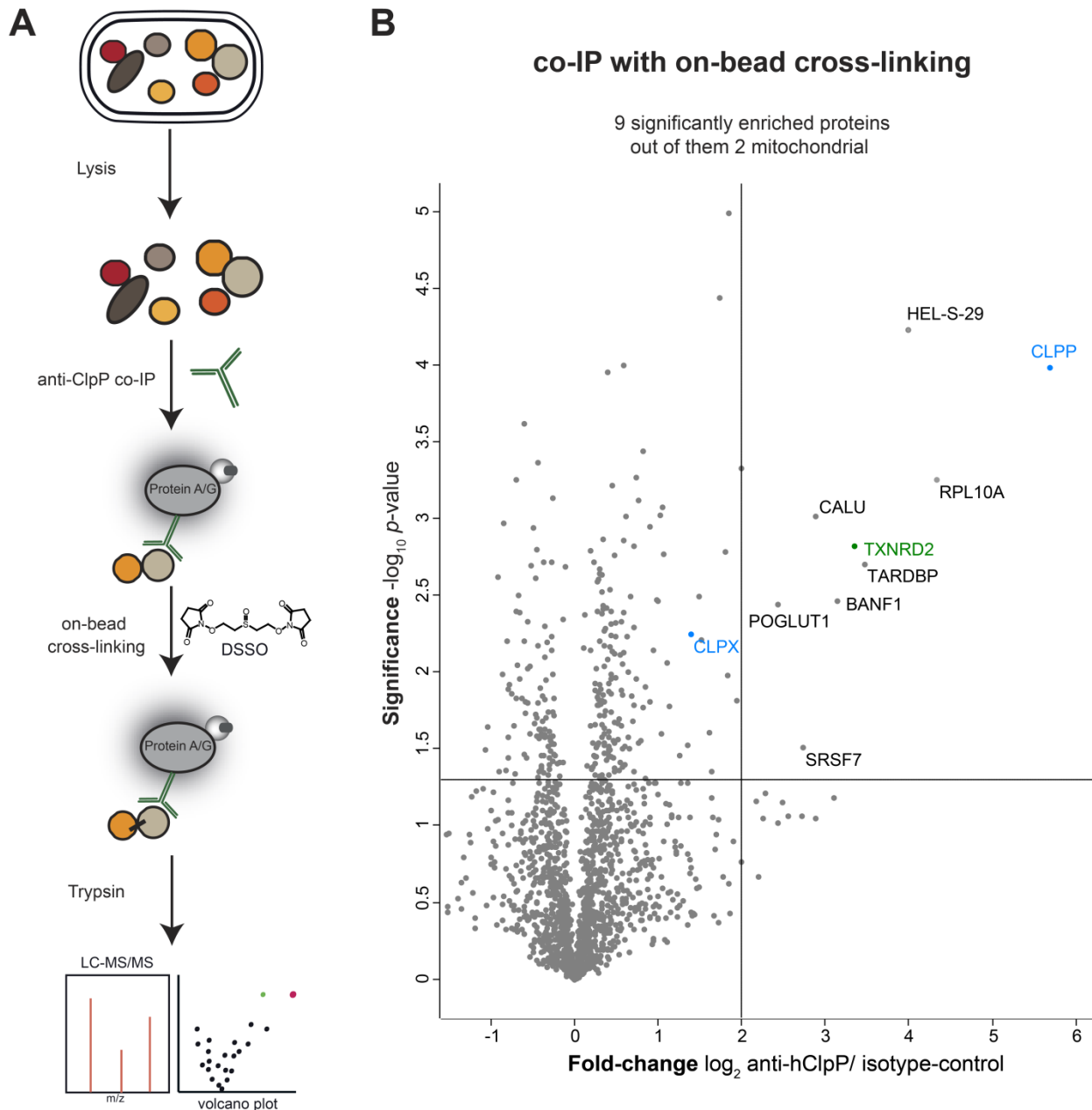
Figure I-19 Co-IP against hClpP from HepG2 combined with on-bead cross-linking. (**A**) Overview of the co-IP workflow combined with on-bead cross-linking using DSSO, including pull-down from cell lysate

and subsequent LC-MS/MS analysis. (**B**) Volcano plot representing t-test results of anti-hClpP co-IP compared to the isotype control co-IP (n=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2$ = 2 (4-fold enrichment) and $-\log_{10}$ (p-value) of 1.3 (solid lines). The major target ClpP and its interacting chaperone ClpX are colored in blue. Mitochondrial proteins are highlighted in green, literature reported ClpP interacting proteins or substrates are colored orange. In the case that a mitochondrial protein was reported as an interacting protein or substrate in literature[77,108,109] it was colored orange.

Subsequent analysis of MS-data revealed that hClpP is the strongest enriched protein (**Figure I-19B**). However, ClpX fell below the enrichment cut-off (4-fold enrichment), tough it still displayed to be more than 2-fold enriched. Only one of the significantly enriched interactors of hClpP was a mitochondrial protein, indicating that on-bead cross-linking did not contribute to stabilization of PPI. In addition, none of the enriched proteins was reported to be a substrate or interactor in previous studies.[77,108,109] As the overall number of interactors was markedly reduced compared to standard co-IP, on-bead cross-linking seemed to disturb interaction networks of hClpP even more compared to the conventional approach.

4.3 hClpP interaction network studied by cross-link/co-IP

As conventional co-IP and co-IP in combination with on-bead cross-linking did not result in a significant overrepresentation of mitochondrial interactors, we applied *in situ* chemical cross-linking using 2 mM of cell-permeable DSSO to freeze interactions insight the cells for 1 h at 37 °C. A subsequent pull-down of hClpP from HepG2 and LC-MS/MS analysis (**Figure I-20A**) revealed a 3-fold increase in interaction partners compared to conventional co-IP (**Figure I-20B**). Notably, more than 75% of the 154 high-confidence enriched hits upon DSSO treatment were mitochondrial proteins. Remarkably, due to the additional cross-linking step, the most prominent interaction partner of ClpP, its ATPase ClpX became the strongest enriched protein. In contrast, ClpX was only moderately enriched with low-confidence by classical co-IP and co-IP combined with on-bead cross-linking, suggesting that the interaction, if not covalently trapped, is transient and comparatively weak.

Upon cross-linker addition we were able to enrich almost 50 previously identified interacting partners or substrates of hClpP from the HepG2 cell line.[77,108,109] Moreover, several enriched proteins represent known bacterial ClpXP substrates such as the mitochondrial phosphoenolpyruvate carboxykinase (*PCK2*) for *C. crescentus* ClpXP[106] or the homolog of mitochondrial medium-chain specific acyl-CoA dehydrogenase (*ACADM*), a substrate of SaClpXP.[107] These interspecies connections point to a conserved role of the protease during

evolution. In addition, it was found very recently that *Drosophila* leucine-rich pentatricopeptide repeat domain-containing protein 1 (LRPPRC1) is a specific substrate of ClpXP.[128] As the human ortholog was significantly enriched upon hClpP pull-down, it could be possible that LRPPRC displays a substrate of the hClpXP complex as well.

A gene ontology enrichment analysis of the significant hits from hClpP experiments with the DAVID tool[129] reveals an overrepresentation of proteins involved in protein folding, protein biosynthesis, metabolism and oxidation-reduction processes (**Table S9**). Several identified interaction partners fit into literature proposed interacting proteins and associated phenotypes of ClpP disturbance. As the expression of the ClpXP protease was found to be altered upon unfolded protein stress it was therefore supposed to be involved in the mitochondrial unfolded protein response (mtUPR). In addition to ClpXP, we detected Lon protease, which is part of the mtUPR as well.[73] Further, we enriched phosphoenolpyruvate carboxykinase 2 (*PCK2*), which was previously identified as ClpP interacting protein and demonstrated to be inducted upon mitochondrial stress.[130] In addition, proteins involved in protein folding (*e.g. CDC37, PPIF, FKBP4, TRAP1*), response to stress (*e.g. HSPA4*) and peptide cleavage (*e.g. PITRM1, PMPCA, PMPCB*) were amongst the enriched hits, suggesting a central role of ClpXP in protein homeostasis. Presequence protease, *PITRM1*, an ATP-independent protease degrades mitochondrial transit peptides after their cleavage[131] and mitochondrial-processing peptidase (*PMPCA* and *PMPCB*) cleaves transit peptides from mitochondrial protein precursors.[132] Interaction of the ClpXP protease system with other degradation and processing systems indicates a highly cooperative protein homeostasis system in the mitochondria.

Remarkably, 16 mitoribosomal proteins were enriched upon combined cross-link/anti-hClpP co-IP, 9 of the large and 7 of the small subunit. This portion represents almost 20% of the entire ribosomal proteins in the mitochondria and suggests an involvement of hClpP in mitochondrial translation. Indeed, analysis of *Clpp* knockout mice revealed an up-regulation of translation machinery, as levels of mitoribosome subunits and proteins involved in translation initiation were increased.[108] In addition, we were able to enrich the mitochondrial elongation factor Ts (*TSFM*), as well as other translational initiation proteins and several tRNA ligases (*e.g. AARS2, EARS2, IARS2*).
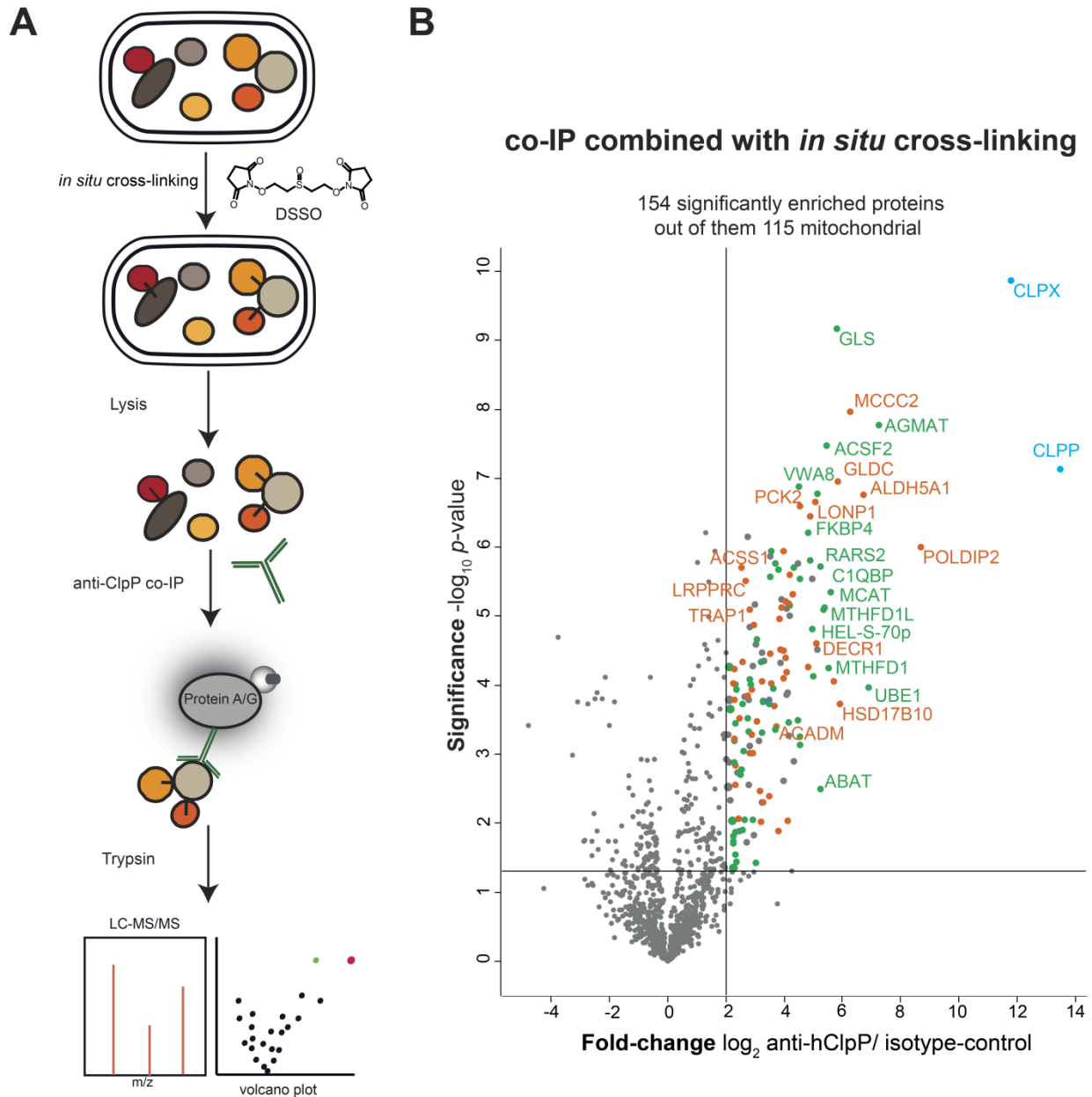
Figure I-20 Co-IP against hClpP from *in situ* cross-linked HepG2. (A) Overview of the combined *in situ* cross-link/co-IP workflow, including DSSO application to intact cells, pull-down from cell lysate and subsequent LC-MS/MS analysis. **(B)** Volcano plot representing *t*-test results of anti-hClpP co-IP compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of log₂ = 2 (4-fold enrichment) and -log₁₀ (*p*-value) of 1.3 (solid lines). The major target ClpP and its interacting chaperone ClpX are colored in blue. Mitochondrial proteins are highlighted in green, literature reported ClpP interacting proteins or substrates are colored orange. In the case that a mitochondrial protein was reported as an interacting protein or substrate in literature[77,108,109] it was colored orange.

Part I-4 An optimized pull-down set-up to study cellular binding partners of hClpP

Further, a central role of ClpP in energy metabolism was proposed in literature, as *Clpp* knockouts in mice, mouse myoblasts and K562 leukemic cells lead to respiratory deficiency[77,78,108] and ClpP was found to interact with mitochondrial respiratory chain proteins.[77,108,109] Additionally, pyruvate dehydrogenase subunits were identified as potential ClpXP substrates[109] and the Lon-ClpP proteolytic quality control was shown to degrade complex I reactive oxygene species (ROS)-generating domains upon mitochondria depolarization.[133] Our results also strongly support an involvement of ClpP in mitochondrial energy metabolism and respiration efficiency, as there were four subunits of the NADH dehydrogenase complex (complex I) and the mitochondrial electron transfer flavoprotein subunit alpha (*ETFA*) amongst the enriched protein hits. Further, we found two components of the pyruvate dehydrogenase (*PDHA1, PDHX*) as well as two components of the isocitrate dehydrogenase complex (*IDH3A, IDH3B*). Our data revealed lots of proteins acting on coenzyme A substrates (*e.g. ACAA2, ACAD9, ACADM, ACAT1*), two adenylate kinases (*AK3, AK4*), CTP-synthase (*CTP1*) and ATP-citrate synthase (*ACLY*), amongst numerous other energy related proteins.

The ClpP protease was shown to play an important role in growth,[76] tumor cell proliferation,[77,78] invasion and metastasis.[79] The protease is overexpressed in different tumor types, *e.g.* acute myeloid leukemia, prostate cancer, sarcomas and breast adenocarcinoma.[77,79] Besides its involvement in mitochondrial energy housekeeping and mRNA translation, a clear reason for up-regulation of ClpP in cancer cells could not be identified so far. Recently, ClpXP was recognized as component of a proteostasis network together with a molecular chaperone (*TRAP1*) and survivin, an inhibitor of apoptosis.[79] With mitochondrial heat shock protein 75 kDa (*TRAP1*) playing a role in tumor progression,[134] two potential cancer promoting proteins could be directly linked to the ClpXP system. In our data, we did not only detect *TRAP1* gene product, but also another protein that is connected to cell proliferation and tumor progression, namely the high confident enriched mitochondrial agmatinase (*AGMAT*), the enzyme for conversion of agmatine to putrescine.[135] Putrescine belongs, besides spermine and spermidine, to the polyamines which are often over-produced in cancer cells as they contribute to cell proliferation and tumor growth.[136] If considered as a direct or indirect interacting partner, the so far quite uncharacterized enzyme could be a candidate to build up a connection from the up-regulation of ClpP to the development of several cancer types.

4.4 Conclusion

In order to examine the subset of cellular binding partners of hClpP, we performed co-IP from liver-cancer derived HepG2. To investigate the stability of hClpP interactions, we compared different experimental conditions. We performed conventional co-IP, co-IP in combination with on-bead cross-linking, as well as co-IP from *in situ* cross-linked cells. The conventional co-IP resulted in few mitochondrial interaction partners, though hClpP was the most enriched protein verifying that the pull-down worked correctly. Only two of the significantly enriched proteins were previously reported to be substrates or interactors of ClpP,[77,108,109] indicating that some of the cellular interactors may have been lost. To investigate whether the interacting network was already disturbed upon cell lysis or if interactions were destroyed during co-IP workflow, we performed cross-linking directly after incubation of the cell lysate with the antibody to capture the interaction network bound to the beads. Though on-bead cross-linking is applied widely, we did not obtain any increase in mitochondrial hClpP interaction partners compared to the conventional pull-down. Further, ClpX as the main interaction partner of hClpP fell below the enrichment cut-off. Finally, we combined co-IP with *in situ* cellular cross-linking to freeze interactions within the cell previous to lysis. Although six (*ABCE1, ALDH9A1, MRPL44, MRPL19, MST065, VARS*) of the mitochondrial proteins that were enriched in the conventional co-IP were not enriched in the combined cross-link/co-IP, we achieved the pull-down of an extraordinary high number of mitochondrial proteins (116), including almost 50 previously declared interacting partners or substrates of ClpP by DSSO addition. Further, the main interaction partner ClpX became the protein with the highest fold-changes compared to the isotype control. Overall, we detected several proteins involved in energy metabolism, translation, or protein folding, indicating a central role of hClpP in mitochondrial housekeeping. Further, some of the potential interactors could display a link to the observed cancer-association of hClpP,[77] inspiring downstream analysis for further exploration of these networks.

Our results demonstrate that some of the identified interaction partners from conventional co-IP workflows may be false-positives, as for a mitochondrial protease like ClpP, cytosolic proteins, which were the biggest number of identified interactors in the conventional pull-down, should not be the main interacting proteins. Further, our data highlight that cellular PPI networks are often of a very sensitive and weak nature and may be already disturbed upon cell lysis, especially in the case of hClpP. This data displays the first comparison of different co-IP strategies in a quantitative manner. Because cross-link/co-IP can be applied to any protein of interest for which antibodies exist and does not require any complex genetic engineering of cells, it is extraordinary useful for studying interaction networks of the protein of choice in live cells.

5 Validation of Lon-protease as hClpP interactor

We demonstrated a superior performance of combined cross-link/co-IP over conventional pull-down methods. Several potential interactors of hClpP were detected, with the majority of them localized in the mitochondria. However, to validate the results of co-IP experiments its of particular importance to investigate if detected interactions can also be obtained by reverse co-IPs.[137] In order to confirm the interaction of hClpP with a high-confident enriched protein, we performed a reverse cross-link/co-IP against the Lon protease (*LONP1*) (**Figure I-21**). As mentioned above, the Lon protease is, like ClpXP, a member of the AAA+ superfamily and coordinates the degradation and processing of proteins in the mitochondrial matrix.[73] In contrast to ClpP, Lon protease contains both, a serine protease domain, as well as an ATPase functional domain.[138,139] Therefore, Lon is able to degrade proteins without the additional association with a chaperone. Satisfyingly, hClpP as well as hClpX were amongst the high-confident enriched hits strongly suggesting a relevant interaction with Lon *in situ*. Approximately 90% of the enriched proteins are predicted to be localized in the mitochondria where the Lon protease is exclusively detected.[73] Further, almost 45 of the proteins have been reported to be Lon substrates or to interact with the protease.[133,140–147] In our anti-Lon cross-link/co-IP, we detected several literature reported substrates of the protease, such as thioredoxin-dependent peroxide reductase (*PRDX3*)[143] or mitochondrial cytochrome c oxidase subunit 4 (*COX4I1*),[147] which were not enriched in the corresponding hClpP experiment. Interestingly, leucine-rich pentatricopeptide repeat domain-containing protein (*LRPPRC*) was detected as a high-confident interaction partner of both, Lon and ClpXP proteases. As mentioned above, it was recently found that the *Drosophila* ortholog of LRPPRC is a specific substrate of ClpXP. Both mammalian and

Drosophila LRPPRC proteins are required for polyadenylation, as well as stability of mitochondrial mRNAs. Moreover, it was suggested that mammalian LRPPRC is degraded not only by the ClpXP system, but also by the Lon protease,[128] which is also supported by our cross-link/co-IP results.
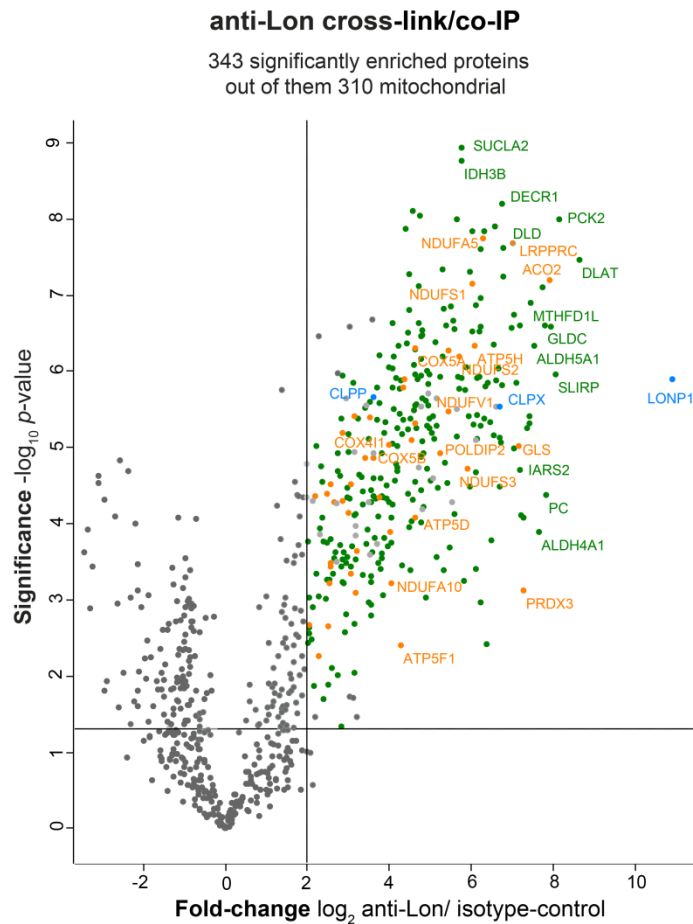
Figure I-21 Co-IP against Lon from *in situ* cross-linked HepG2. Volcano plot representing *t*-test results of anti-Lon co-IP compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 2$ (4-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). The major target Lon, as well as ClpP and ClpX are colored in blue. Mitochondrial proteins are highlighted in green, literature reported Lon interacting proteins or substrates are colored orange. In the case that a mitochondrial protein was reported as an interacting protein or substrate in literature[133,140–147] it was colored orange.[110]

Taken together, we successfully validated one of the high-confidence interactors of our combined anti-hClpP pull-down by reverse cross-link/co-IP, demonstrating the overall reliability of our data sets.

6 Impact of cisplatin on the hClpP interaction network

With a validated and effective co-IP workflow in hand, we utilized our set-up for studying changes in ClpP interaction networks upon cellular perturbations. hClpP has been linked to cell proliferation,[77,78] invasion, and metastasis,[79] and was demonstrated to be overexpressed in different tumor types like acute myeloid leukemia.[77] Furthermore, as mentioned above, genetic disruption or chemical inhibition of hClpP reduces the viability of leukemia cells.[77] We therefore selected cisplatin, a chemotherapeutic agent accumulating in mitochondria (**Figure I-22A**).[148] Among its many targets, cisplatin cross-links to DNA and interferes with transcription, replication, and ATP synthesis. Interestingly, elevated levels of hClpP result in lower sensitivity to cisplatin-induced apoptosis and *vice versa*.[149] These findings suggest that hClpP might be a broadly efficacious target for anti-cancer therapy.

To evaluate whether incubation of cells with cisplatin leads to differences in hClpP interaction levels, we first incubated the chronic myelogenous leukemia-derived cell line K562, which was previously shown to overexpress ClpP,[77] with increasing concentrations of cisplatin for 48 h. Relative metabolic activity was determined *via* a 3-[4,5-dimethylthiazol-2-yl]- 2,5-diphenyl tetrazolium bromide (MTT) assay.[150] Concentrations exceeding 2 µM cisplatin turned out to be toxic to K562 cells (**Figure I-22A**). Therefore, we chose a concentration of 2 µM for co-IP studies as we expected to observe cellular changes due to incubation with the chemotherapeutic agent whereas toxic effects are kept at a minimum. K562 cells were incubated with the agent for 48 h including a control set without cisplatin incubation. Upon cross-linking, cell lysis and pull-down,

proteins were reduced, alkylated, and digested as described above. Peptides were measured by LC-MS/MS previous to downstream statistical analysis.

A

B
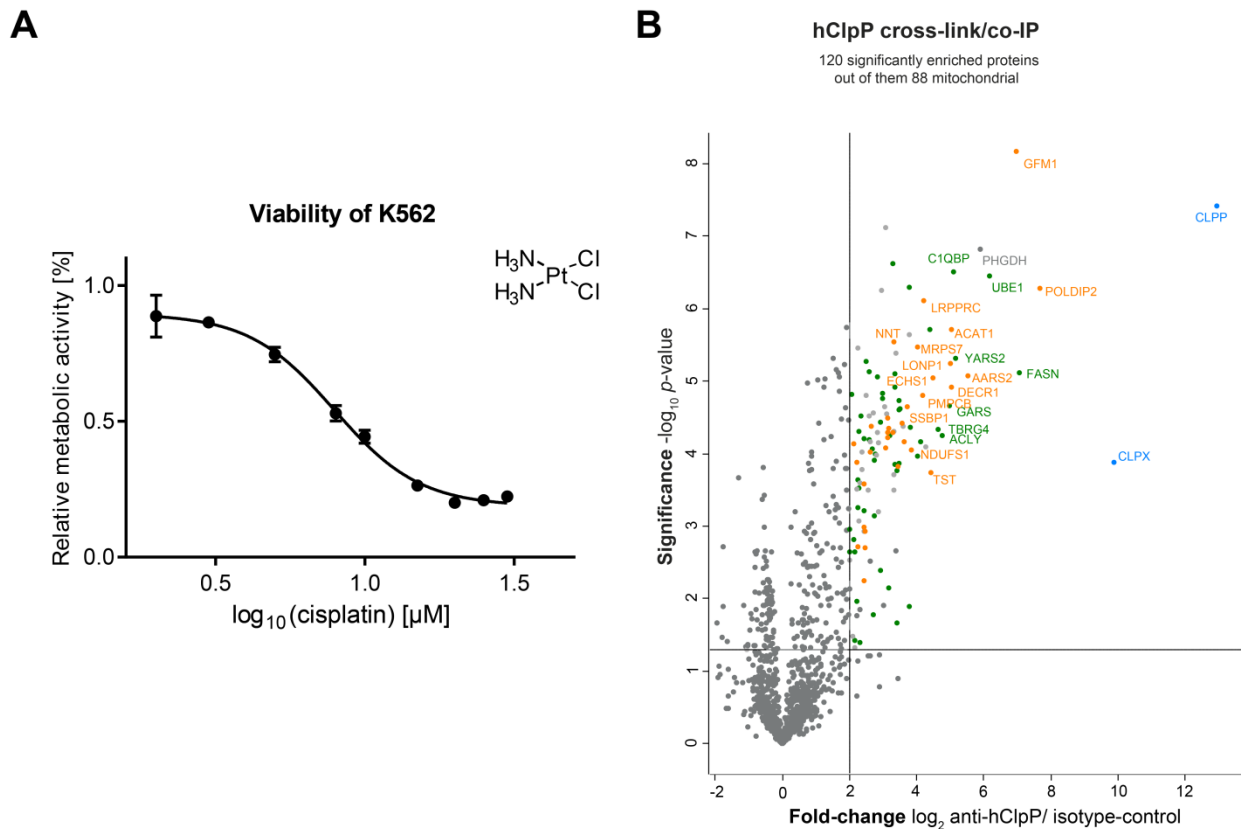
Figure I-22 Toxicity of cisplatin and *in situ* chemical cross-linking combined with co-IP from K562. (**A**) Chemical structure and toxicity of cisplatin in K562 studied by an MTT-assay. Relative metabolic activity of cells after 48 h incubation was normalized to a control without cisplatin incubation (*n*=3). The half-maximal inhibitory concentration (IC$_{50}$ value) was calculated to be 7.908 +/- 0.2715 µM. (**B**) Volcano plot representing *t*-test results of anti-ClpP co-IP from cells without cisplatin incubation compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of log$_2$ = 2 (4-fold enrichment) or depletion factor of log$_2$ = -2 (4-fold depletion), respectively, and -log$_{10}$ (*p*-value) of 1.3 (solid lines). The major target ClpP, together with ClpX are colored in blue, mitochondrial proteins in green and previously reported substrates or interactors in orange.[77,108,109] In the case that a mitochondrial protein was reported as interactor or substrate it was colored orange.

First, we compared the K562 control data set without cisplatin treatment to our anti-hClpP cross-link/co-IP from HepG2. We enriched 35 previously identified interactors or substrates of hClpP[77,108,109] out of 120 high-confidence hits with ClpX displaying the most enriched protein (**Figure I-22B**). Overall, 60 commonly enriched proteins were detected in both experiments. Observed differences could either stem from individual cell line specific pathways or deviating

Part I-6 Impact of cisplatin on the hClpP interaction network

protein expression levels. For example, RNA expression levels of mitochondrial agmatinase (*AGMAT*) and mitochondrial glycine dehydrogenase (*GLDC*), both high-confident interactors of hClpP from the HepG2 pull-down data, are very low in the K562 cell line.[151,152] Vice versa, high-confidence interaction partners such as the polymerase delta-interacting protein 2 (*POLDIP2*) or ubiquitin-like modifier-activating enzyme 1 (*UBE1*), both showing comparable expression levels, were detected in both cell lines. In fact, POLDIP2 was recently shown to be a selective interactor of ClpX in order to regulate substrate degradation.[153]
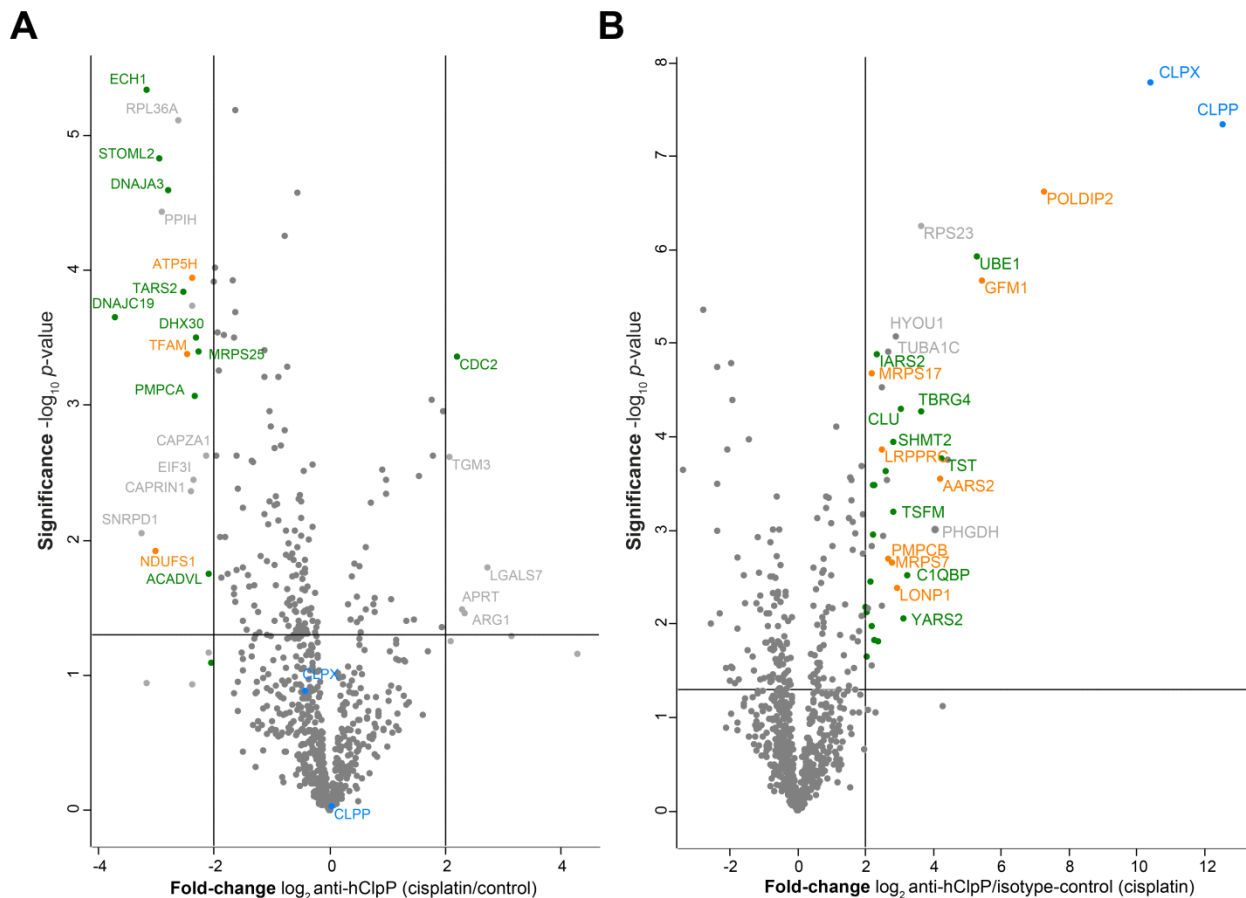
Figure I-23 *In situ* chemical cross-linking combined with co-IP from K562 to study effects of cisplatin. (**A**) Volcano plot representing *t*-test results of anti-ClpP co-IP from K562 cells without cisplatin incubation compared to the anti-ClpP co-IP from K562 cells with cisplatin incubation (*n*=4 biological replicates). (**B**) Volcano plot representing *t*-test results of anti-ClpP co-IP from cells with cisplatin incubation compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 2$ (4-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). The major target ClpP, together with ClpX are colored in blue, mitochondrial proteins in green and previously

reported substrates or interactors in orange.[77,108,109] In the case that a mitochondrial protein was reported as interactor or substrate it was colored orange.

A comparison of hClpP pull-down from cisplatin-treated cells with non-treated control cells revealed an overall reduced number of interacting proteins (**Figure I-23A**). Several of the missing interactors are localized in the mitochondria, suggesting an overall disturbance of protein interaction networks by cisplatin. One of the depleted proteins is mitochondrial ATPase subunit d (*ATP5H*), likely displaying a connection to hampered ATP synthesis upon cisplatin treatment.[149] Further, interaction of hClpP with mitochondrial transcription factor A (*TFAM*) was abolished in cisplatin-treated cells. As TFAM is essential for mitochondrial transcription regulation, depleted levels of the transcription factor in the anti-hClpP co-IP may display a link to cisplatin interfering with mitochondrial transcription.[149] Interestingly, the only mitochondrial protein which displayed an increased interaction with hClpP upon cisplatin treatment is cyclin-dependent kinase 2 (*CDC2*), which plays a key role in the control of the eukaryotic cell cycle. An explanation for the enrichment of CDC2 upon cisplatin treatment is probably its accumulation in mitochondria in G2-arrested cells upon DNA-damage.[154] Although the number of significantly enriched interactors of hClpP was reduced, the interactions with for instance ClpX, as well as with potential substrates like POLDIP2 were not disturbed upon cisplatin treatment (**Figure I-23B**).

In summary, our results serve as proof-of-principle experiment that cross-link/co-IP can be a useful tool to study alterations in protein interaction networks induced by external or internal physiological alterations.

7 Interaction partners of EcClpP

As *in situ* cross-linking followed by pull-down against hClpP turned out to efficiently capture a high number of interacting proteins compared to a non-cross-linked experiment, we aimed at transferring our workflow to bacteria to demonstrate a wide ranging application.

As more than 90 % of cellular proteolysis in *E. coli* is carried out by ATP-dependent proteases as Clp, Lon and FtsH, the ClpXP system plays a key role in protein homeostasis and a deeper analysis of its cellular function is crucial.[155,156] The central role of ClpP is further underlined by several so far identified native substrates of the protease, like the MazE protein, an endogenous regulator of programmed cell death in *E. coli*[157] and two bacteriophage regulators, the λO replication protein[158] and the Mu vir repressor protein.[159] Not only involved in resistance development,[160] ClpXP controls also the expression of type III protein secretion system in the human pathogen strain enterohemorrhagic *E. coli* (EHEC).[161] Although experiments with recombinant EcClpP showed that the protein was not predominantly present as tetradecamer upon cross-linking with DSSO, we anticipated that identification of binding partners *via* fixation of subunits with interacting proteins would be sufficient. Further, as the EcClpP system is better characterized than any other bacterial ClpP and several substrates have already been reported, the validation of pull-down data can be easily facilitated.[86,104] As mentioned previously, potential substrates of EcClpP were mainly identified by the application of inactive ClpP traps.[104,105] Thereby several proteins were detected, with the largest fraction involved in transcription, translation, protein folding, metabolism, and energy production. However, investigation of interaction partners of EcClpP under more native conditions for instance by the co-IP method has never been conducted so far. We therefore chose the *E. coli* ClpXP system for examination by conventional, as well as combined co-IP.

7.1 EcClpP pull-down reveals a central cellular role of the protease

Contrary to hClpP, EcClpP associates not only with ClpX but also with the related ATPase ClpA.[86] Although the conventional ClpP pull-down experiment already revealed 82 significantly enriched proteins (**Figure I-24A**), the two chaperones ClpX and ClpA were not or only slightly enriched. In comparison, upon DSSO addition, the number of enriched proteins was not only doubled (195), but a high-confident enrichment of both ClpX and ClpA could also be achieved (**Figure I-24B**). Notably, the *E. coli* pull-down did not reveal enrichment of the ATPase ClpB, neither with nor without DSSO addition, reflecting that ClpB is not a native chaperone of ClpP.[162]

A high number (71) of enriched proteins upon conventional co-IP was enriched in the combined cross-link/co-IP as well. The non-overlapping proteins which were missing under cross-linking conditions could be a result of background binding or statistical evaluation as besides *deaD* none of the proteins were amongst the top enriched hits in the conventional co-IP. In the case of *deaD* the enrichment in the DSSO experiment fell with a Students t-test difference of 2.33 not far behind the cut-off value of 4.

Overall, upon cross-linker addition we were able to enrich almost 30 substrates of EcClpP[104,105] in our pull-down. Several potential substrates (*e.g. gatY, gatZ, lipA*) were exclusively enriched upon cross-linker treatment, but were missing in the conventional co-IP. Previously identified substrates include the *E. coli* RNA polymerase sigma factor RpoS (σ^s), a central regulator of the general stress response.[160] Intracellular levels of σ^s are controlled by the ClpXP protease system and the recognition factor RssB, contributing to a constitutive degradation during exponential growth phase.[160,163] An increase in the expression of σ^s in stationary phase in *E. coli* appears to be the major trigger for developing highly resistant states during some sorts of starvation and stress,[164] demonstrating a central role of the ClpXP protease in *E. coli* physiology. Interestingly, with application of the cross-linker DSSO we did not only enrich the ClpXP substrate sigma factor RpoS, but also its regulation factor FliZ.

Further, EcClpP was reported to be associated with other chaperones than ClpX or ClpA *in vitro*. Those include for instance the GroEL/GroeES system or trigger factor.[165] As our pull-down resulted in an enrichment of both, GroEL and trigger factor, an interaction *in vivo* may also be possible. In addition, our results suggest a role of ClpP in genome replication and cell division. Besides DNA helicases (*dnaB, rep*), DNA gyrases (*gyrA* and *gyrB*), DNA polymerase I (*polA*), and DNA repair protein RadA (*radA*), several proteins involved in chromosome (*mukB, mukE, mukF*) and cell (*ftsE*) partition were amongst the significantly enriched ones. Interestingly, besides proteins detected in our cross-link/co-IP, the DNA extension and proofreading subunit ε

(DnaQ) of DNA polymerase III,[166] as well as the tubulin-like FtsZ protein that assembles into a ring at midcell marking the positions for dividing of *E. coli*[167] were identified as EcClpXP substrates.
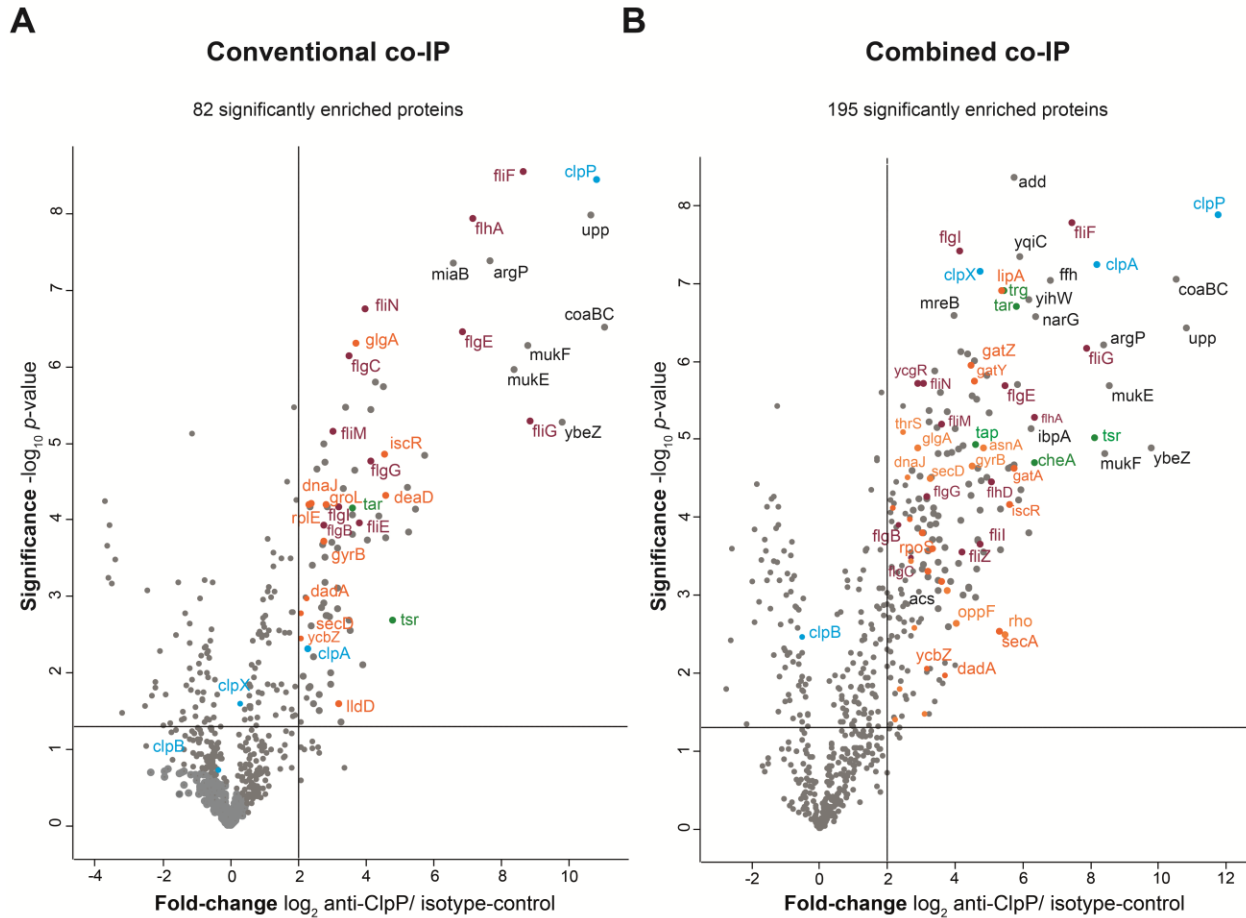
Figure I-24 *In situ* **chemical cross-linking enhances co-IP coverage in** *E. coli*. Volcano plots representing *t*-test results of anti-ClpP co-IP compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 2$ (4-fold enrichment) and -$\log_{10}$ (*p*-value) of 1.3 (solid lines). (**A**) Conventional co-IP in *E. coli* compared with *in situ* cross-linking and pull-down (**B**). The major target ClpP, together with its two interacting chaperones ClpA and ClpX as well as the ATPase ClpB are colored in blue, flagellum-related proteins in magenta, chemotaxis-related proteins in green and previously identified substrates in orange.[110]

Several proteins involved in gene transcription were among the enriched proteins, for instance transcription regulators (*malt, rcsB*), the transcription termination factor Rho (*rho*), an RNA polymerase-associated protein (*rapA*), as well as two RNA polymerase sigma factors (*rpoD and rpoS*). Further, a huge number of hits is related to protein translation, including several

elongation factors (*efp, lepA*), tRNA ligases (*aspS, proS, thrS*), ribosomal proteins itself (*rplE, rpmE*) and ribosomal protein modifying enzymes (*prmB, rimJ, rimO, ycaO*).

ClpP was demonstrated to play an essential role in amino acid and energy homeostasis, as the protease is important for proteolytic degradation during stress.[162] In line with this observation, we detected several proteins linked to amino acid and energy housekeeping among the significantly enriched EcClpP interactors. For instance, we identified enzymes involved in coenzyme A (*e.g. accA, accD*), amino acid (*e.g. dacA, dadA*), fatty acid (*e.g. fabA, fabH, fadB, fadL*), and sugar metabolism (*e.g. gatY, gatZ*), glycogen building and break-down (*glgA, glgP*), as well as proteins involved in lipoprotein and lipopolysaccharide homeostasis (*e.g. lipA, lolD, lptB*). Interestingly, we detected an overrepresentation of NADH-quinone oxidoreductase subunits (*nuo*-genes) amongst enriched proteins, as already observed in the hClpP pull-down dataset. Our results suggest an involvement of EcClpP in cellular respiration, which may be mediated upon interaction with or degradation of subunits of complex I.

Remarkably, we identified 15 proteins that are involved in *E. coli* flagellum formation or are subunits of the flagellum itself. Those include proteins of the basal body (*flgB, flgC, flgG*), several motor switching proteins (*fliG, fliM, fliN*), flagellar hook protein FlgE, the P- and M-ring proteins (*flgI, fliF*), the flagellum-specific ATP synthase (*fliI*), the biosynthesis protein FlhA, and the transcriptional regulator FlhD. This suggests an essential role of the EcClpXP system in flagellar housekeeping, which is in line with literature observations. The protease was previously shown to degrade the master regulators FlhC and FlhD of flagellum biosynthesis and therefore decreases flagellar formation.[85,168]

Furthermore, the soluble chemoreceptor McpB and the integral membrane chemoreceptor McpA were reported as ClpXP substrates in *Caulobacter crescentus*.[169,170] Here, we show co-purification of the methyl-accepting chemotaxis proteins (MCPs) I-IV (*tsr*, *tar*, *trg*, *tap*) and of chemotaxis protein CheA (*cheA*), which are distantly related to McpA and McpB based on homology, suggesting a connection of ClpP to *E. coli* chemotaxis as well. Importantly, almost half of these proteins were only enriched in the cross-linked sample and would have otherwise escaped detection.

7.2 Conclusion

Taken together, we successfully transferred the optimized co-IP strategy from human cells to gram-negative *E. coli*. Comparison of the conventional pull-down with combined cross-link/co-IP revealed twice as much interactors upon DSSO addition. Remarkably, the main interactors of

Part I-7 Interaction partners of EcClpP

EcClpP, the two ATPases ClpA and ClpX were not or only moderately enriched in the conventional pull-down but were two of the most significantly enriched proteins due to cross-linker addition. This suggests, as for the human ClpXP system, that the interactions are of an overall transient nature. We were able to enrich a huge fraction of previously suggested substrates,[104,107] as well as novel interactors, suggesting a central role of EcClpP for instance in chemotaxis or flagellar housekeeping. Our data displays the first subset of cellular EcClpP interactors/substrates under native conditions and may inspire downstream investigations or follow-ups concerning the cellular role of this important bacterial protease.

8 Combination of chemical cross-linking with ABPP

ABPP displays a widely applied method for chemical proteomic profiling of cellular targets of for instance natural products. The ABPP methodology relies on the design of covalently binding probes directed to single proteins or protein families. In addition to their warhead the probes are modified with a moiety enabling for biorthogonal ligation to fluorophores or affinity-handles to visualize or enrich for the cellular target proteins (cf. Part II, Chapter 1.4).[171] A possible and so far not applied development of ABPP could be its combination with chemical cross-linking (cross-link/ABPP) to enrich not only for the specific target proteins of the probe, but simultaneously also for the interactors of this target.

The best suited ABPP probe for investigation of such a combined experimental set-up should display an overall low number of targets not to overwhelm the fraction of enriched proteins, as for each target several interactors are expected to co-purify. Ideal candidates display β-lactone probes, which are derived from natural products. A screen against bacterial proteomes identified virulence-associated protease ClpP as one of the major targets.[172] A quantitative MS-based target identification revealed the β-lactone probes **D3** and **E2** (**Figure I-25A**) to be very specific, with ClpP displaying one of the most prominent enriched protein hits together with an only moderate number of *off*-targets in *S. aureus*.[89] The probes inhibit SaClpP peptidase activity, as well as *S. aureus* extracellular virulence by attenuation of hemolysis.[88,173,174] We therefore chose **D3** and **E2** for our cross-link/ABPP approach, as these β-lactone probes are well-studied and show a limited number of target proteins.

8.1 Efficiency of click-chemistry after chemical cross-linking

Before conducting combined cross-link/ABPP on a cellular scale, we examined whether probe-labeled, DSSO cross-linked recombinant ClpP can still be clicked to azide functionalized handles or if the more compact covalently linked complexes are not accessible anymore. Therefore, we incubated SaClpP with **D3** (1:1 ratio) for 1 h. Afterwards, we cross-linked proteins with increasing concentrations of DSSO (10-200 fold molecular excess) for 1 h at 37 °C. Following quenching of residual cross-linker, we performed CuAAC to attach rhodamine-azide. Subsequent analysis of the products by SDS-PAGE and fluorescence read-out revealed that on the one hand SaClpP was successfully cross-linked and that all of the cross-linked species can be efficiently clicked to the fluorophore on the other hand (**Figure I-25B**). Taken together, cross-linking of **D3**-labeled SaClpP does not interfere with click-chemistry, which is an important requirement for our downstream cellular experiments.
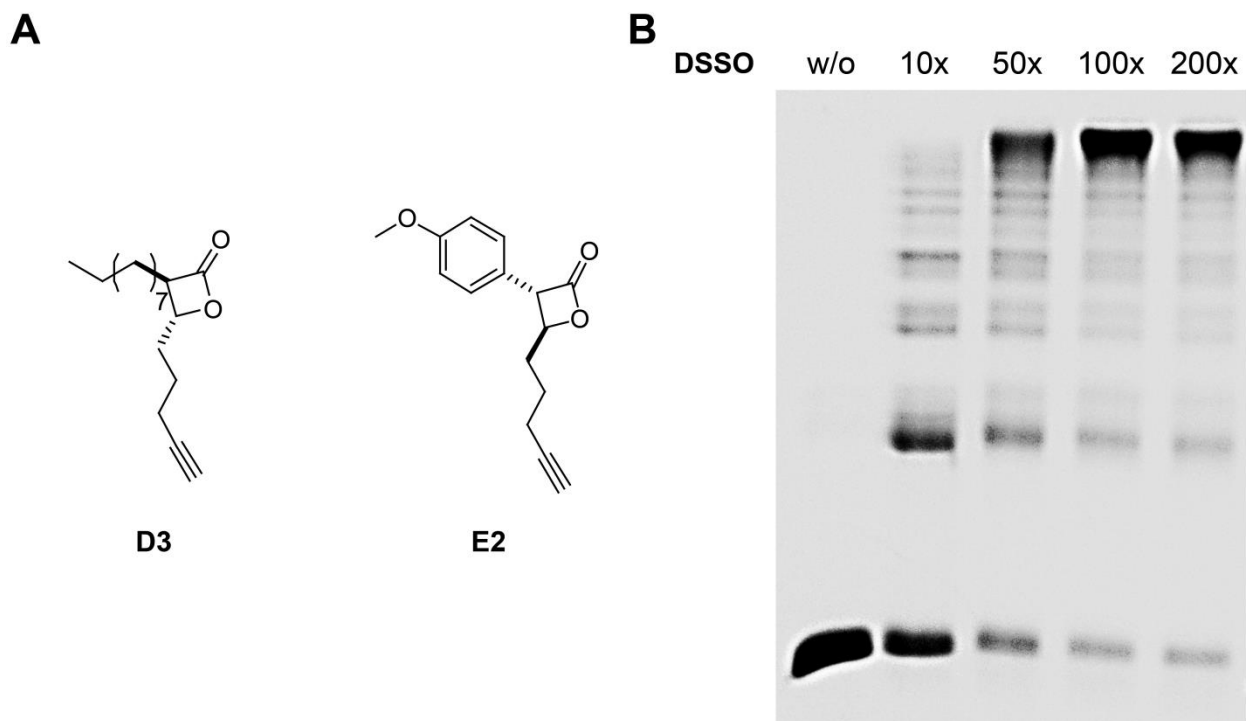
Figure I-25 Labeling of SaClpP with β-lactone probes. (**A**) Chemical structures of **D3** and **E2**. (**B**) **D3**-labeled SaClpP was cross-linked with increasing concentrations of DSSO, including a not cross-linked control. Afterwards, proteins were clicked to rhodamine-azide and fluorescence was detected after SDS-PAGE analysis.

8.2 Combined cross-link/ABPP with D3 and E2 in *S. aureus*

We incubated the obligate pathogenic *S. aureus* strain NCTC8325 with either 50 μM **D3** or **E2** probe, respectively, or DMSO for 1 h at 25 °C.[89] Probe- and DMSO-treated cells were incubated either with 2 mM DSSO for 1 h at 37 °C or with DMSO as control. Afterwards, cells were lysed, proteins were clicked to biotin azide, enriched on avidin beads, and digested with trypsin. Peptides were analyzed with LC-MS/MS and LFQ compared to the DMSO controls.

Labeling of *S. aureus* with **E2** without additional application of the cross-linker revealed ClpP (UniProt ID: Q2G036, marked in blue) to be one of the major targets (**Figure I-26A**). Within the same significance and enrichment cut-off values, previous labeling with **E2** resulted besides ClpP in only one additional target, which displays an uncharacterized protein (UniProt ID: Q2FWX5, marked in orange)[89] that we detected as well. However, our results show a markedly higher number of target proteins, which is most likely caused by LFQ,[123] as previous experiments were performed using dimethyl labeling.[175] This results underline substantial differences of the two quantification methods which were barely compared so far. As the LFQ approach allows for missing value imputation during downstream statistical evaluation, a higher number of probe-enriched proteins can be detected. In contrast, intensity ratios, which are the output from the data analysis of isotopically labeled samples, may lead to loss of several potential protein hits in the case that those hits were not present in the corresponding DMSO control. Comparison of enrichment targets applying **E2** alone and in combination with DSSO cross-linking (**Figure I-26B**) resulted in four overlapping proteins which account for around one-third of the total number of enriched proteins in each experiment. In the cross-link/ABPP experiment we did not detect any enrichment of the main SaClpP interactors ClpX or ClpC which was the case in the anti-hClpP and anti-EcClpP cross-link/co-IPs. Further, we did not enrich previously reported SaClpP substrates identified by the application of a catalytically inactive trap.[107] A possible explanation could be the disruption of ClpP interaction networks upon inhibition with **E2**, as the probe induces deoligomerization into heptamers.[89] It could have further been possible that DSSO is not able to pass the thicker cell wall of gram-positive bacteria as efficiently as it does in the case of gram-negative bacteria. Indeed, *in situ* chemical cross-linking with DSSO has been mainly conducted with gram-negative bacteria.[24,25] To date only one study is published applying *in situ* cross-linking in gram-positive bacteria. Thereby, another cell-permeable cross-linker was used at equal concentration as DSSO was applied in our experiments but at a shorter incubation time of 5 min.[28]
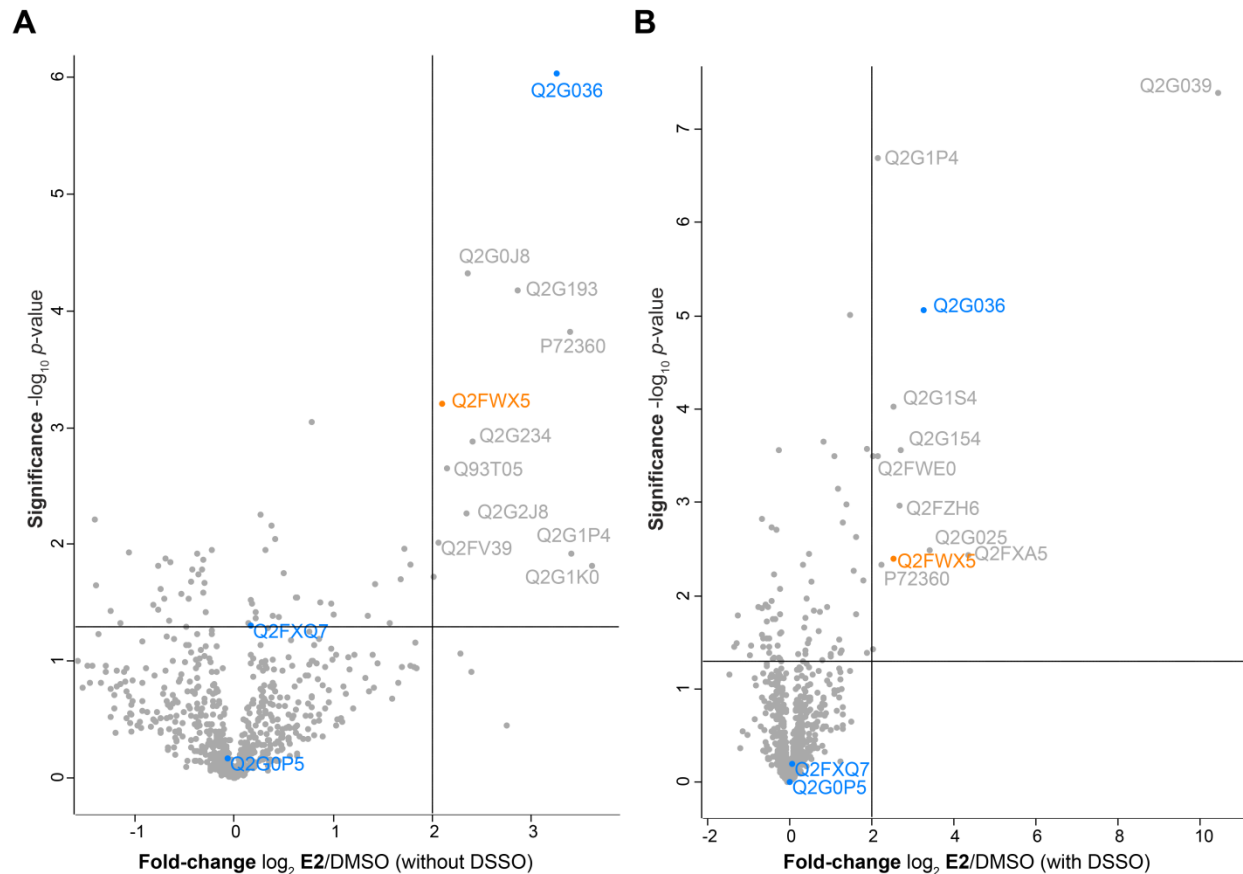
Figure I-26 Combination of ABPP with cross-linking using the E2 probe in *S. aureus*. Volcano plots representing *t*-test results of **E2** enrichment compared to DMSO control from cells without DSSO (**A**) or with DSSO (**B**) treatment (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 2$ (4-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). ClpP (UniProt ID: Q2G036), together with ClpX (UniProt ID: Q2FXQ7) and ClpC (UniProt ID: Q2G0P5) are colored in blue, previously reported **E2** targets in orange.[89]

Upon **D3** labeling in *S. aureus* without additional incubation with DSSO we observed ClpP as one of the most prominent hits (**Figure I-27A**) together with several previously identified target proteins (orange).[89] Notably, ClpP fold-changes upon labeling with **D3** were increased compared to those after labeling with **E2**. This observation can be explained by different degrees of modification by the β-lactone probes. **E2** binds to only 35% of the active sites whereas **D3** modifies 100%.[174] Compared to previous quantitative labeling experiments in *S. aureus* we observed an increased number of **D3** target proteins as already detected for the **E2** probe,[89] which is most likely caused by application of the LFQ strategy. Comparing the control results to **D3** labeling in combination with cross-linking (**Figure I-27B**) we observed an overlap of 11 proteins (about 50%). However, though ClpP was still one of the main targets, we did not obtain

any significant enrichment of ClpX and ClpC, or previously reported potential substrates detected by the inactive ClpP trap[107] which was the case for the combined cross-link/co-IPs.
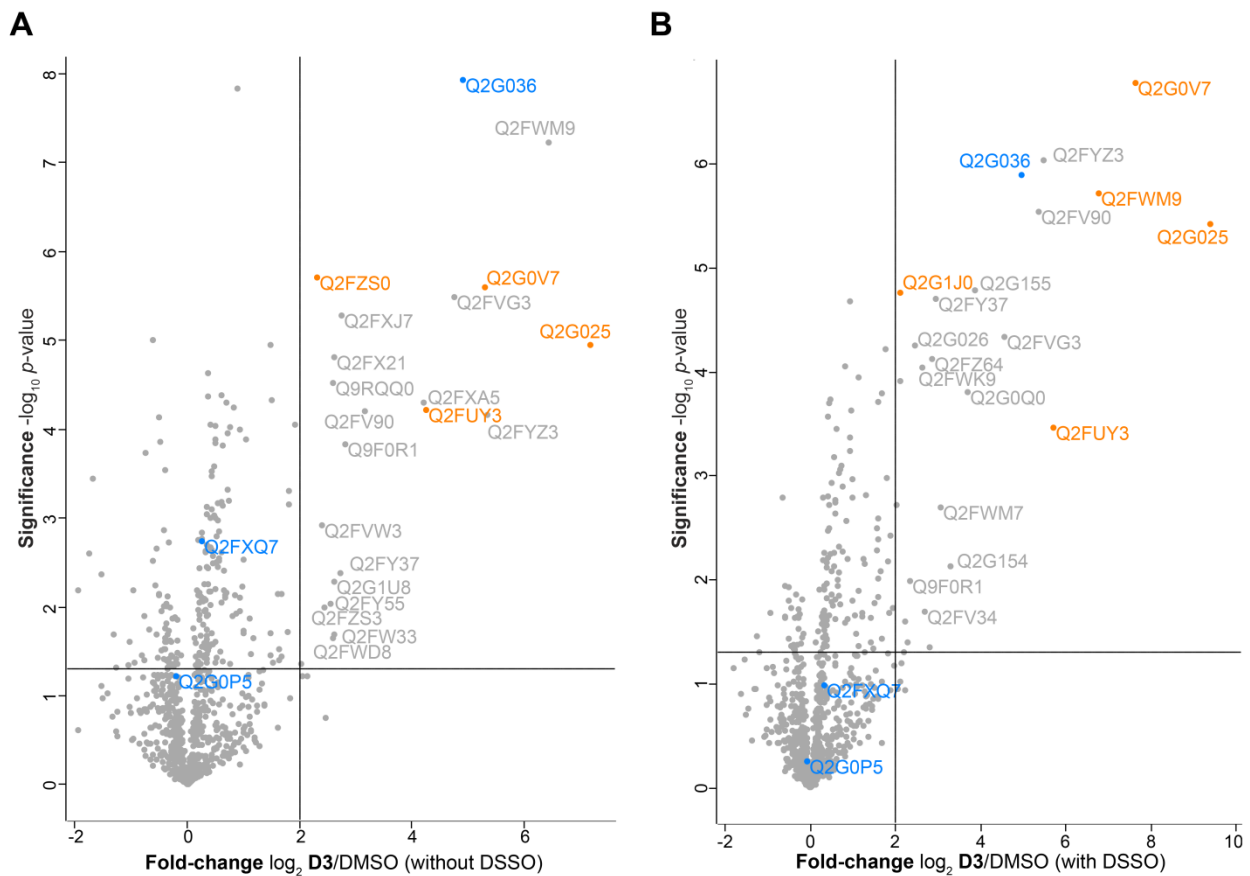
Figure I-27 Combination of ABPP with cross-linking using the D3 probe in *S. aureus*. Volcano plots representing *t*-test results of **D3** enrichment compared to DMSO control from cells without DSSO (**A**) or with DSSO (**B**) treatment (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 2$ (4-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). ClpP (UniProt ID: Q2G036), together with ClpX (UniProt ID: Q2FXQ7) and ClpC (UniProt ID: Q2G0P5) are colored in blue, previously reported **D3** targets in orange.[89]

Although **D3** is retaining the tetradecameric state of the protease,[174] it could nevertheless have been possible that the interaction network of ClpP got disturbed upon inhibition with the β-lactone probe. As mentioned above, another possibility could be that DSSO is not able to efficiently cross the gram-positive bacterial cell wall.

8.3 Combined cross-link/ABPP with D3 in *E. coli*

The initial target screen of β-lactone probes did not identify ClpP as target in *E. coli*[172] and progressive studies with the compounds were mainly conducted in pathogenic *S. aureus* strains.[88,89,173] However, as we wanted to examine whether a missing enrichment of prominent ClpP interactors as ClpX in combined cross-link/ABPP in *S. aureus* were due to ineffective entry of DSSO into the cells, we aimed at repeating the experiments in a gram-negative strain. Therefore, we chose *E. coli* K12, as we were able to receive superior co-IP results upon *in situ* chemical cross-linking, suggesting that DSSO efficiently enters the gram-negative bacteria.

To examine whether **D3** targets EcClpP, we first tested the inhibitory capacity of the β-lactone probe on the recombinant *E. coli* ortholog and compared the results to inhibition of SaClpP. Satisfyingly, we observed for both **D3** and **E2** effective inhibition of EcClpP (**Figure I-28A**). **D3** turned out to be an even more effective inhibitor of EcClpP than of SaClpP (**Figure I-28B**).
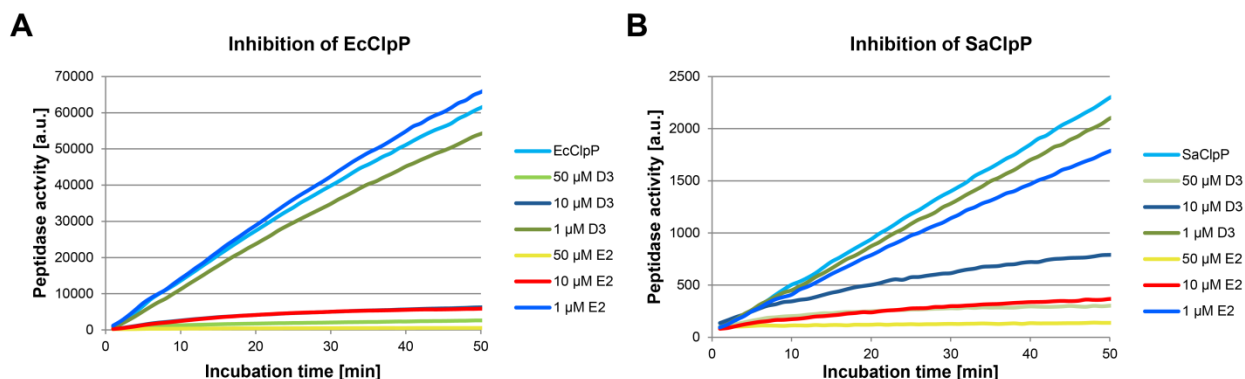
Figure I-28 Inhibition of EcClpP (A) and SaClpP (B) by β-lactone probes D3 and E2. Proteins (1 µM) were incubated with increasing amounts of probes and peptidase activity was followed over the time.

Having validated EcClpP inhibitors in hand, we proceeded with live cell labeling of *E. coli* with **D3**. We chose **D3** instead of **E2** for proteomic studies as it modifies 100% of the active sites[174] and therefore results in an enhanced ClpP enrichment as observed above. We applied the same probe concentration as well as incubation time as **D3** inhibited EcClpP as efficient as SaClpP. Interestingly, although ClpP was not reported as **D3** target in *E. coli* previously,[172] we observed significant enrichment of the protease as one of the main *in situ* targets without application of DSSO (**Figure I-29A**). Amongst others, two esterases (*yheT* and *ygiA*) were enriched by **D3** in *E. coli*. This is not a surprising finding, as esterases were the main targets besides ClpP in *S. aureus* as well.[89] A comparison of the control experiment with the cross-link/ABPP experiment

revealed no overlapping targets besides EcClpP (**Figure I-29B**). In general, the number of **D3** target proteins was markedly reduced upon DSSO addition. As already observed for the combined cross-link/ABPP enrichment in *S. aureus*, the additional cross-linking step did neither result in enrichment of the ClpP associated chaperones ClpX and ClpA nor in co-purification of any previously postulated substrates or interactors by the trapping approaches.[104,105]
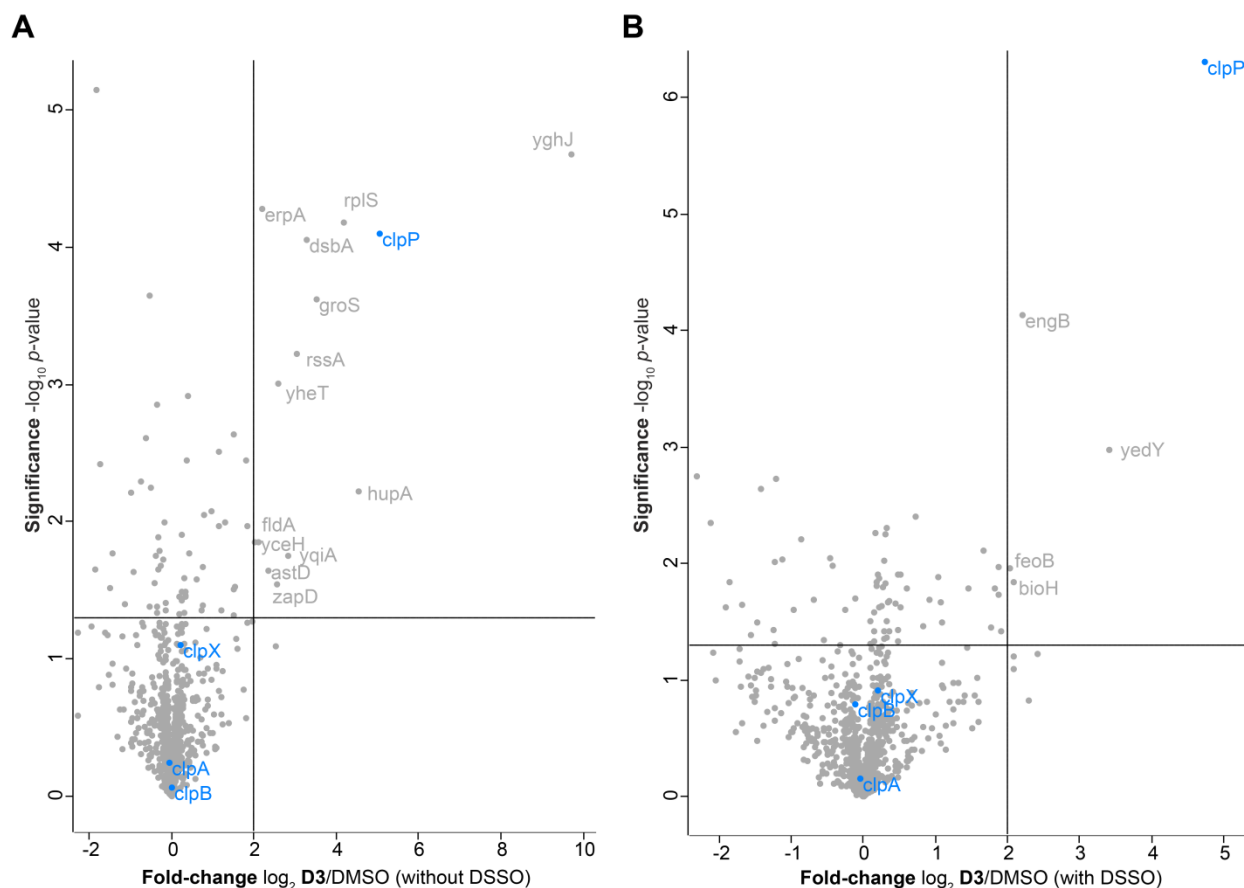
Figure I-29 Combination of ABPP with cross-linking using the D3 probe in *E. coli*. Volcano plots representing *t*-test results of **D3** enrichment compared to DMSO control from cells without DSSO (**A**) or with DSSO (**B**) treatment (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 2$ (4-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). ClpP, together with ClpX, ClpB, and ClpA are colored in blue, previously reported EcClpP interactors or substrates in orange.[104,105]

In conclusion, our results were similar in both bacterial strains and therefore seemed not to be caused by a hampered cell penetration of DSSO in gram-positive bacteria. Most likely, the interaction network of ClpP is already disrupted upon reaction with the inhibitor probes.

To test this hypothesis, we changed the order of cross-linking and incubation with the probe to previously fix cellular interactions. Thereby, we first incubated *E. coli* K12 with 2 mM DSSO for 1 h at 37 °C, then quenched the residual cross-linker, and finally applied 50 µM **D3** or DMSO as

a control for 1 h at 25 °C. Surprisingly, LC-MS/MS analysis and LFQ of the samples revealed that EcClpP was not enriched anymore (**Figure I-30**), indicating that targeting by the inhibitor is blocked upon cross-linking. It could be possible that **D3** is only able to enter the peptidase barrel when ClpP is not associated with an ATPase. A previous cross-linking step could fix most of the cellular ClpP complexes to one of its chaperones therefore preventing its inhibition.
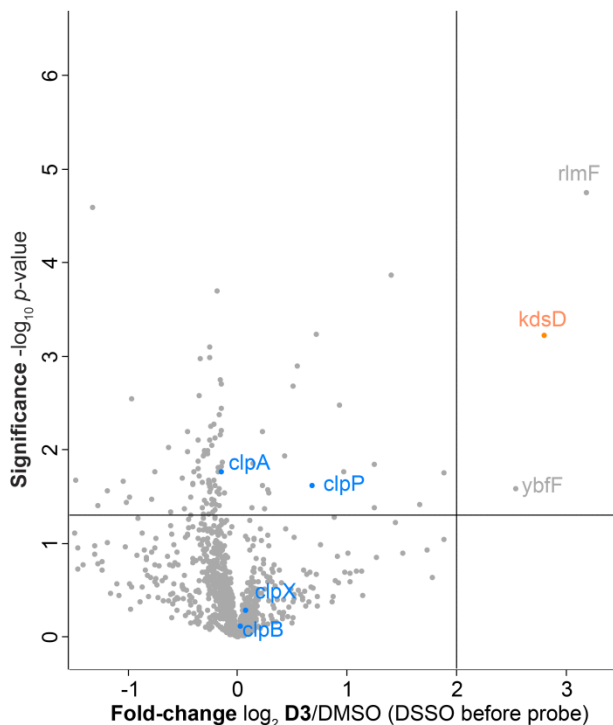
Figure I-30 Reverse combination of ABPP with cross-linking using the D3 probe in _E. coli_. Volcano plot representing _t_-test results of **D3** enrichment compared to DMSO control from cells with DSSO treatment previous to probe labeling (_n_=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 2$ (4-fold enrichment) and $-\log_{10}$ (_p_-value) of 1.3 (solid lines). ClpP, together with ClpX, ClpB, and ClpA are colored in blue, previously reported substrates or interactors in orange.[104,105]

8.4 Conclusion

We conducted here for the first time experimental workflows combining chemical cross-linking with the ABPP methodology. As ABPP probes enable the selective enrichment of their cellular target proteins, a cross-linking step preceding the pull-down should fix the cellular interaction network of target proteins and therefore will lead to simultaneous enrichment of their interactors in theory. We chose the β-lactone probes **D3** and **E2** as they were shown to specifically enrich ClpP with an only moderate number of _off_-targets in _S. aureus_.[89] Initial gel-based labeling

attempts displayed that **D3**-labeled and subsequently cross-linked ClpP complexes are still accessible for CuAAC. However, cellular labeling of *S. aureus* with **D3** and **E2** followed by chemical cross-linking with DSSO did not result in additional enrichment of high-confident interactors like ClpX or previously declared substrates as observed for the combined cross-link/co-IPs. We initially hypothesized that DSSO was not able to effectively cross the thicker cell wall of gram-positive *S. aureus*. Therefore, we transferred our workflow to the *E. coli* system, where we first demonstrated that recombinant EcClpP is as efficiently inhibited by β-lactone probes as SaClpP. We showed here for the first time that **D3** is able to target EcClpP *in situ* in a very selective manner. However, the combined cross-link/ABPP experiment again revealed no enrichment of ClpX, or previously reported substrates. Finally, to rule out that interaction networks of ClpP get completely disturbed upon inhibitor addition, we changed the order of probe labeling and chemical cross-linking. However, previous incubation of *E. coli* with DSSO prevented enrichment by **D3**, indicating that covalently fixed interaction networks protect the peptidase from reacting with the inhibitor.

Taken together, our experiments did not result in successful combination of ABPP with chemical cross-linking at least not for the system studied. However, as such a combination would be a useful progression for the community, it is worth conducting further experiments for instance with different probes and experimental set-ups.

9 Conclusion

PPI are of outstanding importance for cellular processes like signal transduction, DNA replication or transcription. However, studying PPI displays a challenging task as they are often rather transient and of weak nature. Several contacts are disturbed by cell lysis or downstream sample preparation workflows, and thereby artificial contacts may be generated.[26] Although MS-based proteomics became the method of choice for studying proteins even in very complex environments like the whole cell, information on protein structure, as well as topologies of high-molecular weight protein machines is often missing. XL-MS experiments generate distance restraints between surface exposed proximal amino acids from interacting proteins. Thereby, two reactive groups, for instance lysine targeting NHS-esters, that are joined by a spacer with a variable arm length built the cross-linker which connects the interacting proteins.[5] Cross-link restraints can be used to guide molecular modeling or docking experiments, as well as *de novo* protein structure prediction.[13,19] As the cross-linking reaction is carried out in solution where protein structures and interactions are highly dynamic, XL-MS is able to provide a snapshot from an ensemble of protein conformations and interactions under their native states. A manifold of XL-MS studies has been conducted over the decades, including reports on structures of multi-oligomeric complexes like the proteasome[39] or the chaperonin TRiC/CCT.[21] Moreover, not only recombinant protein complexes have been accessed, but also cross-linking contacts from whole proteomes were identified.[18] With the development of novel mass spectrometers and cross-linker chemistries, improved peptide fragmentation and identification using MS-cleavable cross-linkers became possible. Due to CID-labile bonds in the spacer arm, the linker can be fragmented in the mass spectrometer leaving peptide backbones intact, but generating separated peptides for downstream MS analysis.[58] With this new methodology in hand, several studies performed *in situ* cross-linking with cell membrane-permeable, MS-cleavable cross-

linkers as the DSSO linker and thereby reported on novel cellular interactions and processes.[27,176,177]

The ClpXP protease displays a highly conserved protein machinery with essential cellular functions. For instance, ClpP was demonstrated to play an important role in virulence processes of several human pathogens like *S. aureus*[87] and was found to be over-expressed in human cancer types[77] mediating resistance to chemotherapeutic agents such as cisplatin.[149] Although the protease was studied for decades and several crystal structures of bacterial orthologs as well as hClpP have been solved, a high-resolution structure of the complete proteolytic ClpXP complex is still lacking. Interactions between the heptameric ClpP peptidase and its hexameric associated chaperone ClpX have been mainly studied *via* mutations and molecular docking.[66,101,102,115] However, structural data about the interaction are not available. As the contact of a heptameric with a hexameric oligomer displays are structural mismatch, details about the interaction are of outstanding interest to the community.

In this part of the thesis we applied XL-MS to study the interaction between the peptidase ClpP and its associated ATPase ClpX. Therefore, we initially developed a robust and efficient chemical cross-linking workflow with the MS-cleavable linker DSSO[58] and the *S. aureus*, *E. coli*, and human ortholog of ClpP. In addition, we selected the heterooligomeric ClpP1/2 complex of *L. monocytogenes*[81,83,178] where it was not clear whether both parts ClpP1 and ClpP2 are able to interact with ClpX, or if the contact is mediated exclusively *via* one site. We mapped identified cross-linking contacts onto previously solved crystal structures of the ClpP orthologs and thereby confirmed that all identified links are within the maximum distance threshold of DSSO (35 Å).[39]

With a validated cross-linking strategy in hand, we next applied our workflow to study the interaction between ClpX and ClpP from *S. aureus*, human, and *L. monocytogenes*. Most of the identified cross-linking contacts occured between ClpX subunits. This is most likely caused by an increased flexibility of the ATPase along with a higher lysine density across the sequence compared to ClpP. However, especially for LmClpXP1/2 we detected several inter-ClpX-ClpP contacts, mainly mediated by flexible loops of ClpX. Our data consolidated previous assumptions from mutational and docking studies.[66,101,102,115] Moreover, we identified so far not recognized contacts. Taken together, the ClpXP interface revealed to be very flexible which makes it extraordinary suitable for analysis by XL-MS, as this methodology captures different conformations in solution.

In the next part we wanted to identify cellular binding partners of ClpP in a quantitative manner. We therefore applied co-IP to study networks of unmodified ClpP, as antibody-based pull-downs

allow for enrichment of the native proteins without the necessity of tags as required for affinity-based pull-downs (*e.g.* FLAG-tag). Cellular interaction partners of ClpP orthologs have been mainly studied by the application of catalytically inactive proteases that can still bind but no more degrade its corresponding substrates.[104,106,107] Although several potential substrates and interaction partners have been identified, these strategies nevertheless may suffer from their artificial working conditions, as they require heterologous expression of a tagged and inactive peptidase, while the natural ClpP genes have to be knocked-out. To compare whether co-IP leads to similar or additional potential substrates or interactors, we optimized the experimental conditions in a first step. We choose the hClpP system to address whether conventional co-IP, co-IP with on-bead cross-linking or *in situ* chemical cross-linking in combination with co-IP will result in the most confident network. Our data showed clearly that most of the hClpP interactions were lost upon conventional co-IP, as well as co-IP in combination with on-bead cross-linking using DSSO. This underlines that most contacts are rather transient or weak, leading to loss of numerous interaction partners and even in the generation of artificial contacts upon cell lysis. Combined cross-link/co-IP using the cell-permeable DSSO linker led to significant stabilization of the transient contacts and outperformed other co-IP strategies. We were able to increase the number of enriched proteins about 3-fold applying the combined strategy with 75% of those interactors are localized in the mitochondria, the organelle were hClpP is exclusively expressed.[72] In comparison, conventional co-IP revealed only 25% mitochondrial interactors, suggesting that several identified contacts may be false-positives. In addition, the combined experiment led to the enrichment of almost 50 previously declared substrates or interactors of hClpP, whereas conventional co-IP resulted in only four (**Figure I-31A**).[77,108,109]

A

Conventional co-IP — *In situ* cross-linking

| 33 (1) | 22 (3) | 132 (45) |

hClpP pull-down
all enriched proteins

Conventional co-IP — *In situ* cross-linking

| 7 (1) | 7 (3) | 108 (45) |

hClpP pull-down
mitochondrial proteins

B

Conventional co-IP — *In situ* cross-linking

| 11 (1) | 71 (10) | 124 (17) |

EcClpP pull-down
all enriched proteins
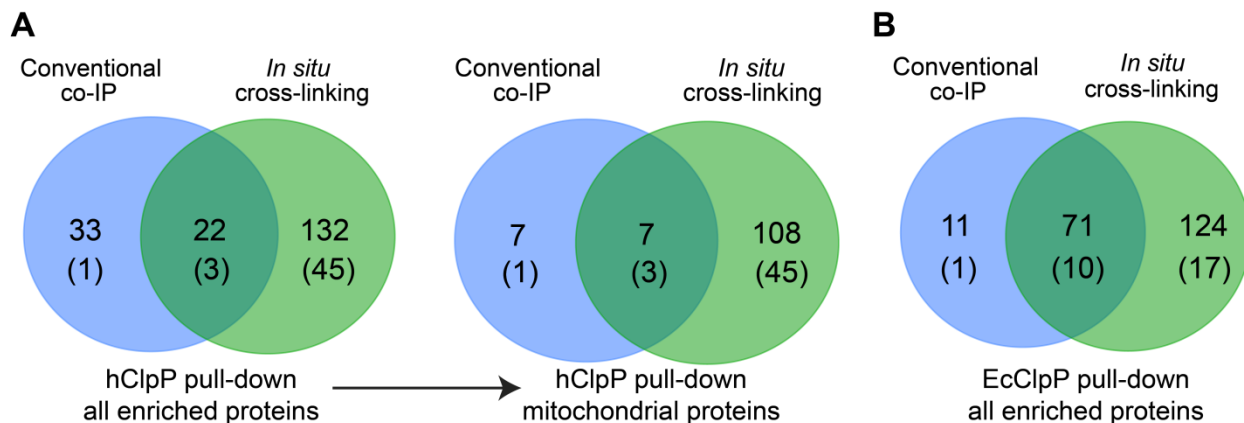
Figure I-31 Venn-Diagrams illustrating the overlap of significantly enriched proteins in conventional and combined cross-link/co-IP experiments of (A) hClpP and (B) EcClpP. The number of mitochondrial proteins detected in the anti-hClpP pull-down after application of the cross-linker

increased 18-fold. Further, we indicated the number of proteins which were previously declared as substrates or interaction partners of ClpPs in brackets.

Moreover, we identified several new potential interactors suggesting an essential and central role of ClpP for instance in cellular energy and protein homeostasis. To validate our approach, we confirmed the protease Lon as hClpP interactor by reverse cross-link/co-IP. As an example of application we compared interaction networks of hClpP in K562 with and without stimulation of the cells with the chemotherapeutic agent cisplatin.

We further transferred our optimized protocol to the study of EcClpP interaction partners, as it is one of the best characterized orthologs with several substrates already identified.[104] As for hClpP, the combined cross-link/co-IP experiments resulted in a significantly higher coverage of interacting proteins compared to co-IP without cross-linking (**Figure I-31B**). The most prominent interactors ClpX and ClpA were not or only slightly enriched in the conventional co-IP, but were one of the strongest enriched proteins after cross-link/co-IP, indicating that the interactions are transient and weak. As for the hClpP pull-down we detected several previously identified potential substrates of EcClpP,[104,105] verifying our approach. Moreover, a huge number of novel potential interactors suggested an important and central role of EcClpP for instance in flagellar housekeeping, chemotaxis, and energy metabolism.

Finally, we aimed at combining ABPP with *in situ* chemical cross-linking to enrich not only for the target(s) of the probe as in conventional ABPP, but in parallel also for its interacting partners. Therefore, we chose previously developed β-lactone-derived probes, which were demonstrated to selectively target ClpP in pathogenic *S. aureus*.[89] After incubation of bacteria with the probes, we inserted an additional incubation step with DSSO to fix cellular protein interaction networks. However, we were not able to achieve a significant enrichment of the prominent ClpP interactors ClpX, ClpC, or ClpA neither for **D3**, nor for **E2** probe treatment. This indicates that either the interaction networks are already disturbed after inhibitor addition, or that DSSO is not able to cross the thicker cell wall of gram-positive *S. aureus*. We therefore repeated experiments in gram-negative *E. coli* where we successfully applied our combined co-IP workflow. However, we did not achieve any enrichment of prominent EcClpP interactors or previously declared substrates, indicating that interactions get disturbed with application of inhibitor. Taken together, these were the first attempts to combine ABPP with chemical cross-linking as such a development would be useful for all reseachers in the field. However, we did not achieve any additional enrichment of target interacting proteins, at least not under the conditions examined.

10 Outlook

Nowadays XL-MS displays one of the methods of choice for studying PPI, protein complex compositions, as well as topologies. With the improvement of MS instrumentation and cross-linker chemistries even complex biological systems as whole tissues are getting accessible.[32] However, the most challenging part in the XL-MS field is the confident identification of especially low-abundance cross-linking contacts mostly resulting from proteins or complexes with comparable lower cellular expression levels. Although several strategies such as SCX or SEC for enrichment of cross-linked peptides exist, most contacts from whole cellular cross-linking experiments result from high abundant complexes as the proteasome. Ongoing optimization in cross-linker chemistry, sample preparation, MS-measurements, and downstream statistical analysis pave the way for studying low-abundant proteins in complex sample backgrounds. For instance, affinity-tagged cross-linkers as the PIR-linker allow for selective enrichment of cross-linked peptides to enhance their MS signal.

The ClpXP system turned out to be an extraordinary suitable candidate to be studied by XL-MS, here with the CID-cleavable cross-linker DSSO. Especially in cases were no high-resolution structures are available XL-MS displays an attractive tool for structural accession of complex protein machineries. As described above, bacterial ClpPs are not only associated with the chaperone ClpX, but also form complexes with other ATPases such as ClpA and ClpC. These high-molecular weight machines may also be interesting candidates for analysis with XL-MS, as no structural information about the peptidase-chaperone interface is available. Our optimized chemical cross-linking workflow can be applied to study any protein interaction of interest.

Further, we demonstrated here for the first time the advantages of combined chemical cross-link/co-IP over conventional co-IP in a direct and quantitative manner. Our workflow is straight-

forward and easy to apply, providing a huge output of information when compared to other experimental set-ups as affinity-enrichments. With this optimized methodology in hand the interaction network of any protein for which antibodies exist can be studied. Further, our comparative studies with hClpP confirmed that, depending on the protein of interest, a huge number of physiological interactions is lost during cell lysis and sample preparation workflows if not covalently trapped. Moreover, several interactions observed upon conventional methods may be false-positives. The most challenging part is, as described above, the low abundance of cross-linked peptides, leading to overall very low signal-to-noise ratios hampering confident downstream data analysis and cross-link site identification. In this thesis we already conducted several possible strategies for the simultaneous identification of cross-link sites within the co-IP data set. We tried to enrich for cross-linked peptides *via* SCX, we reduced the intensity thresholds for peptide sequencing during MS analysis, and we applied the enrichable/clickable cross-linker alkyne-disuccinimidyl bissulfoxide (DSBSO)[54] in varying experimental set-ups. However, we were not able to identify any meaningful cross-linking contacts which were in-line with the co-IP results or any previously reported interactions, probably caused by the nevertheless too low abundance of corresponding peptides. Possible solutions could therefore be an excessive increase in sample material, or the application of other cross-linker chemistries or MS set-ups.

Though we tried to combine chemical cross-linking with the ABPP methodology we were not able to achieve satisfying results at least not for the applied β-lactone probes targeting ClpP.[172] However, a successful combination of both strategies would revolutionize the field, as probe targets as well as corresponding target interacting proteins can be enriched and identified simultaneously. Such a new experimental set-up would be of outstanding interest to the community and picked up by experimenters in the field. Therefore, follow-ups on cross-link/ABPP experiments with different probes or experimental set-ups are important to achieve superior results.

11 Experimental section

11.1 Chemical synthesis

11.1.1 Synthesis of DSSO

Bis(2,5-dioxopyrrolidin-1-yl) 3,3'-thiodipropionate was synthesized according to Kao *et al* [58] from 3,3'-thiodipropionic acid (2.0 g, 11 mmol), *N*-hydroxysuccinimide (2.6 g, 23 mmol) and *N,N'*-dicyclohexylcarbodiimide (DCC) (4.6 g, 22 mmol) in dioxane (50 mL). The crude product was purified by silica gel column chromatography (CH_2Cl_2 : ACN = 1 : 0 to 4 : 1) yielding 2.4 g (59%) of the desired compound. The spectral data[110] are identical to those published in the literature.[58]

Bis(2,5-dioxopyrrolidin-1-yl) 3,3'-sulfinyldipropionate was synthesized following literature procedures.[58] 3-chloroperbenzoic acid (1.00 g, 4.7 mmol, 77%) in $CHCl_3$ (30 mL) was added to a solution of bis(2,5-dioxopyrrolidin-1-yl) 3,3'-thiodipropionate (1.7 g, 4.6 mmol) in $CHCl_3$ (90 mL) at 0 °C and stirred for 20 min at this temperature. The precipitate was filtered, washed with cold $CHCl_3$, cold MeOH and once more with cold $CHCl_3$. The crude product was finally purified by silica gel column chromatography (CH_2Cl_2 : CH_3CN = 4 : 1 to 0 : 1) and individual fractions were analyzed by LC-MS. Yield: 0.61g (34%) DSSO. The spectral data[110] are identical to those published in the literature.[58]

11.1.2 Synthesis of D3 and E2

D3 and **E2** were synthesized as described previously.[172]

11.2 Biological methods

11.2.1 Expression and purification of proteins

Expression and purification of recombinant proteins was performed as described previously. In brief, LmClpX equipped with an N-terminal His$_6$-tag and tobacco etch virus (TEV) cleavage site was overexpressed in *E. coli* BL21 (DE3) (Sigma, Cat# CMC0014) using a pET300 vector.[83] After purification of LmClpX *via* a Histrap column (GE Healthcare, Cat# 17-5248), the protein was digested with a TEV protease and loaded on a size exclusion column equilibrated with 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.6, 200 mM KCl, 1 mM dithiothreitol (DTT), 5% (v/v) glycerol, 0.5 mM ATP, 5 mM MgCl$_2$.[91] LmClpP1/2 was obtained using a dual-affinity purification strategy, with LmClpP1 bearing a C-terminally Strep-II tag and LmClpP2 with a C-terminal His$_6$-tag. Proteins were captured by immobilized metal affinity chromatography (IMAC) and subsequent StrepTactin chromatography (GE Healthcare, Cat# 28-9075-47) followed by purification with a size exclusion column by applying 20 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS) pH 7.0, 300 mM KCl, 15% (v/v) glycerol.[83,91] SaClpP, C-terminally Strep-II tagged, was expressed in *E. coli* BL21 (DE3) using a pET301 vector. After affinity purification, ClpP was purified *via* a size exclusion column (100 mM NaCl pH 7.0, 20 mM Tris(hydroxymethyl)aminomethane (Tris)-HCl).[114] EcClpP bearing a C-terminal Strep-II tag was expressed in *E. coli* BL21 (DE3) using a pET301 vector and was purified as described for SaClpP. N-terminally His$_6$-tagged SaClpX bearing an N-terminally TEV site was cloned into a pET301 vector and transformed into *E. coli* BL21 (DE3). After expression and affinity purification, the protein was digested with TEV protease overnight and subsequently loaded onto a size exclusion column equilibrated with 50 mM HEPES pH 7.5, 300 mM KCl, 15% (v/v) glycerol, 1 mM DTT.[179] hClpP equipped with a C-terminal Strep-II tag was cloned into the pET301 vector and transformed into Rosetta2 (DE3) (Merck, Cat# 71400) cells. After expression and affinity purification the protein was loaded onto a size exclusion column equilibrated with 20 mM HEPES pH 7.0, 100 mM NaCl.[180] hClpX fused to an N-terminal Strep-II tag was cloned into the pET301 vector and expressed in Rosetta2 (DE3) cells. After affinity purification the protein was loaded onto a size exclusion column equilibrated with 20 mM HEPES pH 7.0, 100 mM NaCl, 5 mM MgCl$_2$, 200 mM KCl, 1 mM DTT, 15% (v/v) glycerol.[180]

11.2.2 ClpP activity assay

ClpP activity was monitored following previously established experimental set-ups.[114,174] In brief, 1 µM ClpP was incubated with varying amounts of **E2** and **D3** from a 100-fold stock in DMSO for 20 min at 32 °C in 20 mM HEPES pH 7.6 and 100 mM KCl (EcClpP) or 20 mM HEPES pH 7.0 and 100 mM NaCl (SaClpP), respectively. Afterwards, 200-fold excess of substrate (Suc-Leu-Tyr-AMC) over ClpP was added and release of fluorescence product was followed *via* excitation at 380 nm and emission detection at 440 nm.

11.2.3 Cytotoxicity assay (MTT-assay)

K562 cells were cultured as described under 11.3. Cytotoxicity assays were conducted in 96-well plates. 8,000 cells were seeded in 50 µL media with fetal calf serum (FCS) per well. Cisplatin (Sigma, Cat# C2210000) was dissolved in PBS and pH was adjusted to 7.6. 1 µL of 100-fold cisplatin stocks were diluted with 49 µL media with FCS and added to the cells. A no compound control was included. Incubation carried out for 48 h at 37 °C. Afterwards, 20 µL of an MTT stock (5 mg/mL in PBS, Sigma, Cat# M5655) were added per well and cells were incubated for another 4 h previous to lysis (50 µL of 10% (w/v) SDS, 5% isobutanol (v/v) in 0.012 M HCl per well). After incubation for 16 h at 37 °C, absorbance of reaction product was measured at 570 nm.[150] Triplicates per condition were prepared. Values were normalized to the no-cisplatin control, averaged and plotted against $\log_{10}$ transformed concentration values. Data were fit in Prism 6 (GraphPad) with non-linear regression applying the *log(inhibitor) vs. response - variable slope (four parameters)* function to calculate the IC_{50} value and corresponding error.

11.3 Cell culture

HepG2 (male, ECACC *via* Sigma Aldrich, Cat# 85011430) and K562 (female, ECACC *via* Sigma Aldrich, Cat# 89121407) cells were cultivated in RPMI-1640 media supplemented with 10% L-glutamine and 10% FCS at 37 °C and humidified 5% CO_2 atmosphere. The cells were routinely tested for mycoplasma contamination.

11.4 Microbe strains

E. coli K12 (DSMZ, Cat# 498) were grown in lysogeny broth (LB) media and *S. aureus*

NCTC8325 (*via* Public Health England) in Bushnell Haas broth (BHB) media at 37 °C with shaking at 200 rpm until reaching the exponential phase (OD_{600nm} 0.7).

11.5 XL-MS

11.5.1 Cross-linking of recombinant ClpP orthologs

Crosslinking reactions with all ClpP orthologs were conducted in 25 mM HEPES pH 7.6, 200 mM KCl, 1 mM DTT, 5% (v/v) glycerol, 0.5 mM ATP, 5 mM $MgCl_2$ either for 2 h at 37 °C (with shaking at 300 rpm) or at 4 °C overnight (without shaking). The concentration of ClpP orthologs was 5 µM (based on monomer) and concentrations of DSSO (100 mM stock in DMSO) were varied as indicated. A homogenous solution was observed at all DSSO concentrations tested. Reaction volumes were 50 µL (for SDS-PAGE analysis) or 250 µL (for SEC separation). The reactions were quenched by adding 20 mM NH_4HCO_3 followed by 10 min incubation at r.t. For SDS-PAGE analysis samples were mixed with 2x loading buffer, boiled for 3 min at 95 °C and loaded onto a NuPAGE Tris-Acetate mini gel (Thermo Fisher Scientific, Cat# EA0375BOX) together with a pre-stained high molecular weight marker (Thermo Fisher Scientific, Cat# LC5699). Proteins were stained with SYPRO Ruby Protein Gel Stain (Thermo Fischer Scientific, Cat# S12000). For separation of reaction products by SEC, samples were loaded onto an analytical Superose 6 Increase 10/300 GL column (GE Healthcare, Cat# GE29-0915-96) connected to an ÄKTA system (GE Healthcare) equilibrated with 20 mM HEPES pH 7.6, 200 mM KCl. The molecular weights of the complexes were determined *via* calibration curves. Fractions containing cross-linked complexes with the correct integrity were collected and evaporated *in vacuo* for MS sample preparation. Similarly, the EcClpP ortholog was cross-linked with varying excesses of BS3 (100 mM stock in water, Creative Molecules, Cat# 001SS) as described for the DSSO cross-linker.

To study the impact of cross-linking on **D3** click-efficiency we additionally incubated ClpP orthologs with equimolar amounts of inhibitor for 1 h at r.t. previous to DSSO addition. After cross-linker quenching, SDS was added to a final concentration of 0.4% from a 10% stock in PBS. CuAAC carried out by addition of 100 µM rhodamine-azide (from a 5 mM stock in DMSO), 1 mM $CuSO_4$ (from a 50 mM stock in ddH_2O), 100 µM tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, Sigma, Cat# 678937) ligand (from a 83 mM stock in DMSO) and 1 mM tris(2-carboxyethyl)phosphine (TCEP) (from a 50 mM stock prepared fresh in ddH_2O) for 1 h in

the dark. Reaction was quenched by the addition of 2x loading buffer and samples were analyzed *via* SDS-PAGE.

11.5.2 Cross-linking of recombinant ClpXP orthologs

Cross-linking reactions were carried out using wild-type ClpP and the Walker B mutant of ClpX in order to stabilize the complex.[117] The *S. aureus* system was used for optimization of cross-linking conditions. A concentration of 10 µM of SaClpP (based on monomer) was incubated with 17.14 µM SaClpX (based on monomer) in 250 µL SaClpXP buffer (25 mM HEPES pH 7.6, 200 mM KCl, 1 mM DTT, 5% (v/v) glycerol, 0.5 mM ATP, 5 mM $MgCl_2$). The mixture was incubated for 10 min at 32 °C. Afterwards, DSSO, varying from 40- to 812.5-fold molar excess over SaClpP monomer, was added to the complex (from 100 or 250 mM stocks) and the reaction proceeded for 2 h at 37 °C with shaking at 300 rpm or overnight at 4 °C without shaking. A homogenous solution was observed at all DSSO concentrations tested. The LmClpXP1/2 complex was cross-linked in 25 mM HEPES pH 7.6, 200 mM KCl, 1 mM DTT, 5% (v/v) glycerol, 0.5 mM ATP, 5 mM $MgCl_2$. 10 µM LmClpP1/2 (based on monomer concentration) were incubated with 8.33 µM LmClpX (based on monomer concentration) for 10 min at 32 °C in cross-linking buffer. DSSO was added to an 80-, 100- or 137.5 molar excess (from a 250 mM stock), respectively, over LmClpP monomer and cross-linking was carried out for 2 h at 37 °C with shaking at 300 rpm. Crosslinking of the hClpXP complex was performed in 50 mM HEPES pH 8.0, 100 mM KCl, 10% (v/v) glycerol, 10 mM $MgCl_2$, 1 mM ATP and 1 mM DTT. A concentration of 10 µM of hClpP (based on monomer) was incubated with 17.14 µM hClpX (based on monomer) in 250 µL cross-linking buffer. After pre-incubation of 10 min at 37 °C, an excess of 80x, 100x and 137.5x of DSSO (based on hClpP monomer concentration) was added from a 100 mM stock and cross-linking carried out for 2 h at 37 °C with shaking at 300 rpm. The reaction products were separated using an analytical Superose 6 column equilibrated with 20 mM HEPES pH 7.6, 200 mM KCl. Fractions containing the intact ClpXP complexes were collected and evaporated *in vacuo* for MS preparation. The molecular weights of the complexes were determined *via* calibration curves. Additionally, in order to validate that the cross-linking reaction was successful and the ClpXP complex has the correct integrity after reaction with DSSO, ClpP and ClpX were incubated as described above without DSSO as reference. The untreated ClpXP protease was separated on the analytical Superose 6 equilibrated with the corresponding cross-linking buffer, as the complex requires ATP to stay intact.[181]

11.5.3 Sample preparation for MS

Lyophilized samples were taken up in 100 µL denaturation buffer (7 M urea, 2 M thiourea in 20 mM HEPES pH 7.5) and DTT was added to a final concentration of 1 mM. Samples were reduced for 1 h at 37 °C followed by alkylation with 3 mM iodoacetamide (IAA) for 30 min at r.t. IAA was quenched with 3 mM DTT (final concentration) for 30 min at r.t (all steps were performed under shaking at 450 rpm). Lys-C (Wako, Cat# 125-05061) was added (1:30 (w/w) and samples were digested for 4 h at r.t. under shaking at 450 rpm. Samples were diluted 4-fold with 50 mM triethylammonium bicarbonate (TEAB) and incubated with trypsin (1:30 (w/w), sequencing grade, modified; Promega, Cat# V5111) overnight at 37 °C (shaking at 450 rpm). In order to quench the reaction, formic acid (FA) was added to a final amount of 0.5% (v/v). The ClpP samples were enriched *via* cation-exchange chromatography and desalted with single layer stage-tips (ClpXP samples with two layers due to higher peptide amount). Cation exchange enrichment was carried out using C18 stage tip disks (Empore disk-C18, 47 mm, SUPELCO, Cat# 66883-U) packed on top of cation extraction disks (Empore Cation Extraction 47 mm Disks, SUPELCO, Cat# 66889-U). All following volumes correspond to enrichment and desalting with one layer disks, volumes were doubled for two layer disks. Stage tips were conditioned with (1) 20 µL methanol (MeOH), (2) 20 µL 80% acetonitrile (ACN), 20% water and 0.5% FA, (3) 20 µL aqueous 0.5% FA, (4) 20 µL 1 M NH$_4$OAc, 0.5% FA, 80% water and 20% ACN, and (5) 20 µL aqueous 0.5% FA. Samples were loaded onto the stage tips and washed with 20 µL aqueous 0.5% FA prior to transfer-elution with 20 µL 80% ACN, 20% water and 0.5% FA. Fractionation was carried out by washing with 40 µL 50 mM NH$_4$OAc, 0.5% FA, 80% water and 20% ACN followed by elution with 2x10 µL 1 M NH$_4$OAc, 0.5% FA, 80% water and 20% ACN. Cation exchange enriched samples were further desalted with a C18 stage tip, conditioned with (1) 20 µL MeOH, (2) 20 µL 80% ACN, 20% water and 0.5% FA and (3) 20 µL aqueous 0.5% FA. Afterwards, 60 µL aqueous 0.5% FA were added onto the tip. The 20 µL eluate from the cation exchange column were mixed with the 60 µL aqueous 0.5% FA and passed throw the tip. After washing with 20 µL aqueous 0.5% FA, peptides were eluted with 2x30 µL 80% ACN, 20% water and 0.5% FA followed by evaporation *in vacuo*. For MS analysis, samples were taken up in 25 µL 1% FA in water, pipetted up and down, vortexed and sonicated for 15 min. Finally, the peptide solution was transferred on a filter with 0.2 µm pore size.

11.5.4 MS analysis of DSSO cross-linked proteins

Analysis of the DSSO cross-linked samples was performed using a UltiMate 3000 nano HPLC system (Thermo Fischer Scientific) equipped with an Acclaim C18 PepMap100 75 µm ID x 2 cm

trap (Thermo Fisher Scientific, Cat# ES803A) and an Acclaim PepMap RSLC C18 separation column (75 μm ID x 50 cm; Thermo Fisher Scientific, Cat# 164535) coupled to an EASY-source equipped Thermo Fischer LTQ Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). Samples were loaded onto the trap column with a flow rate of 5 μL/min in aqueous 0.1% trifluoroacetic acid (TFA) followed by a transfer onto the separation column at a rate of 0.3 μL/min. Buffers for the nano-chromatography pump were aqueous 0.1% FA (buffer A) and 0.1% FA in ACN (buffer B). Samples were separated using a gradient increasing buffer B from 5 to 32% over 75 min. Buffer B was further increased to 90% during the next 10 min and held another 10 min at 90%. Buffer B was decreased to 5% within 0.1 min and held until end of the run (total: 105 min).

For EThcD-CID analysis, MS full scans were performed at 60,000 m/z resolution in the orbitrap with quadrupole isolation and internal calibration was performed using the lock mass option. The scan range was from 375 to 1,575 m/z, and RF lens amplitude was set to 60%. The automatic gain control (AGC) target was 4.0 e5 with a maximum injection time of 50 ms. Peptides with at least 5.0 e4 intensity and charge states from 4 to 10 were selected for MS^2 scans with a dynamic exclusion of 60 s (mass tolerance 10 ppm high and low). The MS instrument was operated in a 5 s top speed mode. Two MS^2 scans were performed in the orbitrap on selected precursors. Fragmentation occurred with a CID collision energy of 25% and an isolation window of 1.6 m/z. Orbitrap resolution was set to 30,000 m/z and AGC target was 5.0 e4 during a maximum injection time of 100 ms. The second fragmentation was performed using EThcD with a collision intensity of 15%, using calibrated charge-dependent ETD parameters and ETD supplemental activation. Here, the MS^2 scan was performed with a resolution of 30,000 m/z, an isolation window of 1.6 m/z and an AGC target of 1.0 e5. Maximum injection time was 120 ms and the first mass was 120 m/z.

For MS^2 - MS^3 analysis full scans were performed at 60,000 m/z resolution in the orbitrap using an isolation window of 1.6 m/z and quadrupole isolation. The amplitude of the RF lens was 60% with a scan range of 350 to 1,500 m/z. The maximum injection time was 50 ms and AGC target was 4.0 e5. Internal calibration was performed using the lock mass option. Peptides with charge states 4 to 10 and a minimum intensity of 5.0 e4 were selected for fragmentation with a CID energy of 25%. The isolation window was set to 1.6 m/z. MS^2 scans were recorded in the orbitrap at 30,000 m/z resolution and a dynamic exclusion time of 60 s. The AGC target was set to 5.0 e4 and the maximum injection time was 100 ms. After MS^2 fragmentation, peptides with charge states 2 to 6 and undetermined charge states were investigated for their targeted mass difference. If two precursors differed exactly 31.9721 (mass difference of the two fragmentation

possibilities for DSSO) with 30 ppm low and high mass tolerance, they were both selected for MS^3 fragmentation with an exclusion mass width of 5 m/z (low and high) and an HCD energy of 30%. MS^3 scans were performed in the ion trap with an isolation window of 2 m/z. The ion trap was operated in rapid mode and first mass was set to 120 m/z. The AGC target was set to 2.0 e4 with a maximum injection time of 150 ms. The MS instrument was operating in a 5 sec top speed mode oscillating from MS^3 to full scan. After 4 scans the instrument switched from MS^3 to MS^2 scan.

11.5.5 MS analysis of BS3 cross-linked proteins

Analysis of the BS3 cross-linked samples was performed using a UltiMate 3000 nano HPLC system (Thermo Fischer Scientific) equipped with an Acclaim C18 PepMap100 75 µm ID x 2 cm trap (Thermo Fisher Scientific, Cat# ES803A) and an Acclaim PepMap RSLC C18 separation column (75 µm ID x 50 cm; Thermo Fisher Scientific, Cat# 164535) coupled to an EASY-source equipped Thermo Fischer Q Exactive Plus mass spectrometer (Thermo Fisher Scientific). Samples were loaded onto the trap column with a flow rate of 5 µL/min in aqueous 0.1% TFA followed by a transfer onto the separation column at a rate of 0.3 µL/min. Buffers for the nano-chromatography pump were aqueous 0.1% FA (buffer A) and 0.1% FA in ACN (buffer B). Samples were separated using a gradient increasing buffer B from 5 to 40% over 122 min. Buffer B was further increased to 90% in the next 10 min and held another 10 min at 90%. Buffer B was decreased to 5% in 0.1 min and held until end of the run (total: 152 min). Full scans were performed in the orbitrap mass analyzer with a resolution of 70,000 m/z and a scan range of 300 to 1,500 m/z. The AGC target was set to 3.0 e6 within a maximum injection time of 80 ms. Precursor fragmentation occurred with an HCD collision energy of 27%. MS^2 scans were recorded in the orbitrap at a resolution of 35,000 m/z. The AGC target was set to 5.0 e4 within a maximum injection time of 60 ms. The isolation window was 1.6 m/z and the Q Exactive was operating in a Top$N$ mode of 12. Peptides with unassigned charge states or charge states of 1 to 3 were excluded from the MS^2 fragmentation. Dynamic exclusion was set to 60 s.

11.5.6 MD simulations

An initial structure of the LmClpX-ClpP2 hexamer/heptamer complex was obtained as follows. The structure of the LmClpP2 monomer was first extracted from the crystal structure (PDB: 4RYF) to add missing residues with SWISS-MODEL. The resulting monomer structure was then realigned with PyMOL 1.7.2.1[182] to the other subunits in the complex as given in the crystal structure, yielding a complete LmClpP2 heptamer. For LmClpX, we used SWISS-MODEL to

create a homology model based on the corresponding *E. coli* crystal structure (PDB: 4I81). A separate model was created for each chain in the hexamer, and the resulting models were combined to yield the complex structure. These LmClpP2 heptamer and LmClpX hexamer models were used as inputs for the protein-protein docking webserver HADDOCK,[183] using the template-free setting. In addition, experimental cross-links were set as restraints with an upper limit of 35 Å. The solution with the best stacking of the substrate channels was chosen and the relative heptamer/hexamer orientation was manually slightly readjusted for optimal substrate channel stacking.

From the initial LmClpX-ClpP2 complex, two adjacent LmClpP2 and LmClpX trimers were extracted for MD simulations using the CUDA implementation of the Amber16 software package.[184] The ff14SB[185] force field was used together with the TIP3P[186] water model, and parameters of Meagher et al.[187] for ATP. The complex was placed in a rectangular water box with a minimum protein - boundary distance of 12 Å and ions were added to neutralize the total charge. All bonds involving hydrogens were constrained using SHAKE[188] and periodic boundary conditions were applied combined with Particle Mesh Ewald summation using a nonbonded cutoff of 12 Å. The system was subjected to energy minimization using the XMIN method with a convergence criterion of $0.01 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-1}$. Afterwards the complex was heated up from 0 to 300 K over 1 ns using a stepwise approach (5, 10, 20, 50, 100, 200 and 300 K). During heat up a timestep of 1 fs was used together with a Langevin thermostat ($\gamma = 4.0 \text{ ps}^{-1}$), backbone restraints ($3.0 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$) until 200 K, and Berendsen pressure coupling in the last step at 300 K ($\tau = 2.0$ ps). After heat up a timestep of 2 fs and the slow-coupling Berendsen thermostat ($\tau=10$ ps) and barostat ($\tau=5$ ps) were applied for all following simulations. In total, eleven replicates were simulated for a total of $11 \cdot 40$ ns of simulation time. All MD based analyses presented here are averages over the simulations data of the last 10 ns of all replicates. Simulation postprocessing and analysis were conducted with Cpptraj.[189]

11.5.7 Analysis of cross-linking data

MS raw data for DSSO cross-linked proteins were analyzed using Proteome Discoverer Version 2.2 (Thermo Fisher Scientific, Cat# OPTON-30795) and its XlinkX plug-in. The database contained ClpP and ClpXP sequences, as well as all common contaminants from the MaxQuant (version 1.6.0.1)[124] bin folder to receive statistically significant results. Spectra were analyzed if their S/N ratio was 1.5 or higher. Cross-links were searched against the corresponding database allowing a maximum number of missed cleavages of 2 using the enzyme trypsin. Peptide length had to be at least 5 residues and the maximum number of modifications was set to be 4. Mass

tolerances were 20 ppm for FTMS and 0.5 Da for ITMS with a precursor tolerance of 10 ppm. Carbamidomethyl was set as static modification, oxidation of methionine as dynamic modification. For validation of cross-link identification the Percolator algorithm was used with a FDR threshold of 0.01. Peptides were searched against the corresponding database allowing two missed cleavages using the enzyme trypsin. Precursor mass tolerance was 10 ppm and fragment ion tolerance was set to 0.04 Da. Maximum equal dynamic modifications were set to be 3. Besides oxidation of methionine, all common DSSO modifications were specified as dynamic ones. Carbamidomethyl was set as static modification. Peptide spectrum matches having a maximum deviation of 0.05 were used for Percolator evaluation with a FDR of 0.01. Peptide validation was performed with a FDR of 0.01. Proteins identified with a FDR of 0.01 were reported as valid. Minimum score threshold of ClpP-only analysis was set to be 20, for ClpXP analysis to be 30 to be more stringent as -contrarily to ClpP- high resolution crystal structures are missing for direct validation. In addition, spectra were inspected manually. We combined the data from different fragmentation strategies and different cross-linker excesses for each ClpP ortholog.

Data from samples cross-linked with the non-cleavable cross-linker BS3 were analyzed using the Kojak software (version 1.5.5).[113] The database contained the sequence of EcClpP and a reversed sequence of EcClpP for decoy analysis. Cross-linker was set to be BS3, meaning cross-links were searched between lysine residues and peptides with a mass addition of 138.068074 Da. Fixed modification was set to be carbamidomethylation of cysteine, variable modification was set to be oxidation of methionine. The number of modifications was limited to 2 per peptide. Enzyme was specified to be trypsin with a maximum number of missed cleavages of 2. Precursor mass tolerance was set to be 10 ppm. Maximum peptide mass was set to be 8000 Da, minimum to be 500 Da. We combined the data from different cross-linker excesses.

For all ClpP and ClpXP samples, two measurements were conducted, which were performed with the two fragmentation strategies for DSSO to get a high coverage of cross-links and generate independent datasets. The detected cross-links were combined and mapped onto ClpP crystal structures. The structure of the EcClpX hexamer in the nucleotide bound state (PDB: 4I81) was used as template, on which the sequences of the LmClpX, SaClpX, and hClpX were modeled. For this purpose, the sequence of the *E. coli* crystal structure was aligned with the sequences of LmClpX, SaClpX, and hClpX to generate the highest possible consensus. The model crystal structures were then generated based on the template crystal structure using SWISS-MODEL.[190–193] Cross-links were displayed onto the structures using the XLinkAnalyzer[15] (version 1.1) plug-in into Chimera[194] (version 1.11) applying a distance threshold of 35 Å for

DSSO and BS3 linker, respectively. The smart homo-oligomer mode was activated, meaning that only cross-links are displayed which can be bridged with respect to the distance threshold of the cross-linker. In the case that the cross-link can't be realized within this distance threshold in the crystal structure or model used, the shortest possible distance to build the link is reported. Distance distributions reported derived from median Cα-Cα distances of cross-linked lysine residues.

11.6 co-IP workflows

11.6.1 Preparation of cells

HepG2 and K562 cells were cultivated as described under 11.3. Cells were grown in plates until 80-90% confluency. One separate plate was seeded per replicate. *E. coli* K12 were grown as described under 11.4. One flask was prepared from which the replicates were conducted. Per condition (*e.g.* DSSO cross-linked cells enriched by anti-human ClpP antibody) four replicates were prepared. For cisplatin experiments K562 were additionally incubated with 2 μM Cisplatin (1000-fold stock in PBS dissolved in media with FCS; pH was adjusted) for 48 h. For *in situ* cellular cross-linking human cells were washed once with PBS (25 °C) and then detached in 25 °C PBS (adherent HepG2 cell line). After harvesting at 1000 rpm at r.t., cells were resuspended in either 1 mL 25 °C PBS containing 2 mM DSSO (from a 100 mM DSSO stock in DMSO) or in 1 mL 25 °C PBS containing corresponding DMSO amounts (only HepG2) and incubated at 37 °C and with shaking at 200 rpm for 1 h. Afterwards, cells were harvested (1000 rpm, 4 °C) and washed two times with 50 mM Tris-HCl pH 8.0 in order to quench the reaction and remove excess cross-linker. For *in situ* cellular cross-linking bacteria were harvested at 4,000 rpm and 4 °C for 5 min and washed twice with 25 °C PBS. Per replicate bacteria were resuspended in either 1 mL 25 °C PBS containing 2 mM DSSO (from a 100 mM DSSO stock in DMSO) or in 1 mL 25 °C PBS containing corresponding DMSO amounts to receive an OD_{600nm} of 20. Cells were incubated at 37 °C under shaking at 200 rpm for 1 h. Afterwards, cells were harvested (4000 rpm, 4 °C) and washed two times with 50 mM Tris-HCl pH 8.0 in order to quench the reaction and remove excess cross-linker.

11.6.2 Cell lysis and enrichment *via* antibody

Human cells were resuspended in 1 mL lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 5% (v/v) glycerol, 1% (v/v) NP-40, 1 mM $MgCl_2$) and incubated for 30 min at 4 °C under rotation. *E. coli* K12 cells were resuspended in 1 mL lysis buffer (50 mM Tris-HCl pH 7.4, 5% (v/v)

glycerol, 0.1% NP-40) and lysed by sonication with cooling intervals on ice (3x30 s, 75% intensity). The supernatant was centrifuged for 45 min at 4 °C and 13,0000 rpm. The protein concentration was adjusted with a bicinchoninic acid (BCA) assay (Roti Quant, Roth, Cat# 0120) to obtain 500 µL containing 500 µg protein per replicate. 30 µL Protein A/G agarose beads (Thermo Fisher Scientific, Cat# 20421) were equilibrated with co-IP wash buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 5% (v/v) glycerol, 0.05% (v/v) NP-40). For this purpose, beads were resuspended in 1 mL wash buffer and centrifuged for 1 min at 1000 rpm and 4 °C. 500 µL lysate were added directly to the beads and either 5 µL of the anti-hClpP antibody (1 mg/mL, Abcam, Cat# ab124822), 3.13 µL of the anti-Lon antibody (1.6 mg/mL, Thermo Fisher Scientific, Cat# PA5-50823) or 5 µL of anti-SaClpP (2 mg/mL, polyclonal antibody generated in rabbit against SaClpP, reactivity tested for EcClpP, custom made) were added. To evaluate background binding towards the beads and the antibody constant regions, the same amount of isotype control (Cell Signaling Technology, Cat# 3900) was added to the control samples. The mixture was incubated for 3 h at 4 °C under constant rotation. For on-bead cross-linking 2 mM DSSO (25 µL from a 100 mM stock in DMSO) were added and lysate was cross-linked for 2 h at 4 °C while rotating. DSSO was quenched by adding ammonium bicarbonate to a final concentration of 20 mM from a 1 M stock in ddH$_2$O. Afterwards, samples were centrifuged (1,000 rpm, 4 °C) and supernatant was discarded. To remove unspecific bound proteins, 1 mL wash buffer was added and samples were centrifuged. The procedure was repeated. To remove detergent, 1 mL 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 5% (v/v) glycerol was added and samples were centrifuged. This step was repeated once.

11.6.3 Sample preparation for MS

After co-IP enrichment, samples were reduced and digested in 25 µL 50 mM Tris-HCl pH 8.0 containing 5 ng/µL trypsin (sequencing grade, modified; Promega, Cat# V5111), 2 M urea and 1 mM DTT for 30 min at 25 °C and with shaking at 600 rpm. Afterwards, 100 µL 50 mM Tris-HCl pH 8.0 containing 2 M urea and 5 mM IAA were added and samples were incubated another 30 min at 25 °C under shaking at 600 rpm. Digestion took place overnight at 37 °C under shaking at 600 rpm. The digest was stopped with FA (0.5% final amount). Samples were desalted using double layer C18-stage tips (Empore disk-C18, 47 mm, SUPELCO, Cat# 66883-U) equilibrated with 70 µL MeOH and three times 70 µL aqueous 0.5% FA. Samples were loaded and washed three times with 70 µL aqueous 0.5% FA and eluted three times with 30 µL 80% ACN, 20% water and 0.5% FA. After removal of elution buffer *in vacuo*, samples were taken up in 27 µL aqueous 1% FA, pipetted up and down, vortexed and sonicated for 15 min. Finally, the peptide solution was filtered (0.2 µm pore size).

11.7 Combination of ABPP with chemical cross-linking

11.7.1 Preparation of cells

S. aureus NCTC8325 and *E. coli* K12 were cultured as described under 11.4. One flask was prepared from which the replicates were conducted. Bacteria were grown until an OD_{600nm} of 0.7 and then harvested at 4,000 x g for 5 min at 4 °C, washed once with warm PBS and then resuspended to an OD_{600nm} of 40. 1 mL OD_{600nm} of 40 were then mixed with 10 µL of 5 mM **D3** or **E2** ABPP probe (stocks in DMSO) to result in a final concentration of 50 µM. Bacteria were labeled for 1 h at 25 °C with shaking at 450 rpm. Cells were pelleted and washed two times with cold PBS and then split into two eppies. Bacteria were resuspended in 1 mL warm PBS either containing 2 mM DSSO (20 µL from a 100 mM stock in DMSO) or in 1 mL warm PBS containing 20 µL DMSO as control resulting in an OD_{600nm} of 20. Bacteria were further incubated for 1 h at 37 °C and 200 rpm. Cells were washed twice with 1 mL cold 50 mM Tris-HCl pH 8.0 and once with 1 mL cold PBS. For *E. coli* K12 **D3** labeling, cross-linking and labeling steps were reversed once. Thereby, cross-linking carried out in 1 mL OD_{600nm} of 20, whereas labeling afterwards carried out in 0.5 mL of OD_{600nm} of 40 as described above.

11.7.2 Cell lysis and CuAAC

Bacteria were resuspended in 1 mL cold PBS. *S. aureus* NCTC8325 were lysed in a Precellys homogenizer as described previously.[89] *E. coli* K12 cells were lysed using sonication with cooling intervals on ice in between (3x30 s, 75%). The lysate was centrifuged for 20 min at 21,000 x g and 4 °C to remove cell debris. A BCA assay was performed to adjust for equal protein amounts in a total volume of 800 µL. SDS was adjusted to 0.4% from a 10% stock in PBS previous to CuAAC. Samples were clicked to biotin-azide as described previously[89] and proteins were precipitated adding a four-fold volume of ice-cold acetone (-80 °C) overnight at -20 °C.

11.7.3 Sample preparation for MS

The precipitated proteins were thawed on ice, pelletized at 16,000 x g for 15 min at 4 °C, supernatant was discarded and proteins were washed 2x with 1 mL ice-cold MeOH (-80 °C) using sonication in between to resuspend pellets. Supernatant was discarded and proteins were resuspended in 500 µL 0.4% SDS in PBS previous to enrichment using sonication. Samples were then incubated with 50 µL pre-equilibrated avidin-agarose beads (three times washed with

1 mL PBS with 0.4% SDS for 2 min at 400 x g and r.t.; Sigma, Cat# A9207) under continuous rotation (20 rpm) for 1 h at r.t. Thereafter, beads were pelletized for 2 min at 400 x g and r.t. and washed with 1 mL PBS with 0.4% SDS three times, 6 M urea in water two times, and PBS three times. After each washing step beads were pelletized for 2 min at 400 x g and r.t. and the supernatant was carefully discarded. Beads were resuspended in 200 µL denaturation buffer (7 M urea, 2 M thiourea in 20 mM HEPES pH 7.5) and reduction carried out by addition of 0.4 µL 500 mM DTT prepared fresh in water for 45 min at r.t. and 450 rpm shaking. Afterwards, samples were alkylated with 2 µL of a 500 mM IAA stock prepared fresh in water for 30 min at r.t. and 450 rpm shaking in the dark. IAA was quenched by addition of 2 µL 500 mM DTT and incubation for 30 min at r.t. and shaking at 450 rpm. 2 µL of LysC (0.5 µg/µL, Wako, Cat# 125-05061) were added to each sample and digestion carried out for 2 h at 37 °C and 450 rpm. Thereafter, samples were diluted 4x with 50 mM TEAB (prepared fresh in water from a 1 M stock solution; Sigma, Cat# T7408). 2 µL of trypsin (0.5 µg/µL, Sequencing grade, Promega, Cat# V5111) were added to each sample and digestion carried out for 16 h at 37 °C and 450 rpm. The digest was stopped by addition of FA to a final concentration of 1%. Desalting of samples with C18 columns (Waters, Cat# WAT054960) carried out as described previously.[89] The eluates were evaporated *in vacuo*, samples were re-dissolved in 1% FA in water (pipetted up and down, 15 min sonication), and filtered (0.2 µm pore size) for LC-MS/MS analysis.

11.8 MS analysis of proteomic samples

Samples were analyzed *via* LC-MS/MS using a UltiMate 3000 nano HPLC system (Thermo Fischer Scientific) equipped with an Acclaim C18 PepMap100 75 µm ID x 2 cm trap (Thermo Fisher Scientific, Cat# ES803A) and an Acclaim PepMap RSLC C18 separation column (75 µm ID x 50 cm; Thermo Fisher Scientific, Cat# 164535) coupled to an EASY-source equipped Thermo Fischer LTQ Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). Samples were loaded onto the trap column at a flow rate of 5 µL/min with aqueous 0.1% TFA and then transferred onto the separation column at 0.3 µL/min. Buffers for the nano-chromatography pump were aqueous 0.1% FA (buffer A) and 0.1% FA in ACN (buffer B). Samples were separated using a gradient raising buffer B from 5 to 22% in 112 min, followed by a buffer B increase to 32% within 10 min. Buffer B content was further raised to 90% within the next 10 min and held another 10 min at 90%. Subsequently buffer B was decreased to 5% and held until end of the run (total: 152 min). During sample separation MS full scans were performed at 120,000 m/z resolution in the orbitrap with quadrupole isolation. The MS instrument was operated in a 3 s top speed data dependent mode. The scan range was set from 300 to 1,500 m/z with 60% RF

lens amplitude. The AGC target was set to 2.0e5, the maximum ion injection time was 50 ms, and internal calibration was performed using the lock mass option. Peptides with intensity higher than 5.0 e3 and charge state 2-7 were fragmented with HCD (30%). MS^2 scans were recorded in the ion trap operating in rapid mode. The isolation window was set to 1.6 m/z and the AGC target to 1.0 e4 with maximum injection time of 100 ms. Ions were injected for all available parallelizable time. Dynamic exclusion time was 60 s with 10 ppm low and high mass tolerance.

11.9 Quantification and statistical analysis of proteomic data

MS raw data were searched with MaxQuant software (version 1.6.0.1)[124] and default settings (with the exceptions that the enzyme was trypsin and LFQ and match between runs were activated). All replicates from one condition (*e.g.* DSSO treated and anti-hClpP antibody treated) were defined as one fraction. Searches were performed against Uniprot databases of *H. sapiens* (taxon identifier: 9606, reference reviewed and unreviewed, downloaded on 02.11.2017), *E. coli* K12 (taxon identifier: 83333, reference reviewed and unreviewed proteome, downloaded on 18.12.2017), and *S. aureus* NCTC8325 (taxon identifier: 93061, reference reviewed and unreviewed, downloaded on 06.02.2018). Resulting data were further analyzed using Perseus software version 1.6.0.0.[125] The rows were filtered (only identified by site, potential contaminant, reverse) and $\log_2$ transformed. Replicates (*n*=4) were grouped, filtered for at least three valid values in at least one group and missing values were imputed for total matrix using default settings. A both sided, two-sample Student's *t*-test was performed and derived *p*-values were corrected for multiple testing by the method of Benjamini and Hochberg with a significance level of *p* = 0.05. Volcano plots were generated by plotting $\log_2$ (fold changes) against $-\log_{10}$ (*p*-value).

Mitochondrial proteins were detected searching for GO term *cellular compartment = mitochondrion* with the help of the GO annotation file for *Homo sapiens* submitted on 26.06.2017.[127,195] An analysis of enriched proteins was performed with the functional annotation tool of the DAVID Bioinformatics Resources (version 6.8, 14.02.2018).[129] A modified Fisher Exact test revealed overrepresentation of terms from annotation categories like GO term or KEGG with a *p*-value 0.05.

II – A PROTEOMIC SURVEY ON HUMAN PYRIDOXAL 5'-PHOSPHATE-BINDING PROTEINS

The following part of this thesis is based upon:

<u>Contributions:</u>

A. F. designed, planned and conducted all experiments including kinetic studies, proteomics and metabolomics sample preparation, as well as protein expression and analysis. A. F. performed proteomic and intact protein MS measurements and statistical analysis of the data. V. C. K. contributed expertize to metabolomics sample preparation, measured metabolomics samples, and analyzed metabolomics data. M. P. and A. H. synthesized **PL**-probes. S. A. S. supervised experiments and contributed input throughout the project.

Isabel Buchsbaum performed immunohistochemical imaging under the supervision of Silvia Capello (MPI für Psychatrie). Vadim Korotkov synthesized DSSO cross-linker.

1 Introduction

1.1 The B6 vitamer group

Vitamin B6 is a versatile cofactor referring to the six interconvertible pyridine compounds pyridoxine (**PN**), pyridoxamine (**PM**), pyridoxal (**PL**), and their 5'-phosphorylated versions (**PNP**, **PMP**, and **PLP**). The biologically active form is **PLP**, one of the most ubiquitous cofactors found in nature, catalyzing around 238 different types of biochemical transformations, corresponding to 4% of all classified activities.[196] Although all living organisms rely on vitamin B6, only plants and microorganisms are able to synthesize the cofactor *de novo*. All other organisms, including humans, acquire vitamin B6 from nutrients and interconvert the vitamers to the active form **PLP** by a salvage pathway. Thereby ubiquitously expressed pyridoxal kinase (PLK) phosphorylates **PL**, **PM**, and **PN** to **PLP**, **PMP**, and **PNP**. The latter two are subsequently oxidized to **PLP** by the action of pyridoxine/pyridoxamine 5'-phosphate oxidase (PNPO).[197] Finally, the cofactor is loaded onto **PLP**-dependent enzymes (**PLP**-DEs) *via* formation of a Schiff base (the internal aldimine) between the **PLP** aldehyde group and an ε–amino group of the active site lysine.[198] Hydrolysis of **PLP** by an intracellular specific phosphatase (**PLP**-phosphatase) and subsequent oxidation of **PL** *via* pyridoxal oxidase to pyridoxic acid, which is excreted in the urine, constitute the vitamin B6 degradation pathway (**Figure II-1**).[199–201]

1.2 The diversity of PLP-DEs

PLP-DEs catalyze a diverse set of chemical reactions mainly with amino acid substrates including decarboxylation, transamination, racemization, α, β, γ-substitution or α, β, γ-elimination.[196] Although the scope of **PLP**-catalyzed reactions appears to be manifold, there is a

principle common to all. The cofactor initially forms an external aldimine with the substrate amino acid and subsequently serves as electron sink which stabilizes negative charge development at the Cα-atom in the form of a quinonoid intermediate (**Figure II-2**). One exception display **PLP**-dependent glycogen phophorylases, where **PLP** does not act as electrophilic catalyst, but realizes proton transfer *via* its phosphate group.[202]
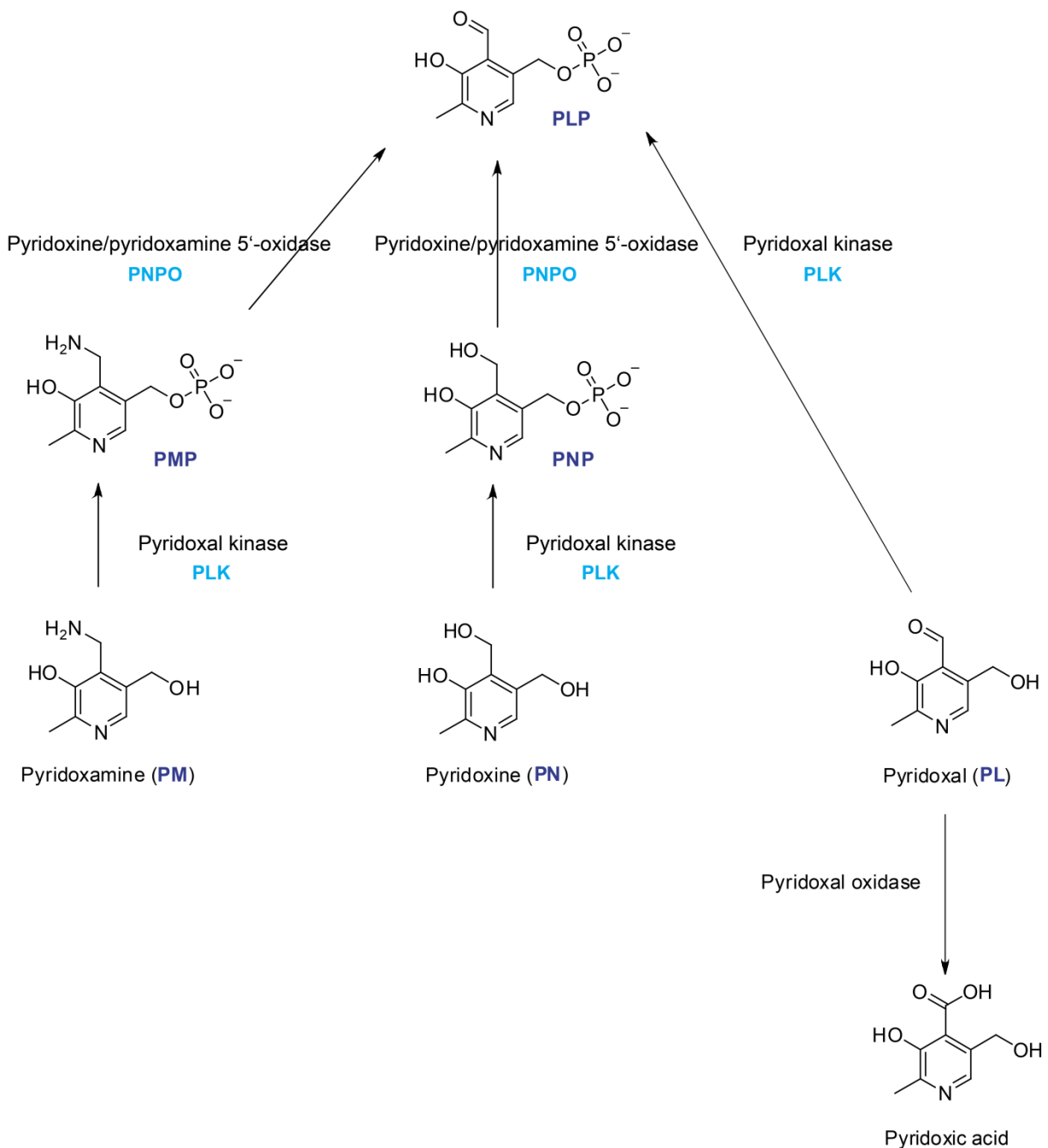
Figure II-1 Interconversion of B6 vitamers in the human body. The membrane permeable precursors pyridoxamine (**PM**), pyridoxine (**PN**), and pyridoxal (**PL**) are phosphorylated by pyridoxal kinase (PLK).

Pyridoxine/pyridoxamine 5'-oxidase (PNPO) subsequently converts **PMP** and **PNP** to the biologically active form **PLP**. Catabolic processes involve oxidation of **PL** to pyridoxic acid.

Although sequence similarity is low among **PLP**-DEs they were categorized into five distinct fold-types guided by structural similarity.[203] It was initially postulated that reaction mechanisms correlate with the structure of **PLP**-DEs,[204] but it has been found that each structural class contains representatives of multiple reaction types.
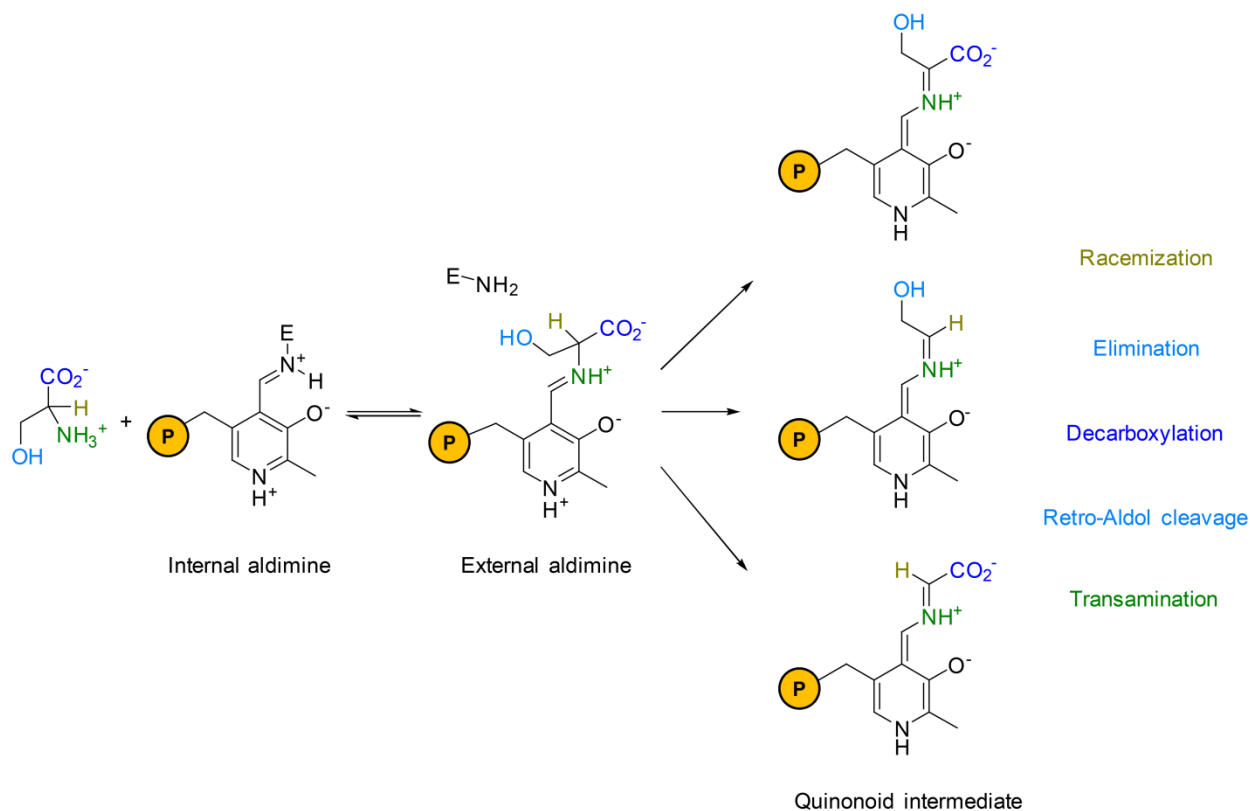
Figure II-2 Chemical reactions catalyzed by the PLP-DEs family. The amino acid substrate (here exemplarily presented as serine) forms in a first common step an external aldimine with the cofactor. Further, the negative charge generated at the Cα-atom is stabilized *via* a quinonoid intermediate, from which **PLP**-DEs realize a diverse set of reactions including racemization and elimination (P = OPO_3^{2-}).

The main part of B6-enzymes belongs to the fold-type I or aspartate aminotransferase family. Members of this fold-type function as obligate homodimers or higher-order oligomers where each subunit displays an active site, mostly localized at the oligomer-interface. Although the members of fold-type II or tryptophan synthase family are structurally similar to those of fold-type I, active site residues are donated entirely from one monomer though the functional form displays a homodimer or higher-order oligomer. Further, members of the fold-type IV or D-amino

acid aminotransferase family are structural similar to fold-type I and II. Fold-types III (alanine racemase family) and V (glycogen phosphorylase family) members are different from other B6-enzymes. As mentioned above, enzymes of glycogen metabolism utilize the cofactor in a different mechanistic way. Fold-type III members consist of an α/β barrel and a second β-strand domain. However, **PLP** binding is similar to fold-types I, II, and IV forming the ubiquitous internal aldimine.[205]

1.3 PLP-DEs in human diseases and oncogenesis

Due to their immense repertoire of chemical reactions, **PLP**-DEs play a crucial role in cellular metabolism ranging from amino acid biosynthesis and turnover to synthesis of neurotransmitters or tetrapyrroles.[206] Several inborn errors due to deficit in **PLP**-DEs are known, including homocystinuria, triggered by defects in cystathionine-β-synthase (CBS), where a high level of homocysteine causes endothelial injury and increases the risk of artery and vein disease. Another example is X-linked sideroblastic anaemia were a deficient in 5-aminolevulinate synthase (ALAS) causes abnormal accumulation of iron in red blood cells.[207] However, also an overactivation of **PLP**-DEs can lead to disease-causing phenotypes. For example, γ-aminobutyric acid (GABA) displays the major inhibitory neurotransmitter in the central nervous system and is converted to succinic semialdehyde *via* the action of GABA-aminotransferase (GABA-AT).[208] Low GABA levels are associated with epilepsy,[209] Parkinson's disease,[210] Huntington's chorea,[211] and Alzheimer's disease.[212] In addition, several human **PLP**-DEs like serine hydroxymethyltransferase (SHMT)[213] or ornithine aminotransferase (OAT)[214] display attractive targets for drug development due to their role in cell proliferation and tumor development. SHMT is a highly conserved enzyme[213] that catalyzes the conversion of serine and tetrahydrofolate (THF) into glycine and 5,10-methylene-THF,[215] which donate carbon for the *de novo* synthesis of purine and pyrimidine nucleotide bases for DNA replication.[216] In order to maintain a high proliferation rate, several human tumors display elevated SHMT activity.[216,217] Human OAT on the other hand was shown to play a role in regulating mitotic cycles in rapidly growing cells.[218] Further, the protein was demonstrated to be overexpressed in hepatocellular carcinoma.[219]

As dysregulated expression levels of **PLP**-DEs correlate with several diseases in human, a large number of B6-dependent enzyme inhibitors have been developed. For example, a main class display mechanism-based inhibitors which are designed to poison specific **PLP**-DEs. For instance, (1*R*,3*S*,4*S*)-3-amino-4-fluorocyclopentane-1-carboxylic acid (FCP) inhibits GABA-AT

via formation of a covalent adduct with **PLP** (**Figure II-3A**).[220] Further, **PL** scaffolds covalently attached to the amino acid substrate or an analogue are applied as transition-state mimics, for example to inhibit human histidine decarboxylase (**Figure II-3B**).[221] However, as all **PLP**-DEs share the same cofactor, the design of selective inhibitors is challenging. For instance, FCP was also shown to inhibit ornithine aminotransferase (OAT),[220] preventing its application as pharmaceutical. Therefore, further optimization concerning inhibitor selectivity, as well as the development of a simple read-out to identify *in situ* target proteins is crucial.
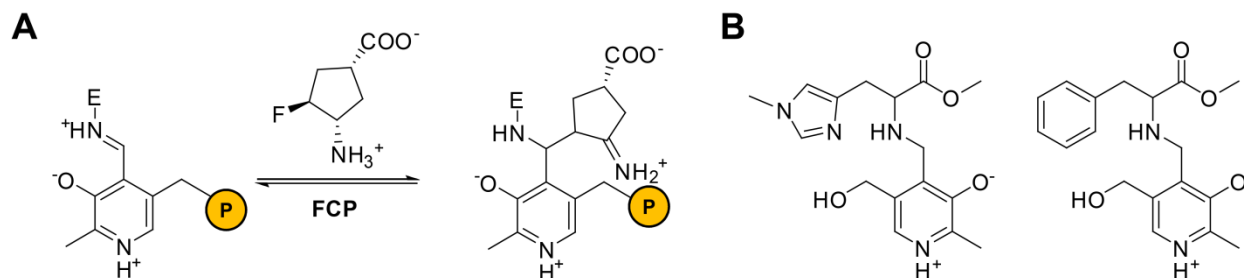
Figure II-3 Inhibitors of PLP-DEs. (**A**) The mechanism-based inhibitor (1*R*,3*S*,4*S*)-3-amino-4-fluorocyclopentane-1-carboxylic acid (FCP) of *γ*-aminobutyric acid-aminotransferase (GABA-AT) forms a covalent adduct with **PLP** (P = OPO_3^{2-}). (**B**) Transition-state analogues for inhibition of human histidine decarboxylase.

1.4 A method for chemical proteomic profiling

A key technology to determine the functional state of proteins in cells and tissues displays ABPP, which relies on the design of covalent probes directed to the active site of single proteins or protein families. ABPP probes utilize a wide range of scaffolds, for instance mechanism-based inhibitors or natural products. The goal is to gain information about structure and mechanistic properties of certain cellular proteins.[111,112] The fundamental building block of an ABPP probe is a small molecule that covalently attaches to a single protein or a set of proteins. A major challenge at this point is to limit cross-reactivity with other cellular proteins termed as *off*-targets. A required feature of an ABPP probe is that target binding occurs in a covalent manner. This can be either achieved by the small-molecule itself, or with the additional introduction of photoreactive groups like diazirines that label proximal residues in target proteins following UV exposure. The second required part of an ABPP probe, which is usually attached to the small-molecule *via* a linker region, is a chemical moiety inert towards cellular functionalities to enable the attachment of a reporter-tag for target identification.[171,222] Commonly used are azides and alkynes, which allow for CuAAC.[52,53,223] Thereby, for instance fluorophores enabling

gel-based target readout or biotin enabling enrichment of target proteins for subsequent identification by LC-MS/MS can be covalently attached (**Figure II-4**).
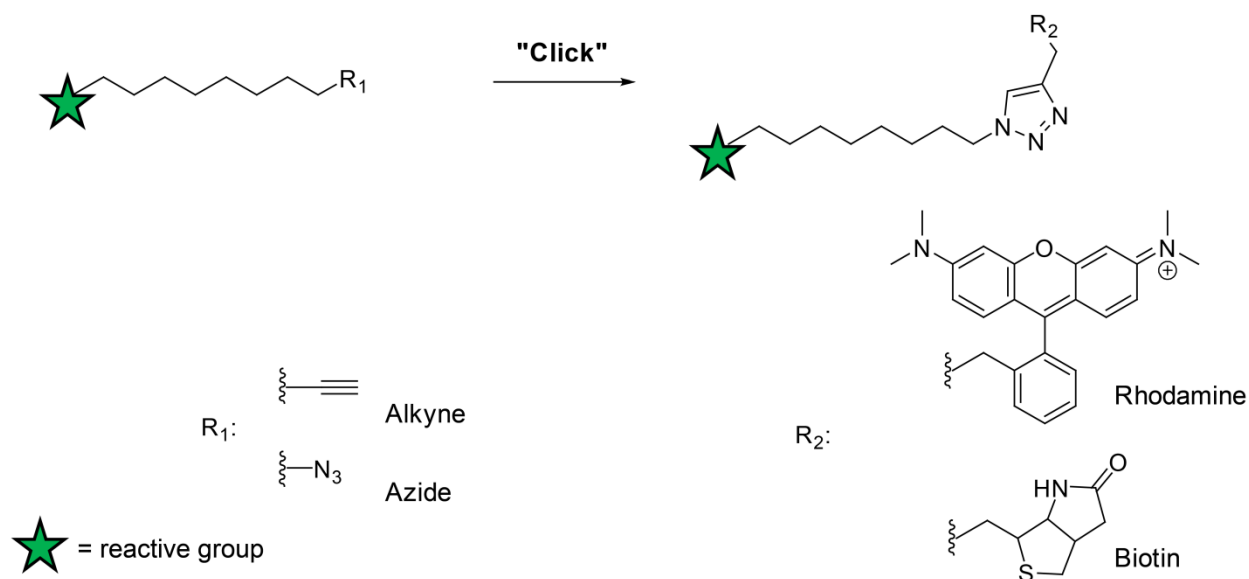
Figure II-4 Schematic representation of an ABPP probe. A reactive group targeting cellular proteins is connected to an alkyne or azide moiety *via* a flexible linker region. Bioorthogonal ligation using CuAAC allows the attachment of rhodamine or biotin for subsequent target visualization or identification, respectively.

Original versions of ABPP utilize a gel-based readout for the detection of cellular target proteins. Probe-treated proteomes are separated by SDS-PAGE and are subsequently analyzed by in-gel fluorescence scanning compared to a non-treated control.[224,225] As the principal limitation of a gel-based analysis is resolution, LC-MS/MS strategies were developed for target identification and quantification. Therefore, biotinylated proteomes are incubated with (strept)avidin beads to enrich for probe-bound proteins, that are further subjected to proteomic workflows including reduction, alkylation and digestion of proteins for shotgun proteomics (**Figure II-5**).[222] Ongoing optimization of sample preparation, instrumentation, and downstream data analysis allows for fast and confident identification of cellular target proteins. Several methods have been developed to additionally quantify the target proteins compared to a background control either by label-based workflows including SILAC,[226] or label-free strategies.[123]
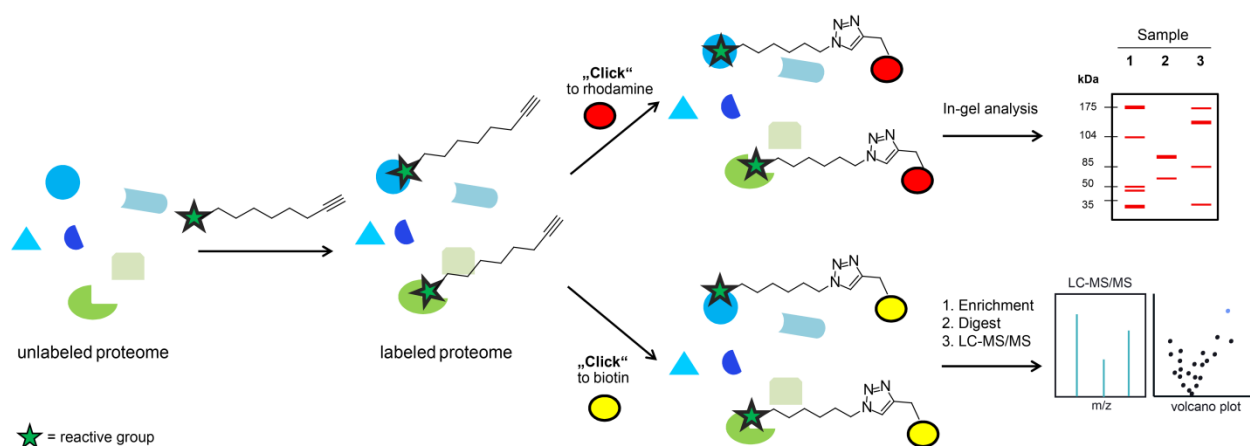
Figure II-5 Schematic overview of a standard ABPP workflow. The proteome is incubated ("labeled") with the alkyne functionalized probe. Thereby target proteins are covalently attached to the probe *via* specific reaction with the reactive moiety (*e.g.* parts of natural products). Subsequent CuAAC to attach rhodamine or biotin enables gel-based fluorescence readout or LC-MS/MS-driven target identification after (strept)avidin-enrichment and protein digestion.

Multiple types of biological applications for the ABPP technology have emerged over the years. The original application is for target discovery of the functionalized small molecules and the readout of target activity in biological systems. This can also be carried out in a comparative way for example by analyzing two or more proteomes which may stem from a healthy and non-healthy individual. For instance, altered enzyme activities could give a hint by which mechanisms the observed phenotypes are caused within the cell. A second application is competitive ABPP where cellular targets of certain inhibitors can be identified by their ability to block probe labeling or displace the probe from the enzyme active site. The advantages over other screening methods are that enzymes are tested under native cellular conditions and that competitive ABPP tests inhibitors against many proteins in parallel, which therefore additionally assigns selectivity and potency.[171,222]

1.5 Strategy and objectives

Although single **PLP**-DEs have been studied extensively, a global proteomic strategy for screening of the complete **PLP**-binding proteome (**PLP**-ome) has been lacking. Inspired by the ABPP approach, vitamin B1 (thiamine), B2 (riboflavin), and B7 (biotin)-derived probes have been developed to study live cell nutrient transport and intracellular protein interactions in *Chloroflexus aurantiacus* (**Figure II-6**).[227] Following incubation with probe and UV irradiation, labeled proteins were clicked to biotin-azide and enriched *via* incubation with avidin beads.

Part II-1 Introduction

Subsequent LFQ after LC-MS/MS analysis identified vitamin transporters and vitamin-dependent enzymes, revealing interesting insights into nutrition transport and cellular cofactor utilization.
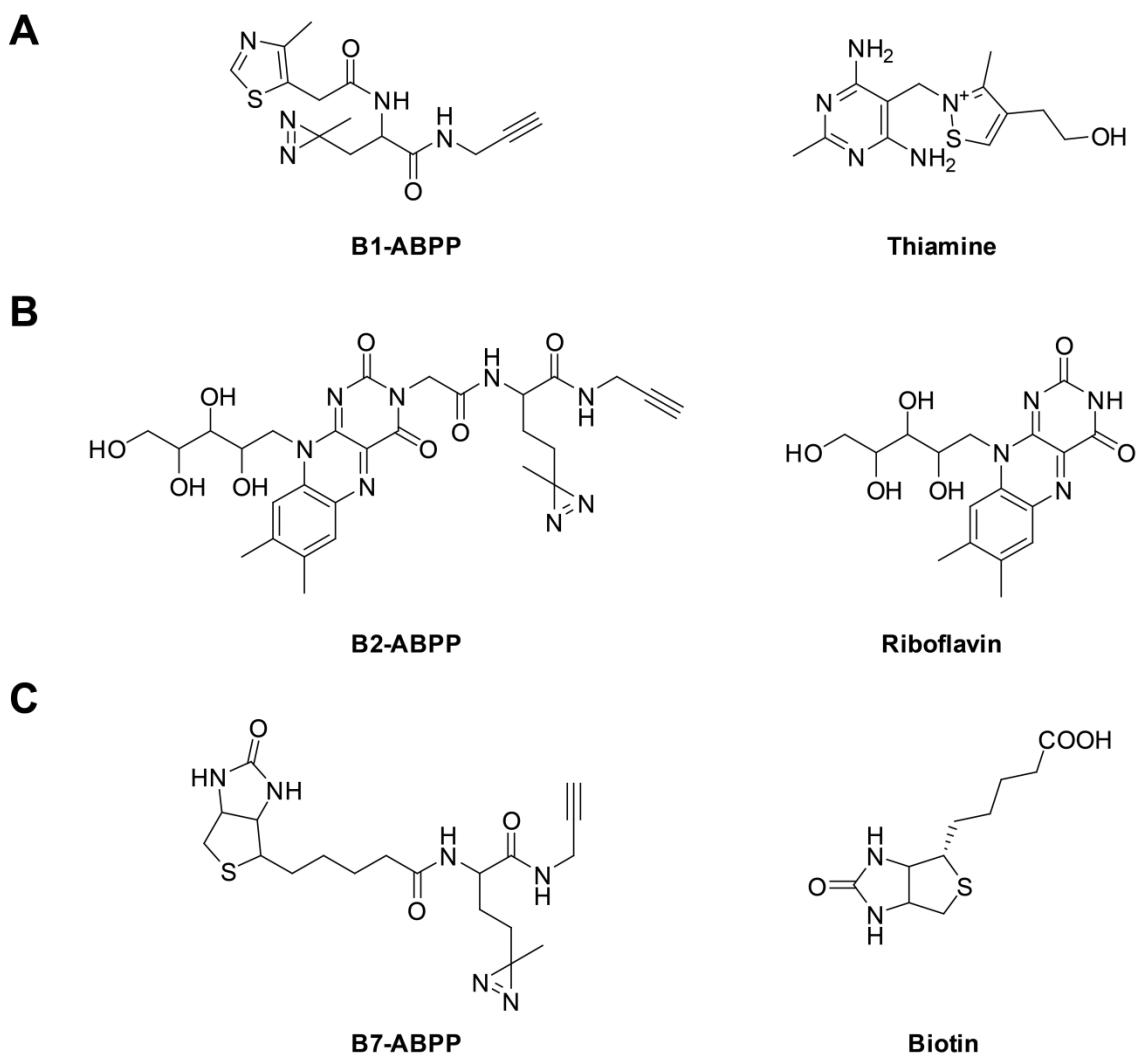
Figure II-6 Vitamin-derived ABPP probes and corresponding native structures. (**A**) B1-ABPP and thiamine, (**B**) B2-ABPP and riboflavin, (**C**) B7-ABPP and biotin.[227]

Although single **PLP**-DEs have been studied intensively, a method for the global and quantitative profiling of this enzyme class is missing. To fill the gap concerning vitamin B6-derived probes and to study the family of **PLP**-DEs with their crucial cellular roles in live cells, our group recently introduced a cofactor mimic library composed of alkyne- and azide-functionalized **PL** probes (**Figure II-7**) to study the *S. aureus* **PLP**-ome.[198] As **PLP** forms an internal aldimine with B6-enzymes, no UV-activatable group like a diazirine for covalent connection to target proteins is necessary. We demonstrated that the **PL**-derived probes are phosphorylated by *S. aureus* pyridoxal kinase (SaPLK) at levels comparable to the natural

vitamer *in vitro*. Further, we have shown that probes are entering the cell and are subsequently activated by the intracellular kinase. Upon *in situ* labeling of bacteria with the probes, we were able to cover 73% of all annotated **PLP**-DEs and assign functions to previously uncharacterized proteins. Further, we demonstrated that different **PLP**-DEs can tolerate modifications at the vitamin B6 scaffold to varying extents and identified cellular *off*-targets of the B6-binding antibiotic D-cycloserine (DCS).[228] Taken together, we proved that our vitamin B6-probe toolset for profiling the **PLP**-ome is compatible with biological systems and addresses shortcomings in the monitoring of the global set of vitamin B6-binding enzymes.
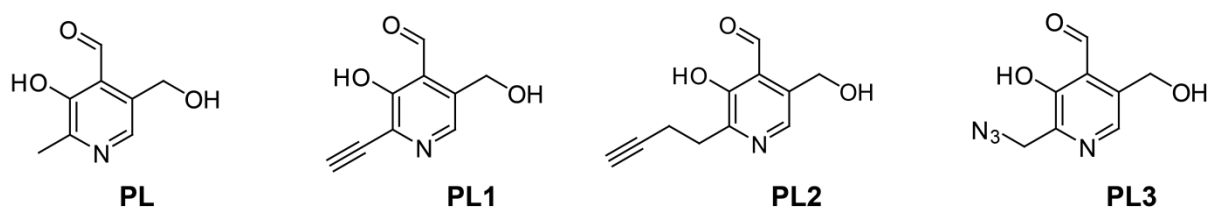
Figure II-7 Structures of natural PL and PL-derived probes. Probes were either functionalized with an alkyne (**PL1** and **PL2**) or an azide moiety (**PL3**) at the 2'-position.

As human **PLP**-DEs have a manifold of physiological and pathophysiological roles within the cell, a method for their global profiling is of overall interest. Therefore, this part of the thesis describes the application of the developed functionalized B6 mimics for higher eukaryotic systems like human cells. After *in situ* labeling of cellular proteins with **PL**-probes, the internal aldimine mediates an irreversible link between cofactor derivative and protein upon reduction with sodium borohydride (NaBH$_4$). Subsequent CuAAC to attach rhodamine or biotin enables gel-based fluorescence readout or enrichment of target proteins for subsequent LC-MS/MS analysis compared to a **PL**-labeled control (**Figure II-8**). Our goals were a significant enrichment of human **PLP**-DEs, as well as the examination of poorly characterized examples. Further, we aimed for *in situ* target screening of B6-binding inhibitors to elucidate their cellular selectivity. Finally, we planned to compare different human cell lines in their labeling pattern to examine if any changes depending on the type of tissue can be observed.
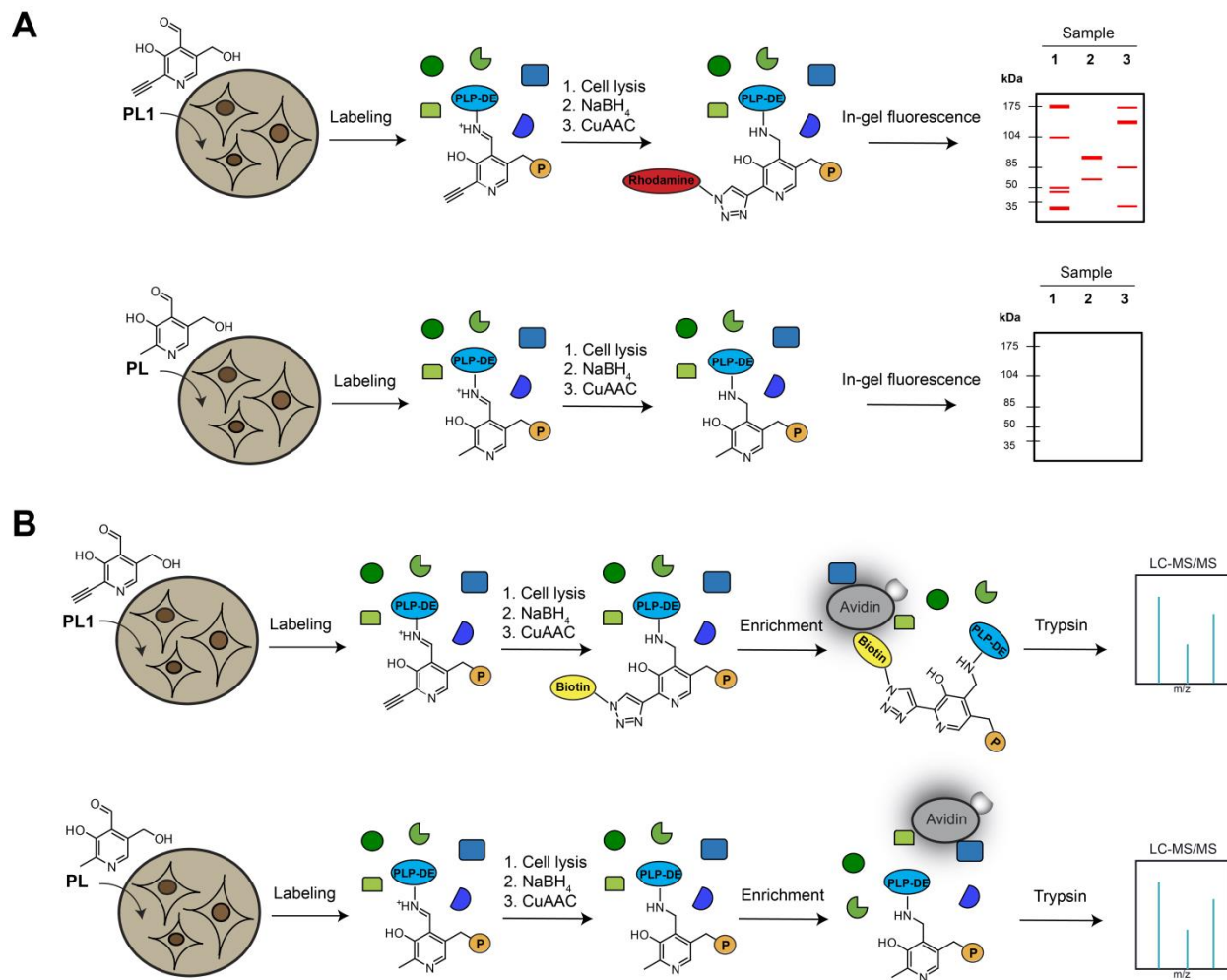
Figure II-8 Overview of the chemical proteomic strategy for profiling the repertoire of PLP-DEs. Cells were incubated in media containing either **PL1** or **PL** as a control to allow for enrichment of **PL1P** or **PLP** in **PLP**-DEs. Following cell lysis, internal aldimine was reduced with sodium borohydride (NaBH$_4$) and proteins were subjected to CuAAC to attach rhodamine (**A**) or biotin (**B**). Proteins were either analyzed by in-gel fluorescence readout (**A**) or by LC-MS/MS after avidin enrichment and protein digestion with trypsin (P = OPO$_3^{2-}$) (**B**).

2 Bioactivation of **PL**-probes by human PLK

As **PL** phosphorylation is the essential step to generate the biologically active form of vitamin B6, we first monitored turnover of **PL**-probes by the human PLK (hPLK). An efficient turnover is crucial for successful capture of **PLP**-DEs inside the cell.

2.1 Bioactivation of PL-probes by wild-type hPLK

PL-probes were synthesized as described previously[198] with slight modifications in the synthetic route. We purified hPLK as a dimer (**Figure II-9A**) in line with literature observations.[229] Turnover of **PL** by hPLK was recorded by tracing the absorbance increase at 392 nm, which corresponds to a new absorbance maxima generated upon phosphorylation (**Figure II-9B**). Similar to the natural occurring vitamin, **PL**-probes show changes in UV/Vis-spectra caused by phosphorylation. The absorbance maxima of phosphorylated **PL**-probes are redshifted with **PL1P** (λ_{max} = 416 nm) exhibiting the most pronounced change compared to **PLP** (λ_{max} = 392 nm). With a K_m-value for **PL** of around 2.16 µM our results are in the range of previous reported data (**Figure II-9C** and **G**).[230] Although the K_m-value for **PL1** increased about 10-fold, catalytic efficiency was only 3-fold reduced compared to the natural vitamer (**Figure II-9D** and **G**). However, the longer 2'-substitution of **PL2** and **PL3** prevented efficient phosphorylation by hPLK (**Figure II-9E, F** and **G**). This was in contrast to our previous results using SaPLK, which converted **PL2** and **PL3** to phosphorylated products at levels comparable to **PL**.[198] The altered affinities of hPLK and SaPLK towards **PL2** and **PL3** could be caused by differences in their catalytic mechanisms. Although both enzymes act on the same substrate, **PL** forms a

hemithioacetal with SaPLK cysteine located in a flexible lid and remains covalently linked during catalysis cycle.[231] In contrast, the cysteine is replaced by an aspartate in hPLK, maintaining its interaction with **PL** solely through hydrogen bonds and hydrophobic interactions.[232,233] Different ways of substrate interaction could allow for a more flexible positioning of longer 2'-substitutions of **PL2** and **PL3** in SaPLK and therefore for a better phosphorylation efficiency.

Substrate	K_m [µM]	v_{max} [µM/min]	k_{cat} [min^{-1}]	k_{eff} [min^{-1} • µM^{-1}]
PL	2.16 ± 0.38	0.79 ± 0.02	3.16 ± 0.09	1.46 ± 0.26
PL1	28.0 ± 5.31	2.87 ± 0.13	11.5 ± 0.51	0.41 ± 0.08
PL2	582 ± 43.8	4.32 ± 0.14	17.3 ± 0.55	0.03 ± 0.00
PL3	516 ± 57.8	8.05 ± 0.36	32.2 ± 1.43	0.06 ± 0.01
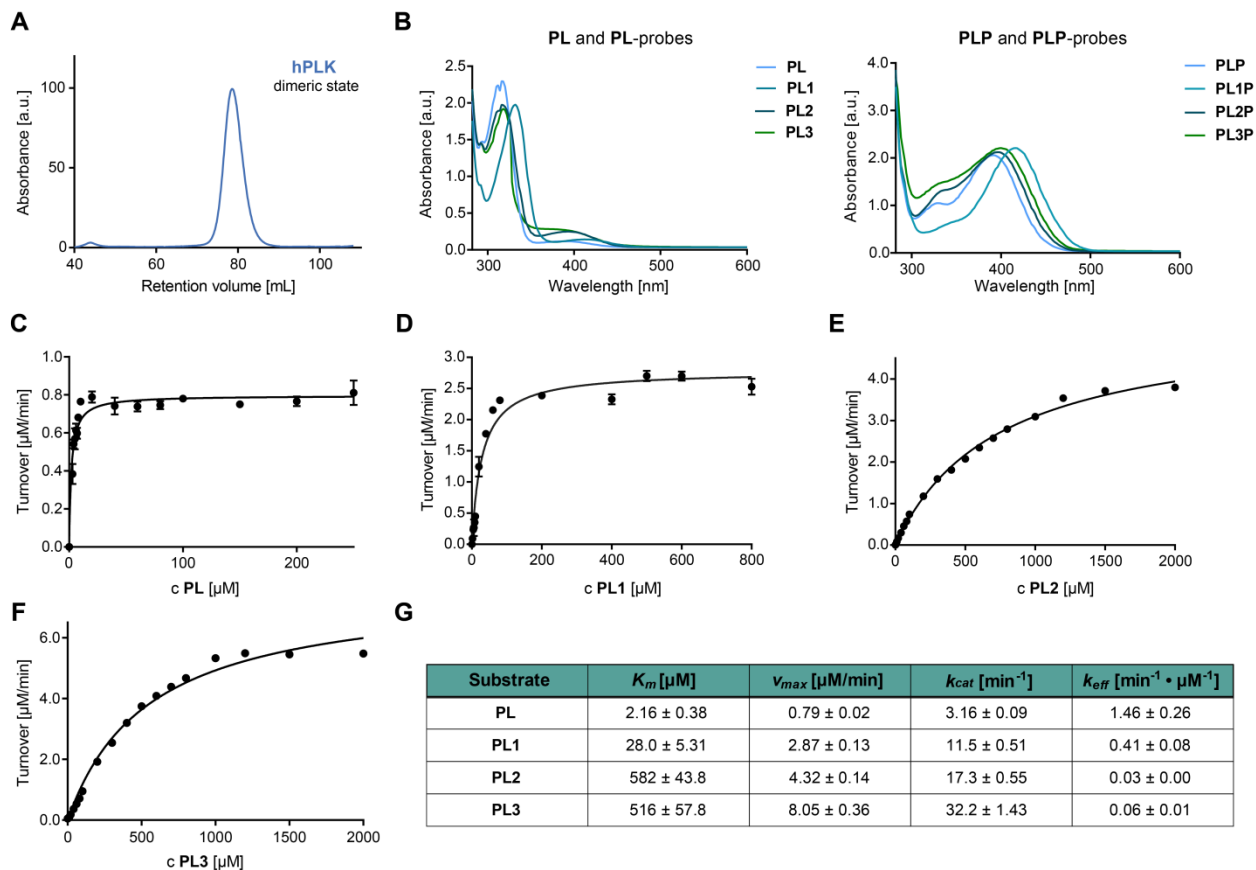
Figure II-9 Bioactivation of PL and PL-probes by hPLK. (**A**) SEC of hPLK revealed a dimeric state upon comparison with standards. (**B**) UV/Vis-spectra of **PL, PL**-probes, **PLP,** and **PLP**-probes reveal additional absorbance maxima for phosphorylated equivalents in the visible region. Michaelis-Menten curves of **PL** (**C**), **PL1** (**D**), **PL2** (**E**) and **PL3** (**F**). Turnover of substrates was monitored by tracing the absorbance increase at 392 nm (**PLP**), 416 nm (**PL1P**), 396 nm (**PL2P**), and 400 nm (**PL3P**) due to phosphorylation (*n*=3 biological replicates; data presented as mean ± SEM). (**G**) Kinetic parameters (*n*=3 biological replicates; data presented as mean ± SEM) for **PL** and **PL**-probes.

2.2 Fine-tuning of hPLK active site geometries inspired by the Bump-and-Hole tactic

As no efficient turnover of longer 2'-substituted probes **PL2** and **PL3** was observed for wild-type hPLK, we inspected the active site of hPLK (PDB: 3KEU)[233] for residues surrounding the 2'-methyl group of the cofactor (**Figure II-10A**). Inspired by the Bump-and-Hole strategy for kinase engineering,[234] where additional space in active site pockets for selective ligand targeting is generated upon introduction of point mutants, we exchanged Val41, Phe43, Val14, and Thr47 to Ala. In addition, we created another point mutant F43G. An engineered kinase with higher affinity towards longer 2'-substitued probes could for instance be transfected into human cells to increase phosphorylation efficiency of **PL2** and **PL3**, and therefore enhance subsequent labeling of **PLP**-DEs with these cofactor mimics.

hPLK variant	K_m for PL [µM]	K_m for PL2 [µM]
wild-type	2.160 ± 0.38	582.0 ± 43.8
V14A	1045 ± 360	1598 ± 299
V41A	154.7 ± 40.7	N/A
F43A	370.2 ± 83.7	705.9 ± 84.6
F43G	2023 ± 158	826.7 ± 241
T47A	103.5 ± 28.3	N/A
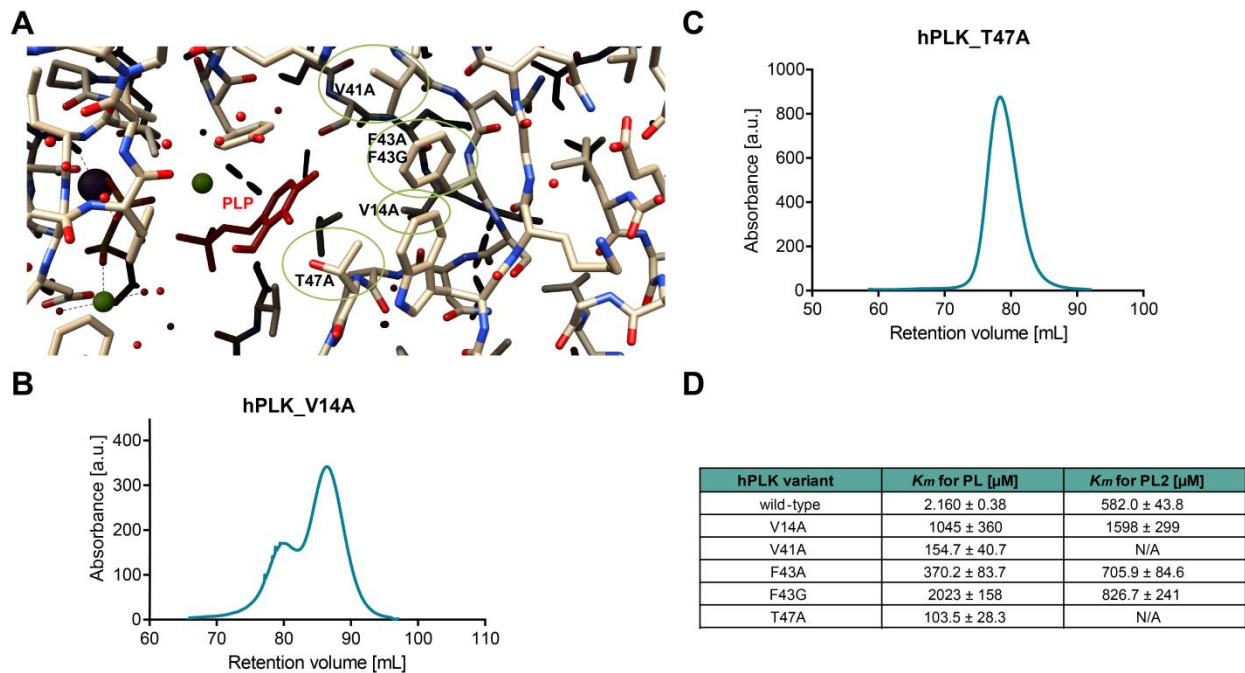
Figure II-10 Engineering of hPLK for more efficient PL-probe turnover. (**A**) Active site pocket of hPLK with residues highlighted that are exchanged in created point mutants (PDB: 3KEU).[233] (**B**) SEC profiles of V14A (**B**) and T47A (**C**) mutants. (**D**) Kinetic values for **PL** and **PL2** for wild-type kinase and point mutants (*n*=3 biological replicates; data presented as mean ± SEM). For V41A and T47A no **PL2** turnover was detected (N/A).

We expressed and purified all hPLK point mutants following the protocol for wild-type kinase. V14A was mainly present as monomer (**Figure II-10B**), as V14 is involved in the dimerization interface.[235] However, all other mutants were purified as dimer like the wild-type hPLK (**Figure II-**

10C). In addition, protein yields of the mutants were comparable to wild-type kinase, except for a significantly lower protein expression of V14A. We proceeded with kinetic assays to monitor turnover of **PL** and **PL2** for all created point mutants. K_m-values for **PL** increased manifold for all mutants, indicating that their active site pocket geometries were not intact anymore. In addition, affinities for **PL2** decreased markedly or no turnover at all was detected in the case of V41A and T47A, suggesting that the created point mutants are not suitable for optimization of *in situ* probe turnover. Due to inefficient turnover of natural **PL** and **PL2**, we did not proceed with investigation of **PL3** phosphorylation.

2.3 Conclusion

Taken together, we observed altered bioactivation efficiency of **PL**-probes by the human ortholog of PLK compared to SaPLK, which may stem from the different evolved catalytic mechanisms. Although we aimed at engineering the active site pocket of hPLK to receive a higher affinity for longer 2'-substitued probes by introduction of point mutants, we were not able to increase the turnover of **PL2**. However, all examined probes were phosphorylated by wild-type hPLK, with **PL1** displaying the most efficient turnover.

3 *En route* to an optimized labeling strategy for **PL**-probes

Having examined their bioactivation, we proceeded with investigating the potential of **PL**-probes to label **PLP**-DEs in human cells. As **PL**-probes turned out to be phosphorylated with lower efficiency than the natural vitamer, we decided to perform all labeling experiments in in-house prepared chemically defined media (CDM) lacking any natural vitamin B6 to increase the uptake, turnover, and loading of cofactor mimics onto **PLP**-DEs. Further, we cultured cells in CDM complemented with dialyzed fetal calf serum (FCS) as this serum-type contains lower amounts of natural vitamins such as vitamin B5 compared to undialyzed standard FCS.[236]

3.1 Labeling with PL1

As **PL1** is most efficiently phosphorylated by hPLK, we initiated our screening with this cofactor analogue. First, we performed *in situ* analytical labeling, incubating live cells with different concentrations of the probe. After cell lysis, $NaBH_4$ reduction and subsequent CuAAC to attach rhodamine-azide, samples were analyzed by SDS-PAGE and fluorescence-based readout. We detected only weak labeling by **PL1** after prolonged exposure times (~30 s; **Figure II-11A**). A weak labeling could be the result of the intrinsic electrophilicity of alkynes, which are therefore prone to be attacked by primary thiols.[237] As the alkyne-moiety in the case of **PL1** is positioned in direct conjugation to the aromatic ring, this intrinsic electrophilicity is increased (**Figure II-11C**, left). To overcome nucleophilic attacks by cellular cysteines, we incubated proteins immediately before CuAAC with 20 or 40 mM IAA[238] as additional activation of **PL1** by Cu(I) was postulated to increase side reactions in the *S. aureus* proteomic background (**Figure II-11A** and **C**, right).

Contrarily to observations for the bacterial system, we did not detect an increased *in situ* labeling signal in human cells upon pre-incubation with IAA.
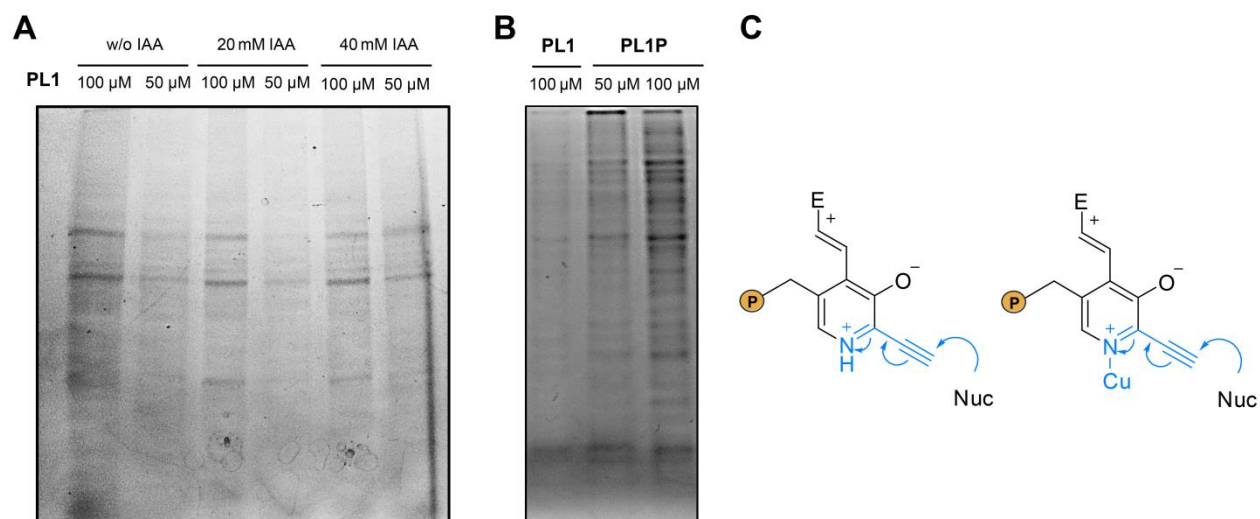
Figure II-11 Optimization of labeling with PL1[P] in HepG2 by a gel-based readout. (**A**) Cells were incubated with different concentrations of **PL1** for 2 h previous to cell lysis, reduction and CuAAC. To prevent the addition of cellular nucleophiles to Cu(I)-activated **PL1,** we incubated samples with 20 and 40 mM IAA previous to click-chemistry (exposure time ~30 s). (**B**) To rescue labeling, either intact cells were treated with **PL1** for 2 h and lysed in the presence of 1 mM maleimide (left) or cell lysate reacted with 1 mM maleimide was labeled with **PL1P** for 1 h (center and right) (exposure time ~5 s). (**C**) Schematic representation of addition of nucleophiles to **PL1** either in the absence (left) or presence of Cu(I) (right). For coomassie-stained gels compare **Figure S7** (P = OPO$_3^{2-}$).

To further evaluate if **PL1** labeling can be improved, we lysed the cells after incubation with the probe in the presence of NaBH$_4$ and 1 mM maleimide to neutralize reactive cellular cysteines from the beginning. Neither with nor without supplementary incubation with IAA previous to CuAAC the labeling efficiency could be increased (**Figure II-11B**, left). Also evaluation of different concentrations of click-reagents (CuSO$_4$, ligand), as well as click-ligands [2-(4-((bis((1-(*tert*-butyl)-1H-1,2,3-triazol-4-yl)methyl)amino)methyl)-1H-1,2,3-triazol-1-yl)acetic acid (BTTAA) and tris((1-hydroxy-propyl-1H-1,2,3-triazol-4-yl)methyl)amine (THPTA)], did not result in a stronger labeling pattern (data not shown). To proof that nulceophilic addition to **PL1** already occur within the cell, we incubated cellular lysate neutralized with 1 mM maleimide with **PL1P**. The pre-phosphorylated version of the probe was obtained by *in vitro* phosphorylation using SaPLK.[198] Indeed, labeling was rescued upon incubation with **PL1P**, suggesting that **PL1** is already attacked insight the cell independent of Cu(I)-activation (**Figures II-11B**, center and right and **C,** left; exposure time ~5 s).

We proceeded with MS-based evaluation of **PL1** cellular targets. *In situ* labeling with 100 µM probe for 2 h resulted in few enriched proteins, including only two **PLP**-DEs (GO term 0030170: **PLP**-binding[127]) named pyridoxal 5'-phosphate-binding protein (PLPBP) and sphingosine 1-phosphate lyase (SGPL1) in line with weak in-gel fluorescence signal (**Figure II-12A**). Although we did not detect increased analytical labeling with **PL1** after IAA incubation previous to CuAAC, we nevertheless evaluated targets *via* MS. Albeit the overall number of significantly enriched proteins was increased, only one **PLP**-DE (PLPBP) was detected, confirming that for the human system IAA incubation previous to click-chemistry does not improve enrichment of vitamin B6-dependent enzymes (**Figure II-12B**). As gel-based optimizations revealed a more intense labeling pattern by incubation of cellular lysate with maleimide previous to labeling with **PL1P**, we proceeded with MS-based evaluation of target proteins under these conditions. We lysed cells in the presence of protease and phosphatase inhibitors for 30 min followed by a 20 min incubation step with 1 mM maleimide and labeling with 1 µM **PL1P** for 1 h. After lysate labeling, the proteome was reduced with NaBH$_4$ and clicked to biotin-azide for subsequent enrichment and LC-MS/MS analysis. Though we observed several significantly enriched proteins, only PLPBP was vitamin B6-dependent, demonstrating that incubation of the proteome with **PL1P** under lysate conditions does not lead to enhanced coverage of **PLP**-DEs (**Figure II-12C**). It could be possible that **PLP**-loading occurs in parallel with protein folding, both processes that may be hampered after cell disruption.
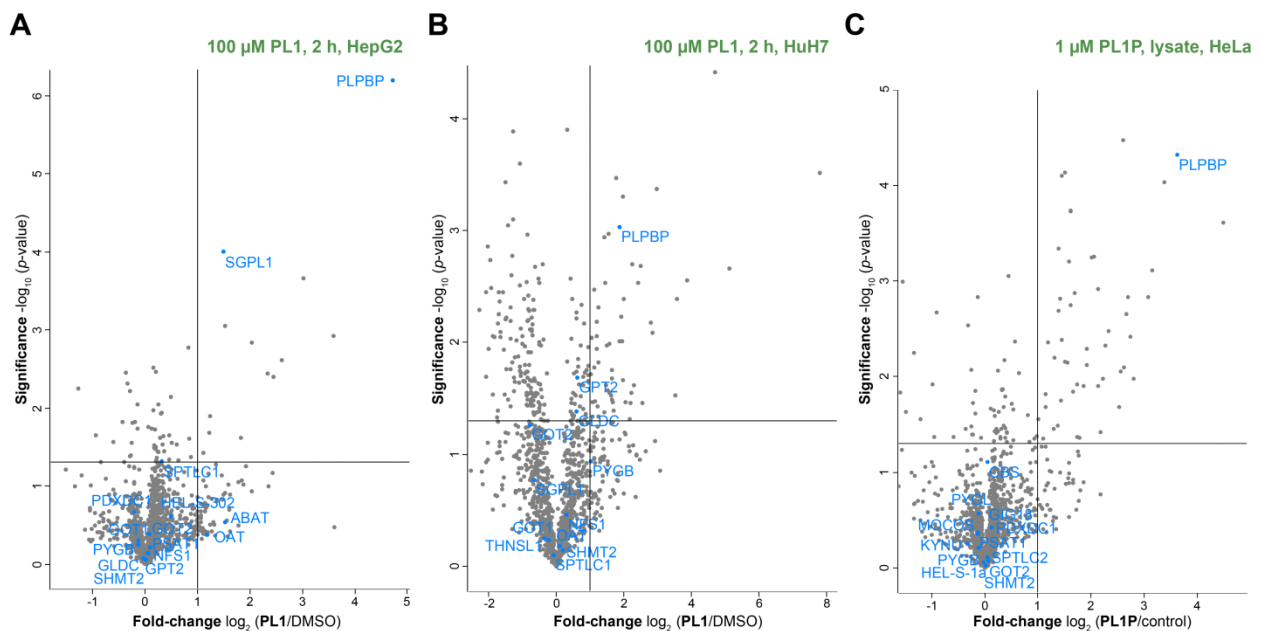
Figure II-12 *In situ* labeling with PL1 and lysate labeling with PL1P evaluated by LC-MS/MS. *In situ* labeling with 100 µM **PL1** for 2 h in HepG2 (**A**) and in HuH7 (**B**), where samples were additionally treated

with 20 mM IAA for 20 min before click-chemistry. (**C**) Labeling of maleimide-incubated HeLa lysate with 1 µM **PL1P** for 1 h. Volcano plots representing *t*-test results of **PL1** or **PL1P** labeling, respectively, compared to the **PL** or **PLP** control (*n*=3 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). **PLP**-DEs (gene names) are colored blue.

As previous experiments in bacteria demonstrated a superior labeling upon supplementation of growth media with **PL**-probes,[198] we cultivated adenovirus-transfected embryonic kidney cells (HEK293) in the presence of **PL1**. We maintained cells for seven passages (35 days) on **PL1**-containing media lacking natural B6-vitamers to allow for complete turnover of the proteome[226] and therefore accumulation of **PL1P** in **PLP**-DEs including a **PL** control. As a balance between low concentrations (5 µM) of natural **PL** in standard cell culture media and increased K_m-value of the probe, we maintained cells in the presence of 10 µM **PL1**. Notably, the doubling rate and morphology of HEK293 was not affected by incubation with **PL1** compared to the **PL** control, indicating that the probe supports essential functions of natural vitamin B6. To evaluate if natural vitamin B6 levels are reduced upon incubation with probe in the chemically defined media lacking **PL**, we performed a quantitative targeted metabolomics analysis. Satisfyingly, MS²-signal intensities of endogenous **PL** and **PLP** levels declined markedly after three passages in **PL**-free media with no natural occurring vitamers left after seven passages (time point of labeling, **Figure II-13**).
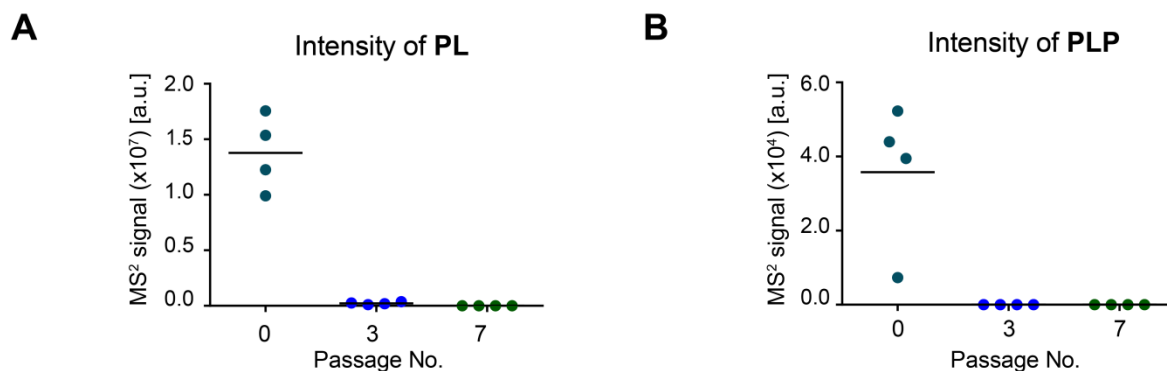
Figure II-13 Targeted metabolomics analysis of PL (A) and PLP (B) in HEK293 cells. MS² signal intensities of vitamers were recorded for standard cells (passage No. zero), and cells that were passaged in chemically defined media with **PL1** but lacking **PL** (*n*=4 biological replicates).[239]

We detected 15 significantly enriched **PLP**-DEs upon **PL1** metabolic labeling compared to the **PL** control, indicating that cofactor loading occurs in parallel with folding of newly synthesized enzymes (**Figure II-14**). Among them were several enzymes related to cancer proliferation and tumor-development (e.g. SHMT1[213] or cytosolic aspartate aminotransferase, GOT1[240]). Further,

we detected the two enzymes serine palmitoyltransferase 2 (*SPTLC2*) and sphingosine-1-phosphate lyase 1 (*SGPL1*) which are involved in *de novo* biosynthesis of sphingolipids in mammalian cells. Sphingolipid metabolites play important roles in differentiation or cell proliferation,[241] as well as pathological processes like Parkinson's disease[242] or atherosclerosis.[243] Enzymes like SPTLC or SGPL are therefore subject of current research including the development of selective inhibitors.[228] Moreover, we hit the **PLP**-DEs cystathionine γ-lyase (*CTH*) and cystathionine β-synthase (*CBS*) which are two out of three enzymes involved in hydrogen sulfide (H_2S) production in humans. H_2S functions for instance as a pro-inflammatory mediator with elevated levels observed in arthritic patients, and as a regulator for heart and blood vessel function.[244,245] PLPBP displayed the highest fold-change comparing LFQ intensities of **PL1** labeling to the **PL** control. Interestingly, PLPBP is a poorly characterized **PLP**-DE with an up to date not defined cellular function. In addition, the *S. aureus* ortholog (uniprot code: A0A0H2XHH8) was significantly enriched with all **PL**-probes in previous experiments.[198]
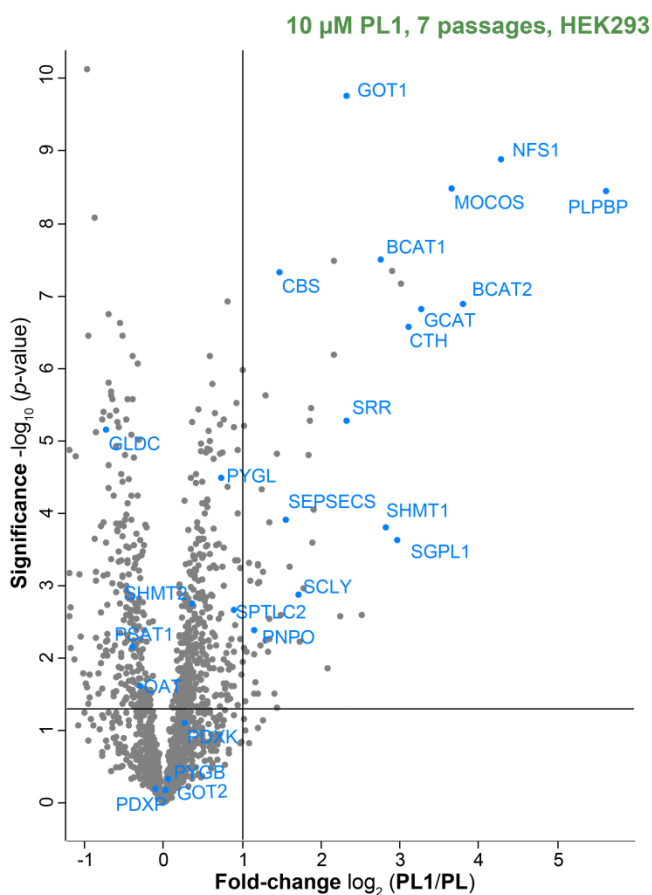
Figure II-14 Metabolic labeling of human cells with PL1 evaluated by LC-MS/MS. PL1 labeling in HEK293 revealed significant enrichment of **PLP**-DEs (gene names, blue). Volcano plot representing *t*-test

results of **PL1** labeling compared to the **PL** control (*n*=6 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines).[239]

3.2 Labeling with PL2

Although no efficient turnover of **PL2** by wild-type hPLK was observed, we nevertheless evaluated labeling efficiency of the probe in human cell lines. First, we performed *in situ* labeling with **PL2**, where we observed concentration-dependent signals in the gel-based fluorescence readout (**Figure II-15A**) with an increased intensity (exposure time ~2-4 s) compared to **PL1** labeling (**Figure II-15B**).
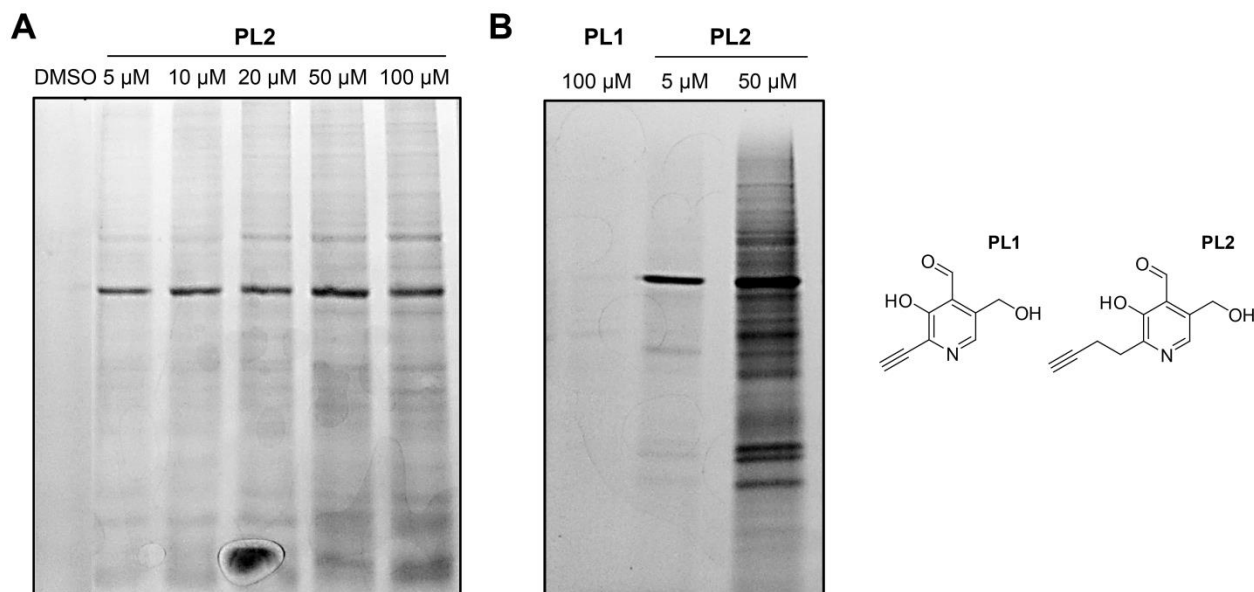
Figure II-15 *In situ* labeling with PL2 probe evaluated by in-gel fluorescence. (**A**) Labeling of HeLa cells with different concentrations of **PL2** for 2 h revealed a concentration-dependent labeling (exposure time ~2 s). (**B**) Comparison of **PL1** and **PL2** labeling for 2 h in HeLa (exposure time ~4 s). For coomassie-stained gels compare **Figure S7**.

The enhanced labeling efficiency of **PL2** indicates that the additional spacer between alkyne and aromatic ring, which decouples the extended conjugated system observed in **PL1**, reduces (at least to a certain extent) attacks by cellular nucleophiles. Strong labeling with **PL2** prompted us to proceed with MS-based target evaluation. In general, high concentrations of 50 and 100 µM **PL2** resulted in a huge amount of non-**PLP**-DE background binders. Albeit this number was reduced upon incubation with 5 or 10 µM of **PL2**, the amount of enriched **PLP**-DEs could not be increased. Also varying incubation times within the range from 30 min to 48 h as well as

screening of different cell lines did not change the labeling pattern, indicating that either the inefficient phosphorylation of **PL2** or uptake of **PL2P** by **PLP**-DEs prevented their enrichment (For MS-based labeling examples see **Figure II-16**).
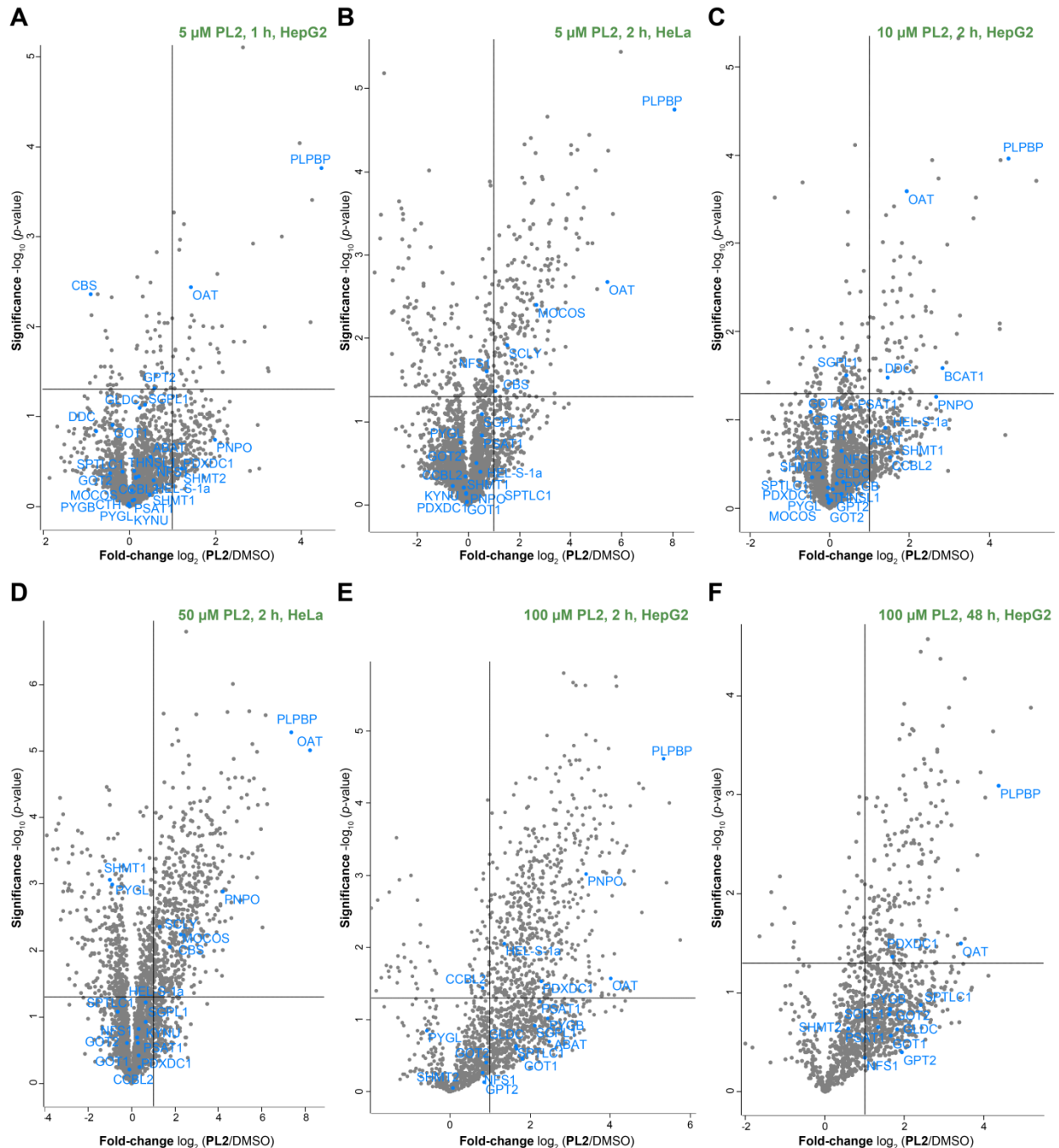
Figure II-16 *In situ* target evaluation of PL2 labeling by LC-MS/MS. Different concentrations of and incubation times with **PL2** were examined in the cell lines indicated. Exemplarily selected volcano plots representing *t*-test results of **PL2** labeling compared to the **PL** control (*n*=3 biological replicates). Cut-off

values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). **PLP**-DEs (gene names) are colored blue.

However, the overall number of enriched proteins upon labeling with **PL2** was strongly increased compared to *in situ* labeling experiments with **PL1**, already suggested by the gel-based readout (**Figure II-15B**). Interestingly, though no efficient capture of **PLP**-DEs was observed, mitochondrial ornithine aminotransferase (*OAT*) could be enriched under all conditions tested. OAT displays a vitamin B6-dependent enzyme which was not enriched upon labeling with **PL1**, indicated a preferred behavior towards a longer 2'-scaffold modification. Notably, a difference in structural constraints within **PLP**-binding sites was already observed for *S. aureus* B6-enzymes.[198] Besides for steric reasons, the more electrophilic nature of the alkyne substitution of **PL1** could for instance have also led to unfavorable reactions within the enzyme active site pocket (*e.g.* with cysteine residues) therefore preventing enrichment by **PL1**. On the other hand, it could have been possible that OAT indeed was taregted by **PL1P**, but probe-bound enzyme could not be enriched due to chemical lability of **PL1**.
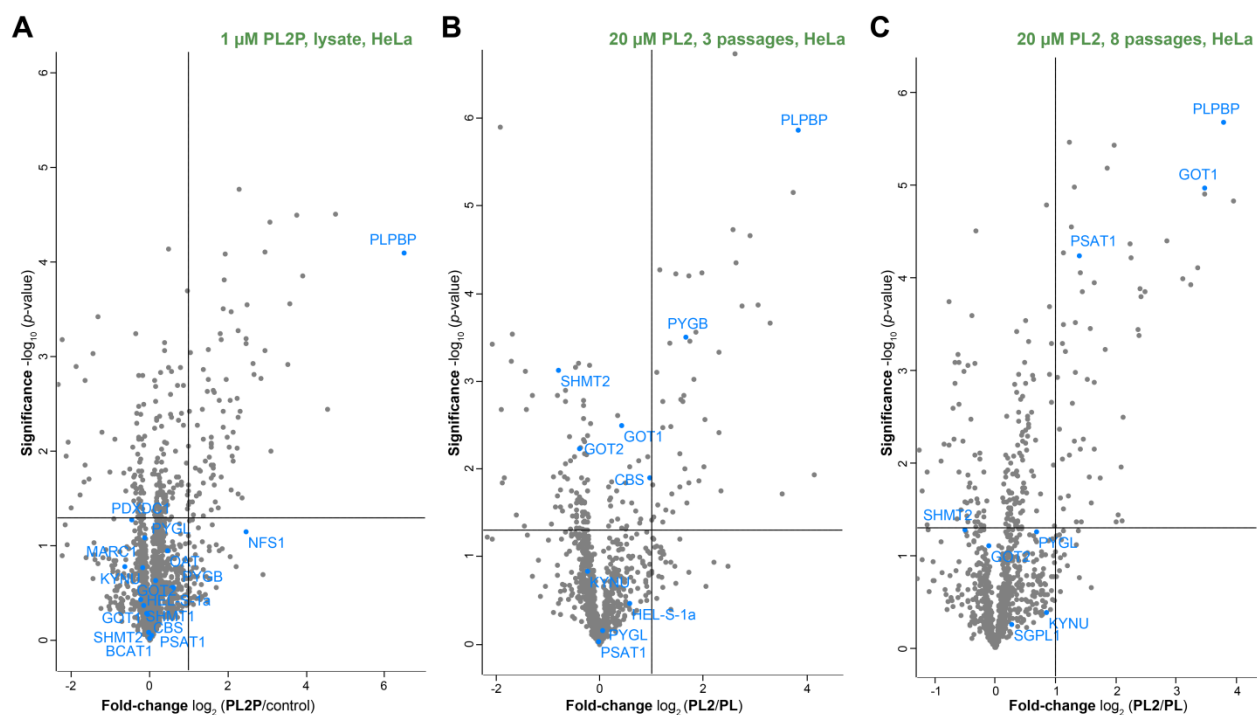
Figure II-17 Target evaluation of labeling with PL2P in the lysate and metabolic labeling with PL2 by LC-MS/MS. (**A**) HeLa cell lysate was incubated with 1 µM **PL2P** for 1 h. HeLa cells were cultured in the presence of 20 µM **PL2** for 3 (**B**) or 8 (**C**) passages, respectively. Volcano plots representing *t*-test results of **PL2** or **PL2P** labeling compared to the **PL** or **PLP** control (*n*=3 biological replicates). Cut-off

values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). **PLP**-DEs (gene names) are colored blue.

To increase the local amount of available phosphorylated probe, we performed labeling in cellular lysate with **PL2P**, pre-phosphorylated by SaPLK. As **PL2** turned out to be more stable towards cellular nucleophiles, we skipped the incubation step with maleimide. As for **PL1P**, we did not observe an increase in **PLP**-DE targets, which could be caused by unfavorable conditions for cofactor uptake under lysate conditions (**Figure II-17A**).

As metabolic labeling led to successful capture of vitamin B6-dependent targets with **PL1**, we grew HeLa cells in the presence of 20 µM **PL2** for 3 (14 days) and 8 passages (40 days), respectively, to allow for direct incooperation of phosphorylated probe into newly synthesized **PLP**-DEs. We decided to double the amount of probe compared to **PL1** which was applied at 10 µM, to account for the hampered phosphorylation efficiency. **PL2** concentrations were not further increased for metabolic labeling as **PLP** carries a very reactive aldehyde that can in principle bind to any ε-amino group in lysine residues which may lead to unspecific effects during cell growth.[246] Moreover, **PLP** was reported as a member of the 30 most damage-prone metabolites[247] and it was shown that cellular growth was inhibited by **PL** concentrations from 100 µM on even after short incubation times of 24 h.[248] Notably, **PL2** was able to support cell growth, suggesting that essential B6-proteins are able to use the analogue to fulfill their biological functions. Though the number of vitamin B6-independent background binders decreased markedly, the overall number of significantly enriched **PLP**-DEs was less than the number resulted from **PL1** metabolic labeling (**Figure II17B** and **C**). Nevertheless, some of the enriched vitamin B6-dependent enzymes like glycogen phosphorylase, brain form (*PYGB*), or phosphoserine aminotransferase (*PSAT1*) were not detected with **PL1**.

Taken together, we did not obtain selective and efficient labeling of **PLP**-DEs with **PL2** comparable to labeling with **PL1** for all conditions and cell lines tested. However, some of the vitamin B6-dependent targets enriched upon **PL2** labeling were not detected in previous experiments with **PL1**. If those **PLP**-DEs are missing because of lower tolerance of human **PLP**-DEs to bulkier 2'-substitution, or because of reduced amounts of phosphorylated probe insight the cell remains to be determined. Further, though we labeled with the same probe, we observed differences in enriched **PLP**-DEs depending on the experimental set-up. For instance, OAT was efficiently captured upon *in situ* labeling experiments up to 48 h but was missing after prolonged incubation times of 14 or 40 days, respectively. In contrast, longer incubation times

with **PL2** were necessary to enrich other vitamin B6-dependent enzymes like GOT1 or PYGB, though they were detected either after 14 or 40 days of incubation.

3.3 Labeling with PL3

Rational gel-based labeling readout with **PL3** was prevented due to huge background labeling observed in the DMSO control when clicked to rhodamine-alkyne (**Figure II-18A**). The reasons for this are unclear, but may stem from lower specificity of inverse click reactions.[223]
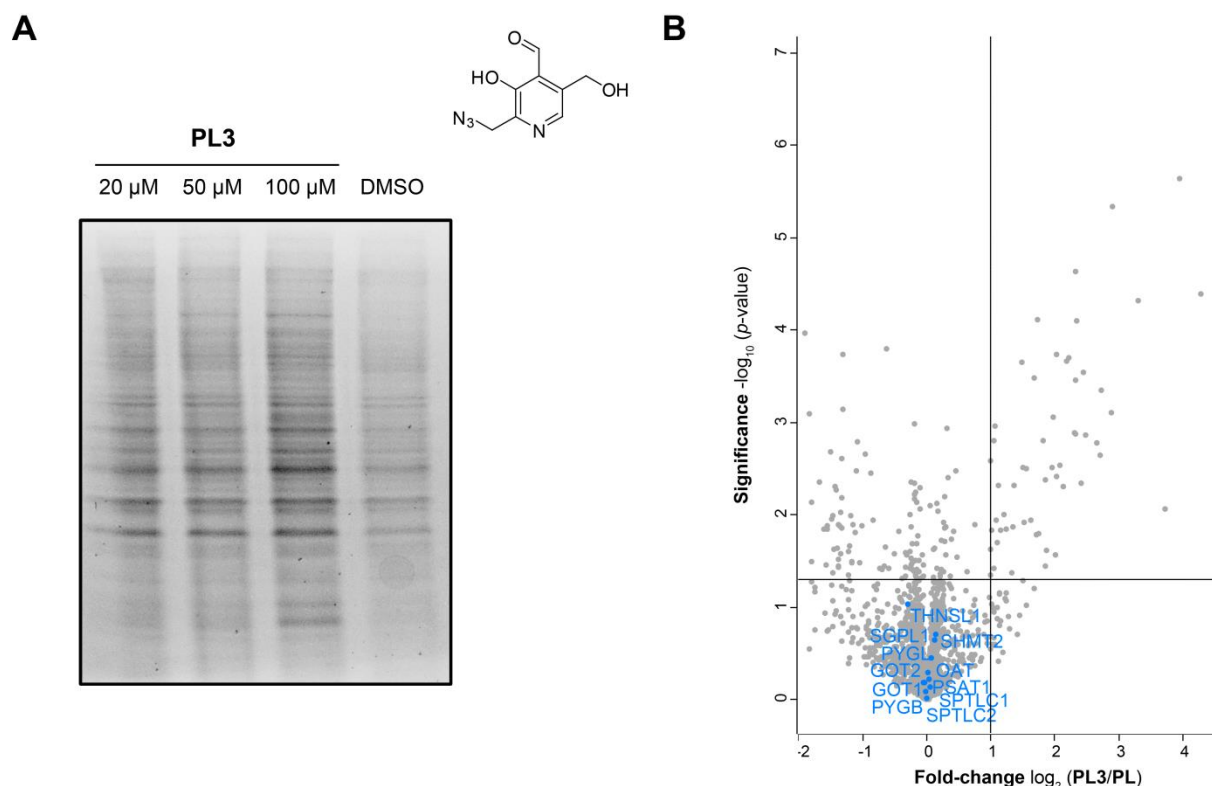
Figure II-18 *In situ* labeling with PL3. (**A**) Analytical labeling with increasing concentrations of **PL3** for 2 h in HeLa and subsequent CuAAC to rhodamine-alkyne. (**B**) LC-MS/MS evaluated labeling of HepG2 with 20 µM **PL3** for 3 days. Volcano plot representing *t*-test results of **PL3** labeling compared to the **PL** control (*n*=3 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). **PLP**-DEs (gene names) are colored blue. For the coomassie-stained gel compare **Figure S7**.

Besides biorthogonal ligations with azide-bearing activity-based probes the *Staudinger* ligation was developed to couple an azide with a triarylphosphine derivative *via* the formation of an amide bond.[249] Intrigued by the high coverage of **PLP**-DEs in *S. aureus* upon labeling with **PL3** and subsequent conjugation to biotin *via Staudinger* ligation,[198] we labeled HepG2 for three days

with 20 µM **PL3** (**Figure II-18B**) and subsequently attached biotin-phosphine using this ligation strategy. Surprisingly, we did not observe any significant enrichment of **PLP**-DEs at all, prompting us to proceed with metabolic labeling. However, several attempts to culture cells in the presence of **PL3** were not successful, as cells were prohibited to grow in the presence of the probe, suggesting that **PL3** could not support the biological function of essential vitamin B6-dependent enzymes. In summary, **PL3** did neither result in any significant labeling of **PLP**-DEs, nor supported the growth of cells, most likely due to the azide-functionality of the compound.

3.4 Conclusion

Taken together, **PL1** was the probe which resulted in the most efficient and selective enrichment of **PLP**-DEs in human cell lines. This is probably caused by the more rapid turnover by hPLK inside the cell. Nevertheless, due to increased electrophilic nature of the alkyne tag, **PL1** can be easily attacked by cellular cysteines, probably reducing the number of detected **PLP**-DEs. Although **PL2** displays a hampered bioactivation, we were able to enrich some vitamin B6-dependent enzymes. These include for instance OAT which was not detected by **PL1**, probably caused by protein-specific active site geometries guiding preferences towards different cofactor modifications. Another reason may be that OAT was indeed labeled with **PL1P**, but its enrichment was prevented by the chemical lability of **PL1**. However, both **PL1** and **PL2** were able to support the growth of human cell lines. In contrast, **PL3** did neither result in any significant enrichment of **PLP**-DEs nor supported cell growth for metabolic labeling, suggesting that essential human **PLP**-DEs could not fulfill their biological functions with an azide-functionalized **PL**-scaffold. Such an essential candidate could likely be PLPBP, as it was strongly enriched throughout all labeling experiments using **PL1** and **PL2**, independent of probe concentration or the incubation time, but was missing upon labeling with **PL3**. As PLPBP was already detected after short labeling times of 30 min, it may incooperate the cofactor faster than other **PLP**-DEs and show increased exchange rates of **PLP**. These observations display an interesting finding as *S. aureus* was able to grow in the presence of **PL1** and **PL3** but **PL2** did not support the growth,[198] indicating significant differences in the identity of essential **PLP**-DEs or their behavior towards the probes.

Taken all significantly enriched **PLP**-DEs by both **PL1** and **PL2** from all conditions tested together, we were able to cover for 23 **PLP**-DEs which account for 44% of the total human pool (**Table II-1**) (53 reviewed proteins annotated with GO term 0030170: **PLP**-binding).[127,250]

Part II-3 *En route* to an optimized labeling strategy for PL-probes

Table II-1 Significantly enriched PLP-DEs (GO term 0030170: PLP-binding) with PL1 and PL2 from all conditions tested.

PLP-DE	Gene name	Targeted with probe
5-aminolevulinic acid synthase	ALAS1	**PL1**
Branched-chain amino acid aminotransferase, cytosolic	BCAT1	**PL1**, **PL2**
Branched-chain amino acid aminotransferase, mitochondrial	BCAT2	**PL1**
Cystathionine-β-synthase	CBS	**PL1**, **PL2**
Cystathionine-γ-lyase	CTH	**PL1**
Aromatic L-amino acid decarboxylase	DDC	**PL2**
Glycine C-acetyltransferase	GCAT	**PL1**
Aspartate aminotransferase, cytosolic	GOT1	**PL1**, **PL2**
Molybdenum cofactor sulfurase	MOCOS	**PL1**, **PL2**
Cysteine desulfurase, mitochondrial	NFS1	**PL1**
Ornithine aminotransferase	OAT	**PL2**
Pyridoxal-dependent decarboxylase domain-containing protein 1	PDXDC1	**PL2**
Pyridoxal 5'-phosphate binding protein	PLPBP	**PL1**, **PL2**
Pyridoxine/pyridoxamine 5'-phosphate oxidase	PNPO	**PL1**, **PL2**
Phosphoserine aminotransferase	PSAT1	**PL2**
Glycogen phosphorylase, brain form	PYGB	**PL2**
Selenocysteine lyase	SCLY	**PL1**, **PL2**
o-phosphoseryl-tRNA(Sec) selenium transferase	SEPSECS	**PL1**
Sphingosine-1-phosphate lyase 1	SGPL1	**PL1**
Serine hydroxymethyltransferase, cytosolic	SHMT1	**PL1**
Serine hydroxymethyltransferase, mitochondrial	SHMT2	**PL1**
Serine palmitoyltransferase 2	SPTLC2	**PL1**
Serine racemase	SRR	**PL1**

4 Phosphorylated cofactor-mimics bind to recombinant human **PLP**-DEs

As **PLP** displays a highly reactive aldehyde which can spontaneously form aldimines with any ε-amino groups of lysine residues of non-**PLP**-DEs,[246] it is not surprising that we detected some background binders among the significantly enriched proteins although these are mostly less enriched than known **PLP**-DEs. To validate that **PLP**-DEs are targeted by the phosphorylated versions of the probe, we exemplarily selected SHMT1 which was highly enriched by the metabolic labeling approach with **PL1**, as well as PLPBP (cf. Part II, Chapter 8) as common target of both **PL1** and **PL2** under all examined conditions.

4.1 PL1P binds to recombinant human SHMT1

Inspection of the active site pocket of SHMT1 with bound **PLP** (PDB: 1BJ4[213]) reveals sufficient space for the additional alkyne group of **PL1** (**Figure II-19A**). Strep-tagged SHMT1 was initially purified as the holo-enzyme and transformed to the apo-form by nucleophilic displacement of bound **PLP** using hydroxylamine.[251] Internal aldimine absorbance reappeared upon reconstitution of apo-SHMT1 with **PL1P** and **PLP** as a control (SHMT1-**PL1P** and SHMT-**PLP**) (**Figure II-19B**). Notably, the absorbance maximum of SHMT1-**PL1P** was redshifted compared to holo-SHMT1 and SHMT1-**PLP** which is in line with changes observed for unbound **PL1P**. Thermal denaturation studies revealed a selective increase in melting temperature of apo-SHMT1 upon incubation with increasing **PLP** amounts, whereas un-phosphorylated vitamer did not lead to changes (**Figure II-19C**, left). **PL1P** resulted in significant stabilization of the apo-enzyme as well, suggesting that probe and natural vitamin B6 behave similarly (**Figure II-19C**,

right). Finally, apo-SHMT1 was incubated with **PL1** or **PL1P**, respectively, the internal aldimine reduced with NaBH$_4$ and the protein conjugated to a rhodamine tag using CuAAC. SDS-PAGE with fluorescence read-out revealed specific and concentration-dependent labeling of SHMT1 with **PL1P** whereas only weak background labeling when incubated with excess of **PL1** was observed (**Figure II-19D**).
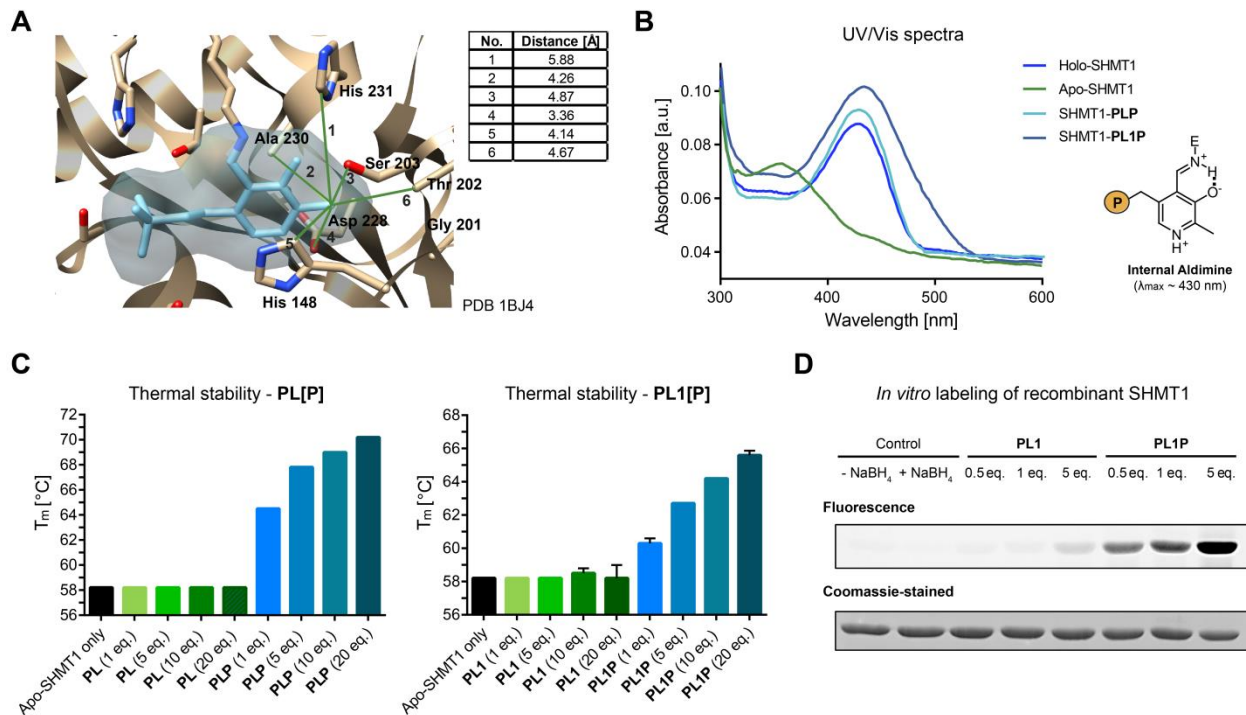
Figure II-19 PL1P is the species that binds to SHMT1. (**A**) Active site pocket of SHMT1 with corresponding distances to residues surrounding the 2'-methyl group of **PLP** (PDB: 1BJ4[213]). (**B**) UV/Vis spectra of holo-, apo-, and apo-SHMT1 reconstituted with **PLP** and **PL1P** (SHMT1-**PLP**, SHMT1-**PL1P**) displayed a successful recovering of internal aldimine absorbance from apo-SHMT1 (P = OPO$_3^{2-}$). (**C**) Thermal denaturation studies of apo-SHMT1 revealed a selective stabilization upon **PLP** (left) and **PL1P** (right) incubation (*n*=3). (**D**) *In vitro* labeling of recombinant apo-SHMT1 displayed selective and concentration-dependent labeling by **PL1P**.[239]

4.2 PL1P and PL2P bind to recombinant human PLPBP

We purified Strep-tagged PLPBP and initiated our validation experiments with **PL1** and **PL1P**. Thermal stability assays with PLPBP dimer displayed a melting temperature shift from 57 °C to around 66 °C upon incubation with increasing amounts of **PLP** whereas no changes were observed upon incubation with **PL** (**Figure II-20A**, left). As for SHMT1, similar results were obtained with **PL1** and **PL1P**, where only **PL1P** lead to a stabilization of PLPBP (**Figure II-20A**,

right). In line with thermal stability assays, *in vitro* labeling of recombinant PLPBP dimer revealed a concentration-dependent fluorescence increase exclusively for **PL1P** whereas no labeling was observed for **PL1** (**Figure II-20B**) indicating that **PL1P** is the species bound under cellular conditions. Similar results were obtained for PLPBP monomer, resulting from a unique behavior towards **PL1** and **PL1P** independent of the oligomeric state (**Figure II-20C** and **D**).
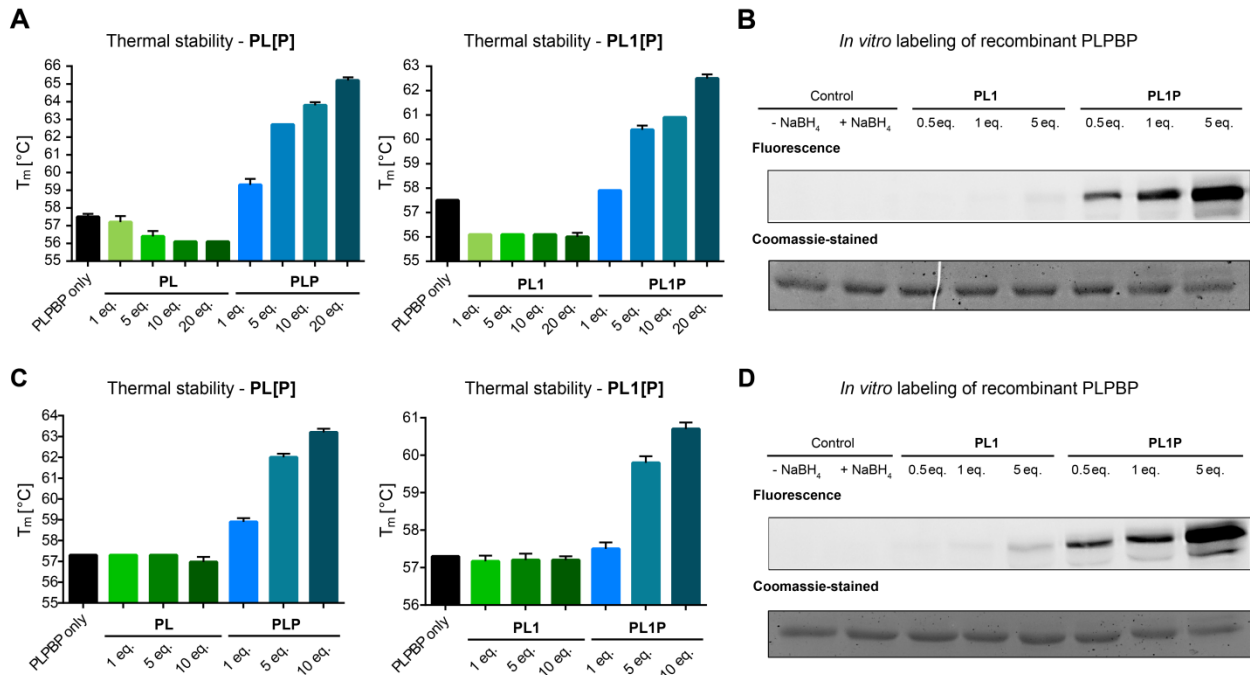
Figure II-20 PL1P is the species that binds to PLPBP. Thermal stability assays displayed a selective stabilization of PLPBP dimer (**A**), as well as monomer (**C**) by **PLP** (left) and **PL1P** (right). In contrast, **PL** and **PL1** did not lead to an increase in melting temperature (*n*=3). In line with thermal shift data, recombinant dimeric (**B**), as well as monomeric (**D**) PLPBP were exclusively labeled upon incubation with **PL1P** in a concentration dependent manner.[239]

We further performed thermal stability assays with **PL2** and **PL2P**, as well as *in vitro* labeling. Therefore, we exemplarily selected monomeric PLPBP, as we demonstrated that **PLP** has no effect on oligomerization (cf. Part II, Chapter 8), and that monomer and dimer behave similar in validation experiments with **PL1** and **PL1P**. As for **PL1P** we detected an exclusive stabilization (**Figure II-21A**), as well as concentration-dependent labeling upon incubation with **PL2P** (**Figure II-21B**). Only a slight background labeling was observed for an excess of **PL2**.
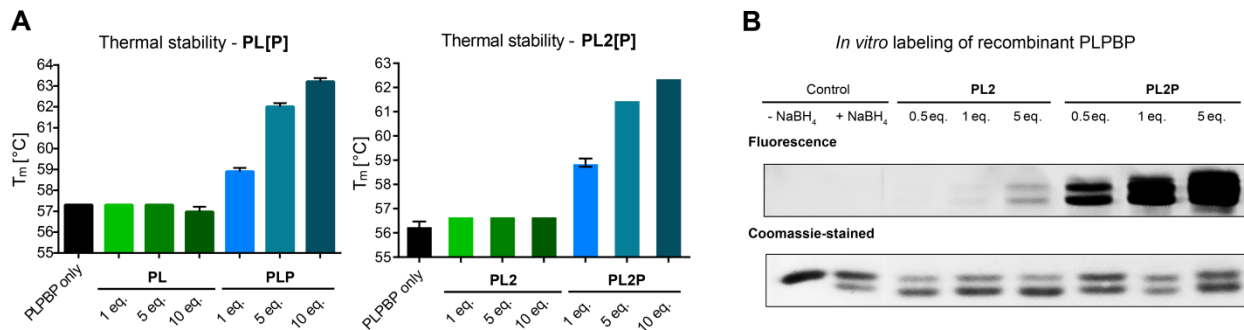
Figure II-21 PL2P is the species that binds to PLPBP. (**A**) Thermal stability assays displayed a selective stabilization of PLPBP monomer by **PLP** (left) and **PL2P** (right). In contrast, **PL** and **PL2** did not lead to an increase in melting temperature (n=3). (**B**) In line with thermal shift data, recombinant monomeric PLPBP was mainly labeled by **PL2P** in a concentration dependent manner.

4.3 Competition with PLP

To further validate that **PL**-probes bind to **PLP**-DEs active sites, we performed competitive labeling with the natural cofactor. Therefore, we exemplarily incubated **PL1P**-labeled SHMT1 and PLPBP with increasing excess of **PLP** over probe, reduced samples and performed CuAAC to attach rhodamine for fluorescence readout following SDS-PAGE analysis (**Figure II-22**).

Satisfyingly, a 10-fold molar excess of **PLP** efficiently outcompeted **PL1P** labeling of PLPBP monomer and dimer. In contrast, a 100-fold molar excess was required in the case of SHMT1. This is most likely caused by the fact that SHMT1 was applied as the apo-form previous to **PL1P** labeling, which can therefore bind a higher amount of probe that needs to be outcompeted further on. In the case of PLPBP, **PLP** could not completely be removed during the purification. Attempts to prepare the apo-version of the protein were not successful as PLPBP misfolds or aggregates when incubated with hydroxylamine or cysteine to displace **PLP** (data not shown). Residual amounts of **PLP** in PLPBP therefore reduce the amount of **PL1P** than can be taken up previous to competition experiments.

A SHMT1 - **PLP** excess over **PL1P**

0 0.5 1 2 10 20 50 100 200

Fluorescence

Coomassie-stained

B PLPBP Dimer - **PLP** excess over **PL1P**

0 0.5 1 2 10 20 50 100 200

Fluorescence

Coomassie-stained

C PLPBP Monomer - **PLP** excess over **PL1P**

0 0.5 1 2 10 20 50 100 200

Fluorescence

Coomassie-stained
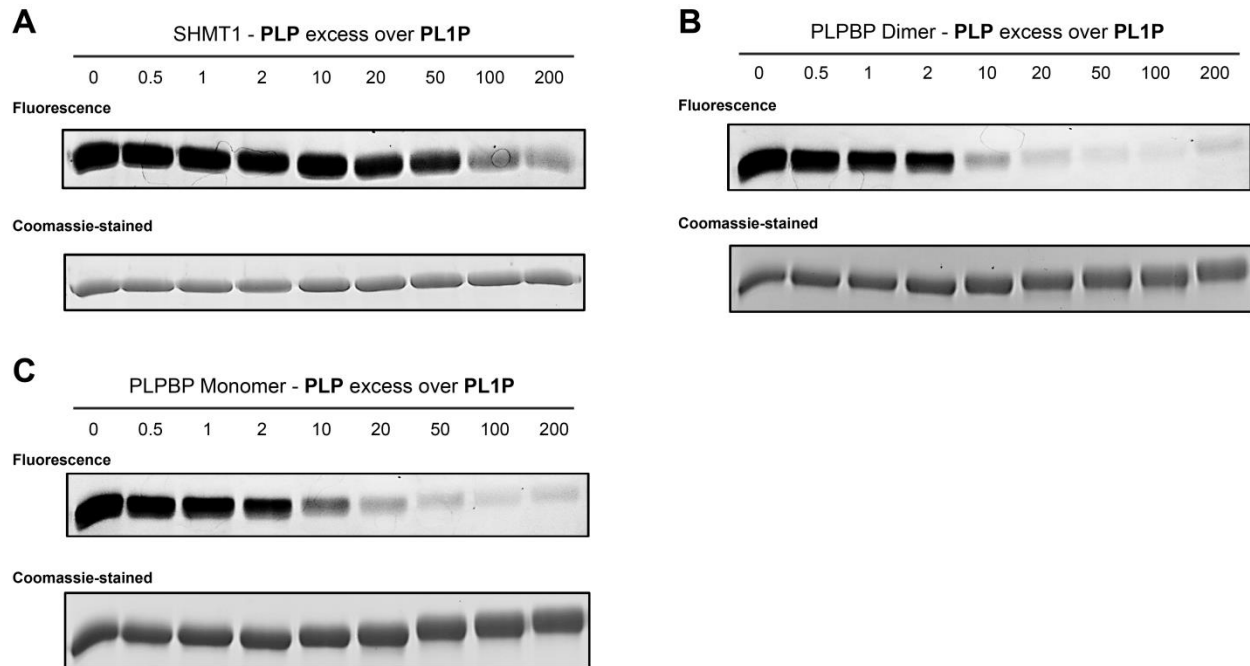
Figure II-22 Competition of PL1P-labeling with PLP. Apo-SHMT1 (**A**), PLPBP dimer (**B**), and monomer (**C**) were labeled with **PL1P**, previous to addition of varying molar excess of **PLP**. Proteins were reduced and clicked to rhodamine for fluorescence readout.

4.4 Conclusion

Taken together we demonstrated that phosphorylated **PL**-probes selectively bind to equal lysine residues than natural cofactor in both SHMT1 and PLPBP. Although we observed especially for cellular **PL2**-labeling under several conditions a huge number of background binders which are not vitamin B6-dependent, these validation experiments support that the species bound to **PLP**-DEs under cellular conditions is the phosphorylated version of the probe.

5 _In situ_ target screening of vitamin B6 antagonists

As **PLP**-DEs inhibitor development is often challenged by conserved cofactor binding sites, _in situ_ target screening of vitamin B6 antagonists displays an important step to characterize their cellular selectivity. However, a method for fast and efficient readout is still lacking. We therefore previously applied our cofactor mimics to uncover new _off_-targets of DCS in _S. aureus._[198] Interestingly, the main vitamin B6-dependent cellular targets of DCS turned out to be D-alanine aminotransferase, diaminopimelate decarboxylase, as well as Orn/Lys/Arg decarboxylase instead of alanine racemase as previously believed.[252] To adapt this selectivity screen for human cells we exemplarily selected the **PLP**-DEs inhibitors penicillamine (Pen) and aminooxyacetic acid (AOAA). Both inhibitors display poorly characterized cellular specificities and thus represent excellent candidates for B6-dependent competition experiments in **PL1**-labeled HEK293.

We initiated the selectivity profiling with AOAA, an inhibitor of hydrogen sulfide (H_2S) biosynthesis. Humans produce small amounts of H_2S which exerts multiple effects as a signaling molecule.[253] For instance, H_2S participates in the regulation of cellular metabolism and immunological as well as inflammatory responses.[253,254] Further, it suppresses the proliferation of colorectal cancer cells and reduces tumor growth.[255] AOAA covalently binds to **PLP**-DEs _via_ formation of an external aldimine with the cofactor (**Figure II-23A**).[256,257] We selected an inhibitor concentration of 1 mM for our competition experiments, as previous metabolomics studies with intact HCT116 cells were performed at a concentration of 1 mM for 24 h.[245] Indeed, competition experiments confirmed binding to **PLP**-DEs involved in the H_2S pathway (**Figure II-23B**), in particular to cystathionine γ-lyase (CTH). This finding is in line with studies suggesting that CTH is the major target of AOAA and not cysthationine β-synthase (CBS) as previously believed.[258]

Remarkably, although proposed as a general **PLP**-DEs inhibitor *in vitro* (e.g. for CBS[258] or GOT1[245]), AOAA did not show competition with other detected **PLP**-DEs indicating a high selectivity *in situ*. However, several non-**PLP**-DEs were depleted upon incubation with AOAA. This could be explained by the general high reactivity of *N*-derived nucleophiles to the human proteome,[259] which might have led to secondary effects or destabilization of some proteins.
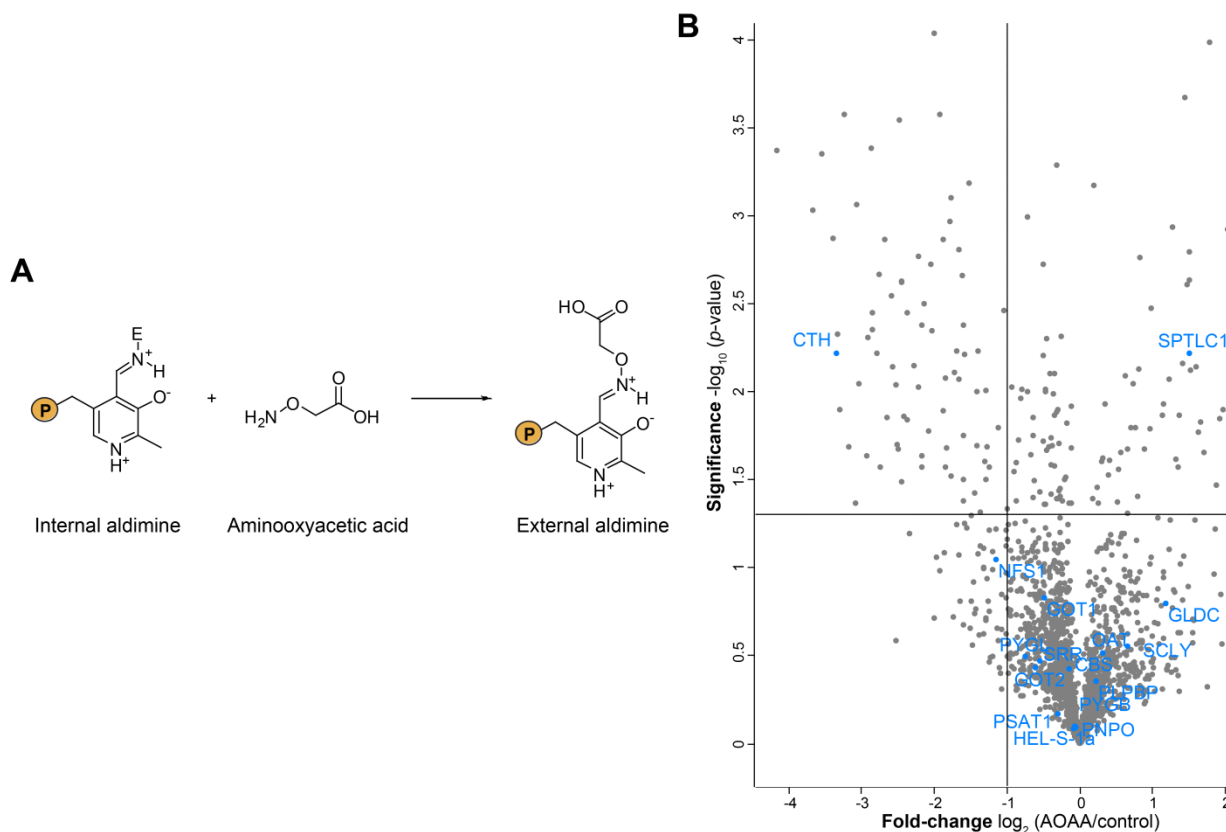
Figure II-23 *In situ* target screening of AOAA. (**A**) Reaction mechanism of AOAA with the **PLP** cofactor (P = OPO_3^{2-}). (**B**) **PL1**-labeled HEK293 cells were incubated with 1 mM of AOAA for 2 h. Volcano plot representing *t*-test results of competition compared to the **PL1**-labeled control revealed significant depletion of one **PLP**-DE (blue) (*n*=3 biological replicates). Cut-off values were defined as depletion factor of log_2 = -1 (2-fold depletion) and $-log_{10}$ (*p*-value) of 1.3 (solid lines).

We next analyzed Pen, a degradation product of penicillins[260] which is therapeutically applied in the treatment of rheumatoid arthritis[261] and Wilsons's disease.[262] Further, it is used to treat heavy metal poisonings as it forms complexes with several toxic metals like cadmium, mercury, and gold.[263] Although the D-enantiomer is the natural occurring isomer, both configurations of Pen display an anti-vitamin B6 effect.[264–266] Several **PLP**-DEs including CTH,[244] SHMT1,[267] and GOT1[268] were shown to be inhibited by Pen *in vitro*. However, our competitive proteomics studies in HEK293 using L- and D-Pen revealed only a limited number of **PLP**-DE targets

Part II-5 *In situ* target screening of vitamin B6 antagonists

(**Figure II-24A** and **B**) suggesting a higher selectivity *in situ*. We choose 10 mM Pen for competition experiments, as experiments with recombinant **PLP**-DEs required for a higher concentration of 10 mM to display inhibitory effects.[244,269] The cofactor is displaced by Pen *via* formation of a thiazolidine adduct (**Figure II-24C**).[270]
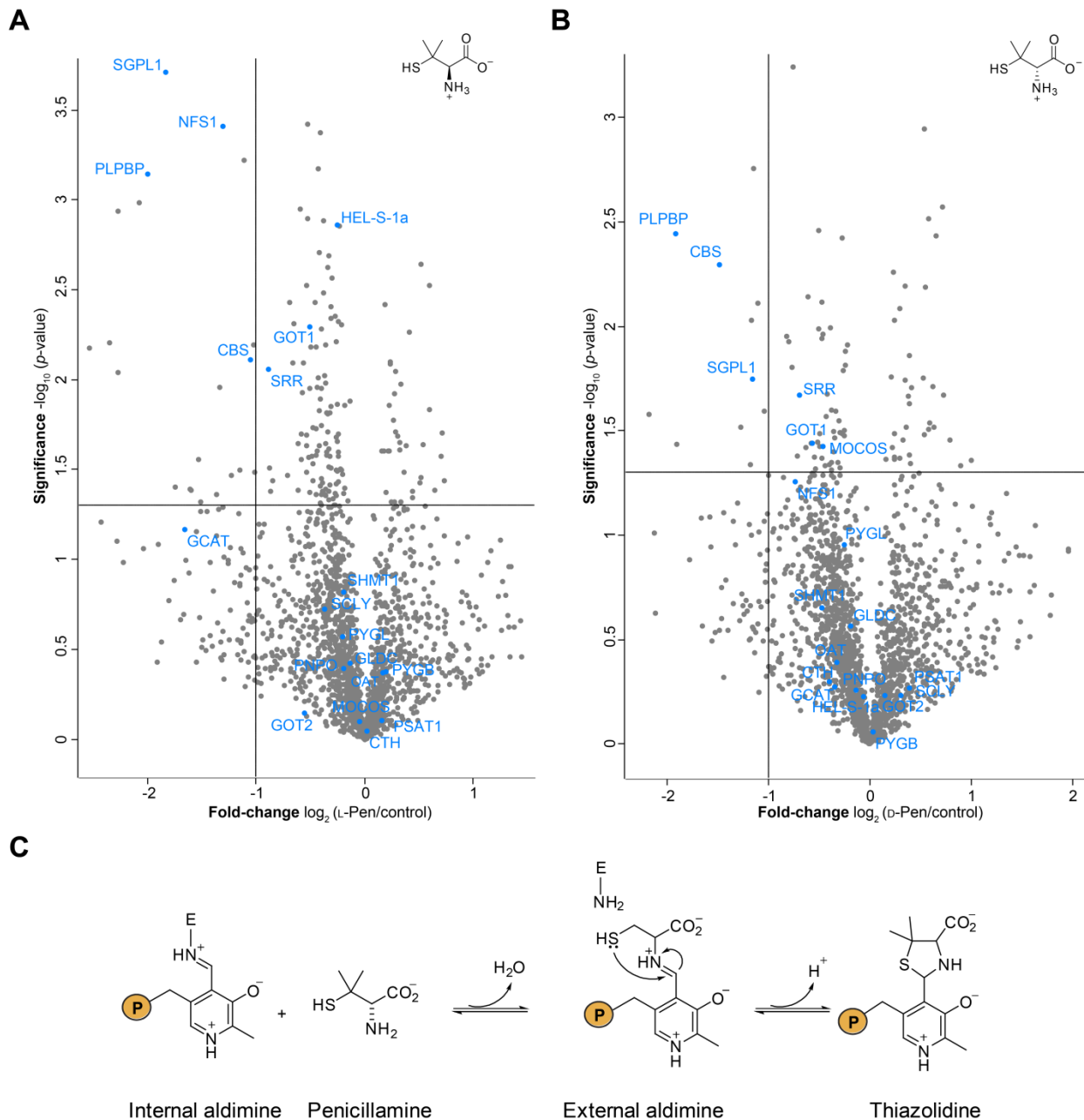
Figure II-24 *In situ* target screening of penicillamine (Pen). PL1-labeled HEK293 cells were incubated with 10 mM of L- (**A**) and D-Pen (**B**) for 2 h, respectively. Volcano plots representing *t*-test results of competition compared to the **PL1**-labeled control revealed significant depletion of **PLP**-DEs (blue) (*n*=3 biological replicates). Cut-off values were defined as depletion factor of log$_2$ = -1 (2-fold depletion)

and -log$_{10}$ (*p*-value) of 1.3 (solid lines). (**C**) Mechanism of thiazolidine-adduct formation of Pen with **PLP** (P = OPO$_3^{2-}$).[239]

Both enantiomers show a similar target scope including PLPBP and SGPL1. Contrary to previous postulations, the primary target within the H$_2$S pathway is CBS instead of CTH.[244] Interestingly, mitochondrial cysteine desulfurase (NFS1) was competed exclusively by L-Pen indicating an absolute configuration dependent selectivity.

To validate our competition results, we incubated recombinant PLPBP with increasing concentrations of L- and D-Pen. UV/Vis spectra displayed a decrease in absorbance of the internal aldimine for both enantiomers (**Figure II-25**), confirming that Pen indeed targets PLPBP. Interestingly, D-Pen turned out to displace **PLP** more effectively from recombinant PLPBP, though both configurations display similar competition *in situ*. This observation is not surprising, as many **PLP**-DEs inhibitors, including AOAA and Pen, show effects on various vitamin B6-dependent enzymes *in vitro*, but display a different selectivity *in situ* as demonstrated by the competition data.
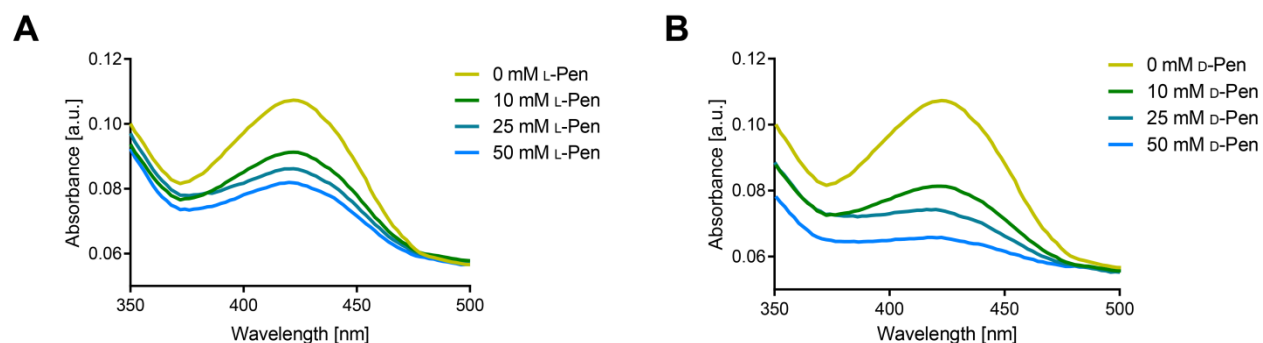
Figure II-25 UV/Vis spectra of PLPBP dimer incubated with increasing concentrations of L- (A) and D-Pen (B).[239]

6 **PL1** labeling profiles in different human cell lines

Though several **PLP**-DEs display altered activity at equal expression levels in different kinds of cancer,[136,215] mechanisms behind this phenomenon are largely unexploited.

As we were able to set up an efficient workflow for labeling of **PLP**-DEs in human HEK293 with the **PL1**-probe, our goal was the screening of other cell lines in a next step. We wanted to determine if labeling of this enzyme class across various cell lines would reveal signature profiles characteristic for an individual cell line. For this, we cultivated the human cancer-derived cell lines K562 (chronic myelogenous leukemia), HeLa (cervix epitheloid carcinoma), as well as HCT116 (colorectal carcinoma) in the presence of 10 µM **PL1** for seven passages as we did for HEK293. For all cell lines we did not observe any changes in doubling rate and morphology upon incubation with **PL1**. Satisfyingly we observed significant enrichment of several **PLP**-DEs in the different cell lines (**Figure II-26**).

As for the HEK293 labeling, several non-**PLP**-binders were among the enriched proteins, though vitamin B6-dependent proteins showed mostly the highest fold-changes. Enrichment of non-**PLP**-DEs by the vitamin B6-probe may be a result of background binding to lysine residues. Indeed, **PLP** displays a highly reactive aldehyde which can in principle bind to any ε-amino group on lysine residues.[246] In addition, we compared significantly enriched proteins from our data sets with proteins enriched by an amine-reactive alkyne probe, which was applied for profiling global lysine reactivity in the human proteome.[271] Interestingly, a huge number of the non-**PLP**-DEs enriched by **PL1** were detected upon incubation with the amine reactive probe as

well, indicating that lysine residues within that proteins show nucleophilic characters. Several proteins annotated to not bind vitamin B6 enriched by **PL1** showed medium reactivity towards the amine-reactive compounds (*e.g. ACADVL, ACTR2, PDIA3, P4HB*), some were identified as hyper-reactive (*e.g. GCLM, EPHX1*) and could therefore also bind the **PL1** probe in an unspecific manner.
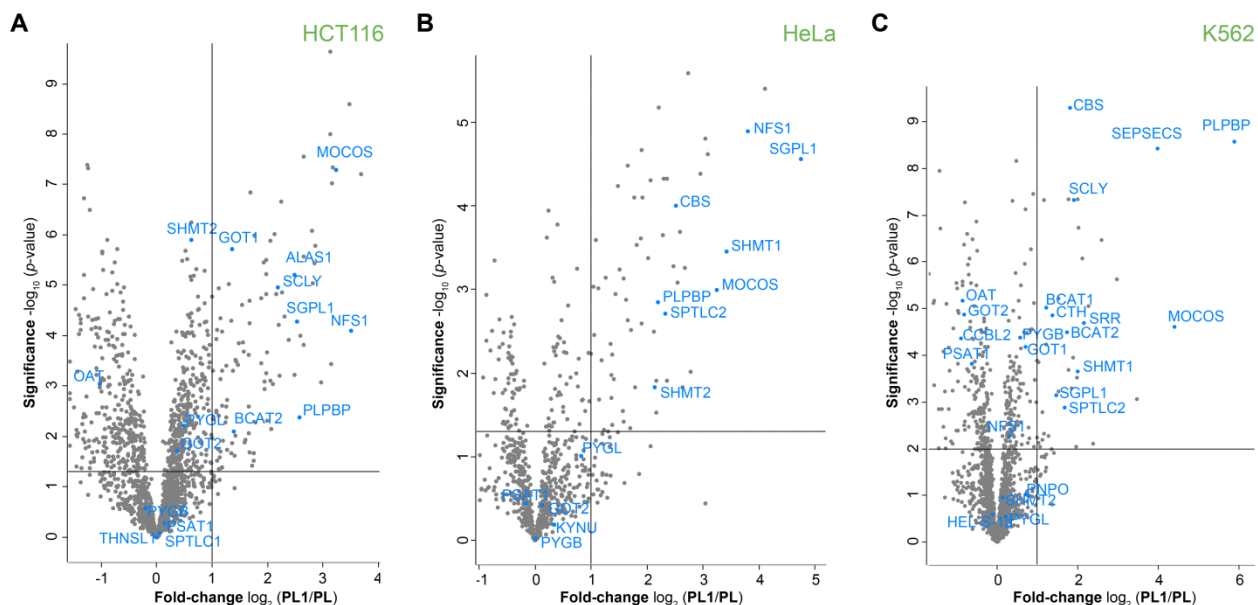
Figure II-26 Metabolic labeling of human cancer-derived cells with PL1. Volcano plots representing *t*-test results of **PL1** labeling in HCT116 (**A**), HeLa (**B**), and K562 (**C**) compared to the **PL** control revealed significant enrichment of **PLP**-DEs (blue) (n=6 biological replicates for K562 and HCT116, n= 4 biological replicates for HeLa). Cut-off values were defined as enrichment factor of $\log_2$ = 1 (2-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines).[239]

However, to evaluate if one of the proteins annotated as non-**PLP**-dependent does indeed bind to vitamin B6, we selected glucosylceramidase (*GBA*) as it was the only common significantly enriched hit among the non-B6 binders throughout all cell lines examined (**Figure II-27A**). As a characteristic feature of **PLP**-DEs is the absorbance of the internal aldimine at around 430 nm, we recorded an UV/Vis spectrum of GBA (**Figure II-27B**). We did not observe a corresponding absorbance maximum in the visible region, suggesting that GBA is not a **PLP**-DE. Further, as we demonstrated that a markedly increase in melting temperature upon incubation with **PLP** and **PL1P** is a second characteristic feature of **PLP**-DEs, we incubated GBA with each 20-fold excess of **PL**, **PLP**, **PL1**, and **PL1P** (**Figure II-27C**). Neither for the natural vitamer nor for **PL1P** we detected such an substantial increase in the melting temperature, indicating that GBA is not a **PLP**-DE in line with the UV/Vis data.
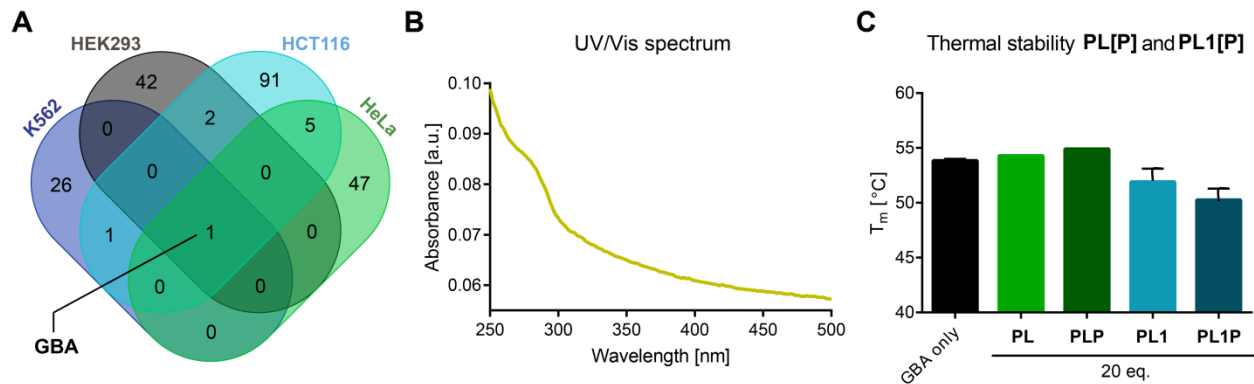
Figure II-27 GBA as common non-vitamin B6-binding target is not a PLP-DE. (**A**) GBA is the only non-vitamin B6-binding protein significantly enriched in all cell lines examined. (**B**) UV/Vis spectra of GBA. The characteristic absorbance maximum of the internal aldimine at around 430 nm is missing. (**C**) Thermal stability data display that GBA is not markedly stabilized by **PLP** or **PL1P** as observed for genuine **PLP**-DEs (*n*=3).[239]

Taken all studied cell lines together, we were able to enrich 18 **PLP**-DEs upon metabolic labeling with **PL1**, which account for 34% of the complete human **PLP**-ome (53 reviewed proteins annotated with GO term 0030170: **PLP**-binding).[127,250] Interestingly, differences in the number and fold-changes of significantly enriched **PLP**-DEs were observed dependent on the type of cell (**Figure II-28A**). Enzymes like molybdenum cofactor sulfurase (*MOCOS*), sphingosine-1-phosphate lyase 1 (*SGPL1*) or PLPBP were enriched throughout all cell lines studied. Remarkably, other **PLP**-DEs were exclusively detected by our method in particular cell lines like mitochondrial 5-aminolevulinate synthase (*ALAS1*) in HCT116 or mitochondrial 2-amino-3-ketobutyrate coenzyme A ligase (*GCAT*) in HEK293.

To examine if different labeling intensities stem from cell-type specific protein expression levels, we performed western blots against various **PLP**-DEs in the different cell lines (**Figure II-28B** and **S8**). Interestingly, protein amounts were comparable throughout all cell lines, suggesting an altered cofactor loading of **PLP**-DEs dependent on the type of cell. This observation is also supported by expression levels reported in the human proteome database.[272] To further investigate whether different labeling patterns are caused by differences in expression levels of the vitamin B6 salvage enzymes, we examined protein levels of hPLK and PNPO as well (**Figure II-28C** and **S8**). As for **PLP**-DEs, protein amounts of salvage enzymes were alike throughout the different cell lines.
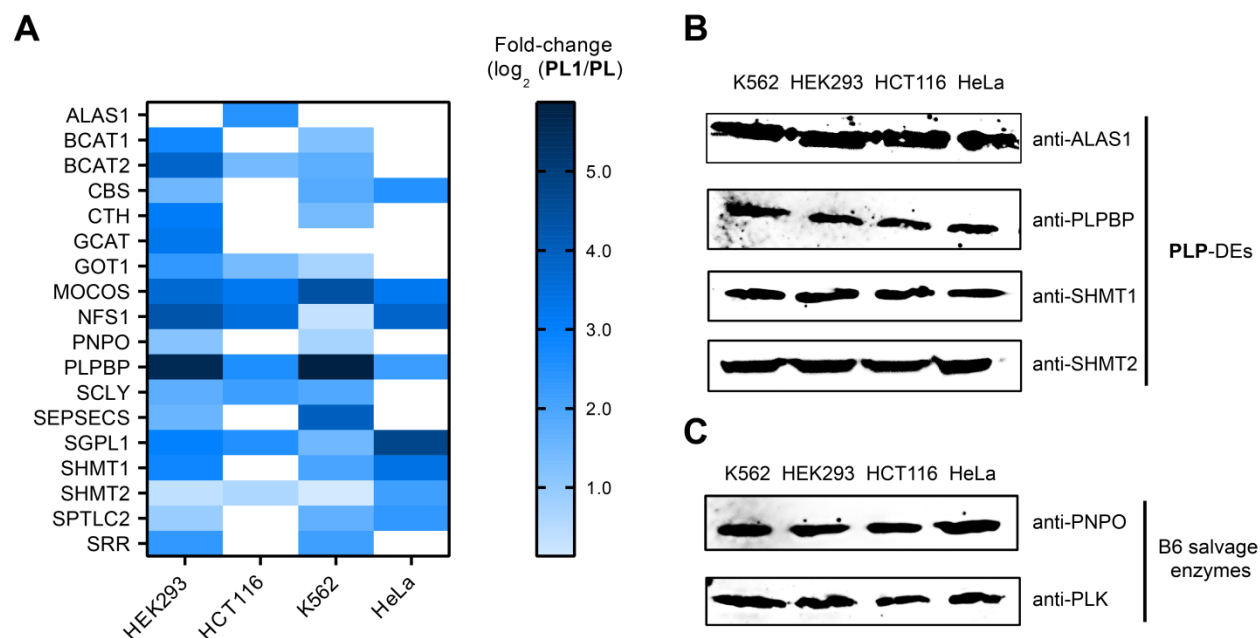
Figure II-28 Comparison of PL1-labeling profiles in various human cell lines. (**A**) Heat-map illustrating the corresponding fold-changes (log$_2$ **PL1/PL**) of **PLP**-DEs in the different cell lines. The map depicts only **PLP**-DEs which are significantly enriched in at least one cell line. In the case that a **PLP**-DE was not detected in samples of a certain cell line, the corresponding rectangle is colorless. Western blots against the **PLP**-DEs ALAS1, PLPBP, SHMT1, and SHMT2 (**B**), as well as vitamin B6 salvage enzymes PNPO and hPLK (**C**) revealed a similar expression of corresponding proteins in all cell lines studied. For Ponceau S-stained membrane compare **Figure S8**.[239]

It could be possible that loading with cofactor may be cell-type specific and does not necessarily correlate with abundance, which may display an additional means of enzyme activity regulation. To validate this hypothesis independent of probe labeling we performed a western blot analysis of NaBH$_4$-reduced and not reduced proteomes with an anti-**PL** antibody raised against reduced **PL** in conjugation to BSA.[273] Specificity was validated by incubation of the antibody with hPLK, which does not covalently bind **PLP**, NaBH$_4$-reduced forms of PLPBP and SHMT1, as well as the non-reduced PLPBP and SHMT1 counterparts (**Figure II-29A**). Satisfyingly, the antibody recognizes reduced **PLP**-DEs well with slighter labeling of their non-reduced counterparts. Moreover, hPLK lacking **PLP** was not recognized at all. In line with the chemoproteomic data the western blot analysis revealed different loading states of **PLP**-DEs as well (**Figure II-29B**, left). Non-reduced proteomes displayed only slight background labeling further validating antibody specificity (**Figure II-29B**, right). Notably, for HeLa cells we observed the lowest number of **PLP**-DEs in both, the proteomic, as well as the western blot analysis.
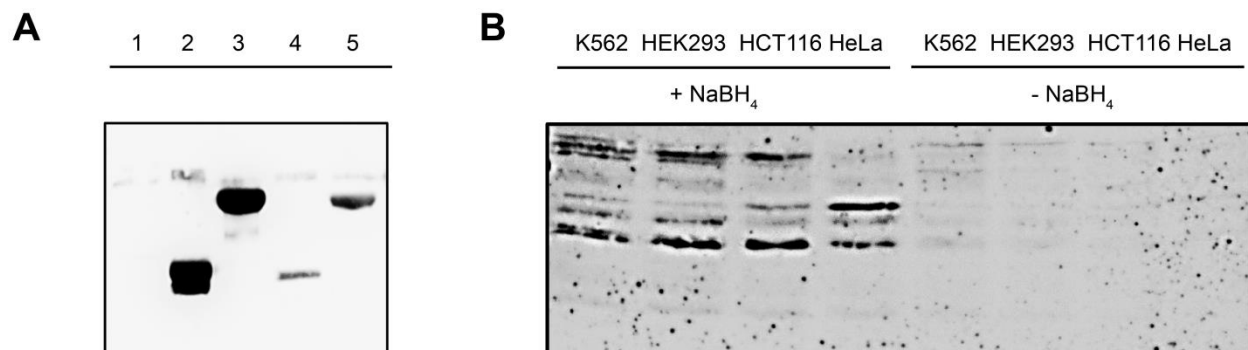
Figure II-29 PLP-ome analysis using an anti-PL antibody raised against reduced cofactor in conjugation with BSA. (**A**) Antibody specificity was validated upon incubation with hPLK (1), which does not covalently bind **PLP**, NaBH$_4$-reduced forms of PLPBP (2) and SHMT1 (3), as well as the non-reduced PLPBP (4) and SHMT1 (5) counterparts. (**B**) Western blot analysis of cofactor loading states of **PLP**-DEs in NaBH$_4$-treated as well as not reduced proteomes. For Ponceau S-stained membrane compare **Figure S8**.[239]

Taken together, the results support cell-type specific alterations in **PLP**-DEs cofactor loading states independent of labeling with the probe, which might indicate a so far not recognized means of enzyme activity regulation. Regulation on the level of cofactor loading can occur faster than regulation on the protein expression level and may therefore serve the cell as fast adaption to changing environmental conditions.

7 Cellular binding partners of vitamin B6 salvage enzymes

Although mechanism and structure of **PLP**-DEs have been studied extensively, little is known on how the cofactor is supplied to apo-enzymes. Further, **PLP** contains a very reactive aldehyde[246] and therefore a tight regulation of biogeneration and free cellular pools is necessary.
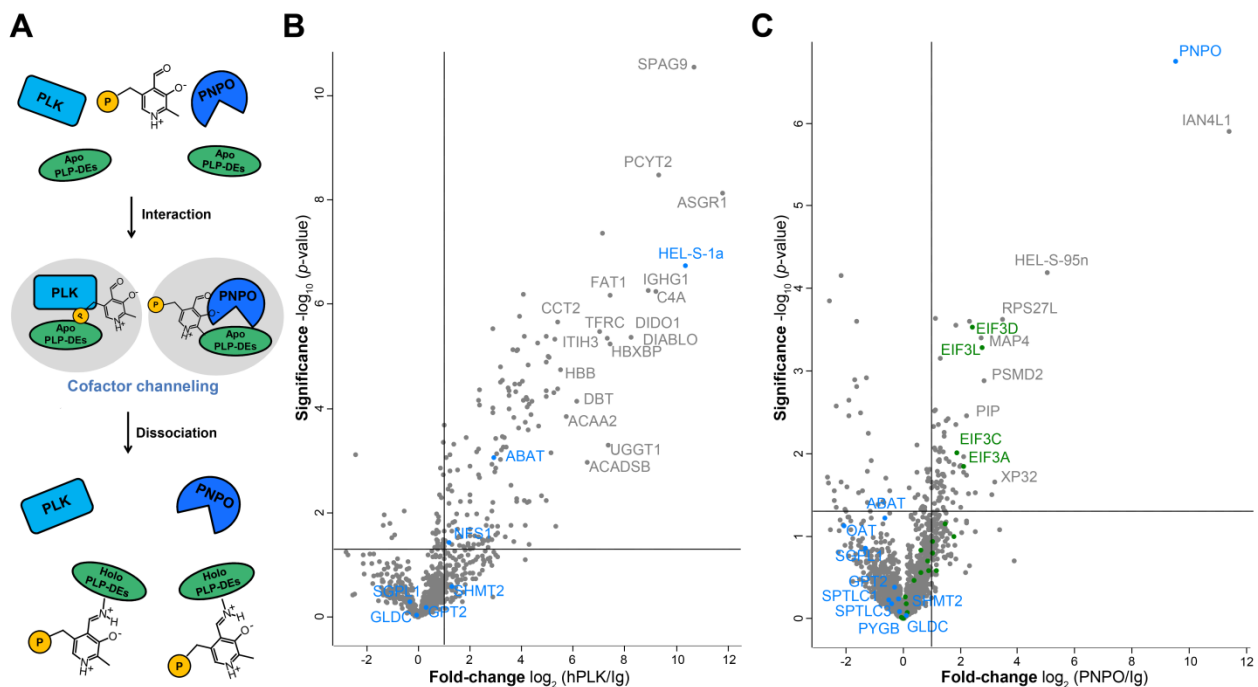
Figure II-30 Cellular binding partners of hPLK and PNPO. (A) Channeling mechanism proposed for **PLP** transfer to apo-B6-enzymes. PLK and PNPO form a specific complex with **PLP** and apo-B6-enzymes to mediate the transfer of the cofactor (P = OPO_3^{2-}). Combined cross-link/co-IPs against hPLK (*HEL-S-1a*)

(**B**) and PNPO (**C**) from HepG2. Volcano plots representing *t*-test results of anti-hPLK or anti-PNPO co-IP, respectively, compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). **PLP**-DEs are highlighted in blue, and components of the eukaryotic translation initiation factor 3 are colored green for the PNPO co-IP.

Albeit free **PLP** can react with apo-B6-enzymes, the intracellular concentrations are too low to meet the requirements of many **PLP**-DEs,[274] which raises the question how the cofactor is transported to newly synthesized B6-enzymes. One possible proposed mechanism postulates an interaction of the salvage enzymes PLK and PNPO with apo-B6-enzymes to directly channel the cofactor to its acceptors (**Figure II-30A**).[275–277] A channeling mechanism requires that both PLK and PNPO form specific interactions with B6-enzymes inside the cell. To address this hypothesis, we performed combined cross-link/co-IP with the DSSO cross-linker[110] from HepG2 to investigate the set of cellular binding partners of kinase and oxidase (**Figure II-30B** and **C**). Though we identified several interactors, we detected no significant overrepresentation of **PLP**-DEs neither for the hPLK nor for the PNPO pull-down.

In order to trigger a possible interaction with **PLP**-DEs we incubated HepG2 with a 20-fold excess of **PN** (100 µM) compared to standard medium concentration overnight preceding cross-link/co-IP against hPLK and PNPO (**Figure II-31A** and **C**). In an additional experiment cells were incubated with 200 µM **PL** (40-fold excess) for 2 h previous to pull-down of hPLK, as **PL** is a primary substrate of the kinase as well (**Figure II-31B**). Though the cells were supplied with a huge amount of cofactor we did not observe enhanced interaction with **PLP**-DEs under no condition examined. However, for PNPO we detected a markedly increase in subunits of eukaryotic translation initiation factor 3 (ETIF3) upon **PN** stimulation. A gene ontology analysis[278,279] of enriched proteins revealed a significant overrepresentation of the biological process *translation initiation* with a corrected *p*-value of 3.2158E-13 (8% gene cluster frequency). Although a STRING analysis[280] of PNPO and subunits of ETIF3 did not result in any functional association, a connection to ETIF3 displays an interesting observation which may inspire further investigations.
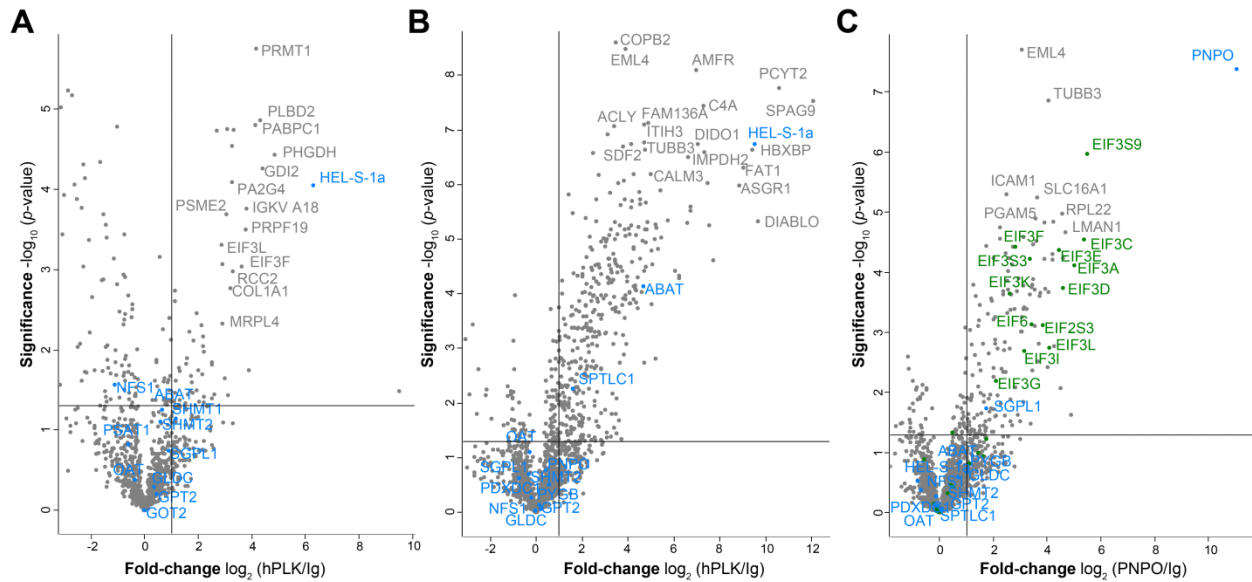
Figure II-31 Cellular binding partners of hPLK and PNPO under excess concentrations of cofactor. Combined cross-link/co-IPs against hPLK (*HEL-S-1a*) after stimulation with 100 µM **PN** (**A**), and 200 µM **PL** (**B**), as well as against PNPO after stimulation with 100 µM **PN** (**C**) from HepG2. Volcano plots representing *t*-test results of anti-hPLK or anti-PNPO co-IP, respectively, compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of log_2 = 1 (2-fold enrichment) and -log_{10} (*p*-value) of 1.3 (solid lines). **PLP**-DEs are highlighted in blue, and components of the eukaryotic translation initiation factor 3 are colored green for the PNPO co-IP.

Taken together, our combined cross-link/co-IP experiments against **PLP** salvage enzymes hPLK and PNPO did not result in a significantly overrepresented interaction with **PLP**-DEs neither under normal cellular conditions nor upon cofactor-triggering. These results do not support the hypothesis of a direct interaction of kinase and oxidase with B6-proteins to channel the cofactor to the apo-enzymes, at least not under the experimental conditions examined.

8 Characterization of PLPBP

PLPBP was significantly enriched upon labeling with **PL1** and **PL2** in all cell lines and under all conditions studied. Moreover, PLPBP displayed the highest fold-changes in most labeling experiments and could be an essential **PLP**-binding protein, as discussed above (cf. Part II, 3.4). Human *PLPBP* was first isolated in 1999 and demonstrated to be ubiquitously expressed in all tissues. Although PLPBP is highly conserved throughout evolution with orthologs occurring across all domains of life including plants and bacteria,[281] its exact cellular function is not defined yet. PLPBP was shown to bind **PLP**, but no enzymatic activity towards any of the 20 proteinogenic amino acids and their corresponding D-enantiomers was detected, though the main substrate class of **PLP**-DEs are amino acids.[282] Several authors have postulated an involvement in **PLP** homeostasis as knockout of PLPBP orthologs leads to increased pyridoxine toxicity in *E. coli*[283] and cyanobacteria[284], and mutations in human *PLPBP* are causing vitamin B6-dependent epilepsy (VitB6-EP).[285] Biochemical mechanisms associated with VitB6-EP caused by mutations in other genes are inactivation of **PLP** by small accumulating compounds or impaired **PLP** formation or transport.[286] However, pathological mechanisms by which *PLPBP* variants cause VitB6-EP remain to be elucidated. Interestingly, hepatocellular carcinoma (HCC)-derived cell lines were shown to have reduced gene expression levels of *PLPBP*. As re-expression inhibited cell colony formation, migration and tumorigenesis, PLPBP was described as tumor suppressor gene for HCC.[287] Because PLPBP is poorly characterized and the only **PLP**-binding protein enriched throughout all labeling experiments, our goal was an examination of structural features as well as cellular roles of this conserved vitamin B6-binding protein in the next step.

8.1 Structural properties of recombinant wild-type human PLPBP

First, we expressed and purified recombinant human PLPBP to investigate its oligomeric and spectral properties. Therefore, we performed affinity chromatography against Strep-tagged PLPBP to isolate the protein from Rosetta 2 (DE3) cells, followed by a TEV-digest to remove the affinity tag and subsequent SEC. Human PLPBP is mainly dimeric with a residual amount of monomer even under reductive conditions indicating that dimerization is not caused by a disulfide bridge (**Figure II-32A**). Notably, the Strep-tagged version of the protein displayed the same oligomeric properties. This was in contrast to previous observations postulating that the biggest amounts of recombinant human PLPBP count for the monomer with only a minor fraction of dimer.[288] Separation, storage, and reinjection of both species into an analytical size exclusion column revealed that monomer and dimer are stable and not in rapid equilibrium with each other (**Figure II-32B** and **C**). Oligomerization was not dependent on **PLP**, as incubation with excess **PLP** did not lead to dimerization of the protein or *vice versa* (**Figure II-32D** and **E**). We further confirmed that both species have **PLP** bound by high-resolution MS (TEV-digested, NaBH₄-reduced proteins; MW = 30,480 Da) and UV/Vis-spectroscopy, where both species display the characteristic internal aldimine peak (**Figure II-32F**). In addition, we determined the **PLP** binding site (Lys47) by LC-MS/MS analysis (**Figure II-32G**). Notably, as the yeast (PDB: 1B54)[289] and several other bacterial PLPBP orthologs like the one from *E. coli* (*yggS*) (PDB: 1W8G) were shown to be monomeric proteins, a predominant dimeric species displays a special property of human PLPBP, which might indicate an additional function. This is also supported by the fact that some of the lysines involved in cross-linking contacts (see below) were missing in several other orthologs (K81, 125, and 133).[290] However, a monomeric state displays a noteworthy feature as the majority of **PLP**-DEs are dimeric, shielding the cofactor by the other subunit.[291] In contrast to that, **PLP** is solvent exposed in monomeric PLPBP orthologs like *E. coli* YggS.[291] The exposed **PLP** could facilitate its exchange with other molecular partners supporting the hypothesis of PLPBP orthologs being involved in **PLP** homeostasis.[285]

In order to compare solvent accessibility of human dimeric and monomeric PLPBP we incubated both species with increasing concentration of DCS[228] and monitored the absorbance decay at a wavelength of 425 nm as DCS is a **PLP**-binding antibiotic (**Figure II-32H**).[292] Internal aldimine absorbance of the monomer was almost completely lost upon incubation with 10 mM DCS. Contrarily, dimer absorbance disappeared not until 50 mM DCS, indicating a reduced solvent accessibility compared to the monomer. In addition, the peak intensities resulting from the enolimine (λ_{max} ~ 336 nm) and ketoenamine (λ_{max} ~ 425 nm) forms of the internal aldimine were

different in monomer and dimer (**Figure II-32F**). In line with DCS titration, these differences suggest a change in solvent exposure of **PLP** in dimeric PLPBP compared to the monomer, as the enolimine form is favored in a non-polar environment.[293]
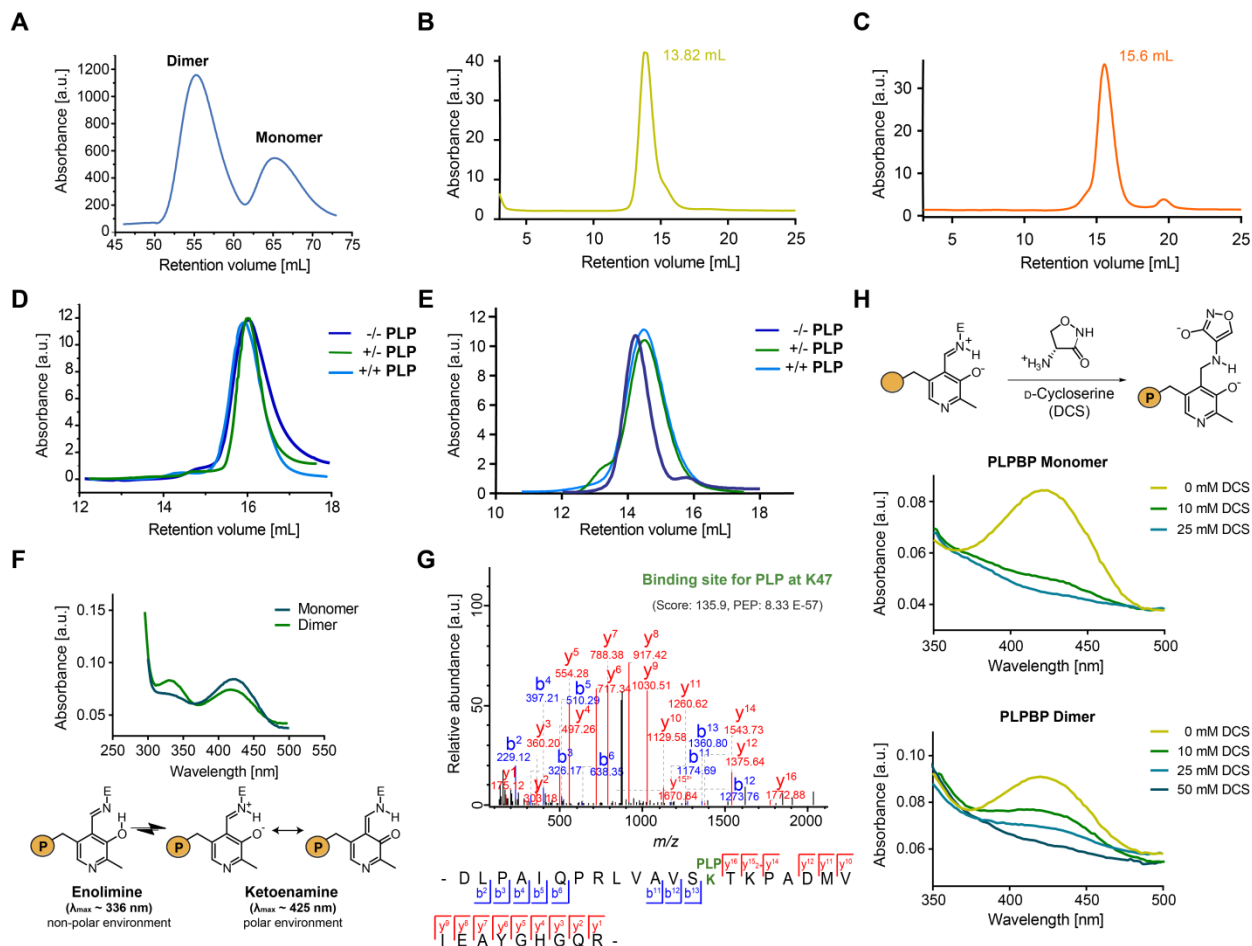
Figure II-32 Properties of recombinant wild-type human PLPBP. (**A**) The main oligomeric state of PLPBP is dimeric, with a small fraction of monomer according to SEC (preparative scale). Separation, storage, and reinjection of dimer (**B**) and monomer (**C**) (analytical scale). Oligomerization of PLPBP is not dependent on **PLP**, as pre-incubation of monomeric (**D**) or dimeric (**E**) PLPBP with excess **PLP** and separation in buffer without **PLP** (+/-), as well as separation in buffer with **PLP** (+/+) did not lead to differences in retention time compared to PLPBP not pre-incubated with excess **PLP** and separated in the absence of **PLP** (-/-; analytical scale). (**F**) UV/Vis-spectra of PLPBP monomer and dimer displayed characteristic absorbance peaks corresponding to the internal aldimine. Peak intensities resulting from the enolimine (λ_{max} ~ 336 nm) and ketoenamine (λ_{max} ~ 425 nm) forms of the internal aldimine differed in monomer and dimer (P = OPO_3^{2-}).[293] (**G**) MS/MS spectra identifying the **PLP**-binding site of PLPBP (K47) with corresponding score and posterior error probability (PEP). (**H**) Incubation with 10 mM of the **PLP**-binding antibiotic DCS led to almost complete displacement of **PLP** from monomeric species whereas internal aldimine absorbance of PLPBP dimer disappeared not until 50 mM DCS (P = OPO_3^{2-}).[292]

Further, we performed XL-MS using the MS-cleavable cross-linker DSSO[58] to monitor the dimerization interface of PLPBP. The dimer could be covalently linked upon DSSO addition (50-fold excess over PLPBP, reaction time: 1 h at 37 °C) as determined by SDS-PAGE (**Figure II-33A**). Cross-linked dimer and monomer fractions were separated by SEC (**Figure II-33B**) prior to reduction, alkylation, and digestion. Cross-linked peptides were measured by LC-MS/MS and data were analyzed using XlinkX (**Table S10**).[18,110] As no crystal structure of human PLPBP is available, a model was generated using the SWISS-MODEL platform and the ortholog from yeast (PDB: 1B54)[289] as template.[191,192] Cross-links solely occurring in the dimer fraction served as restraints for generation of a PLPBP dimerization model with the HADDOCK software (**Figure II-33C**, left).[19] Cross-links were mapped onto PLPBP dimer and displayed all to be within the maximum range of DSSO (35 Å; **Figure II-33C**, right).[39] In line with spectroscopic data, the PLPBP model generated with cross-link restraints displayed the dimerization interface around the **PLP** binding site K47, suggesting that the cofactor is shielded from the surrounding upon PLPBP oligomerization.

Cross-links in PLPBP dimer	Inter-chain cross-link distance (Å)
K47 - K133	21.45
K47 - K81	12.37
K49 - K81	9.15
K49 - K125	32.85
K49 - K133	23.72
K104 - K133	5.52
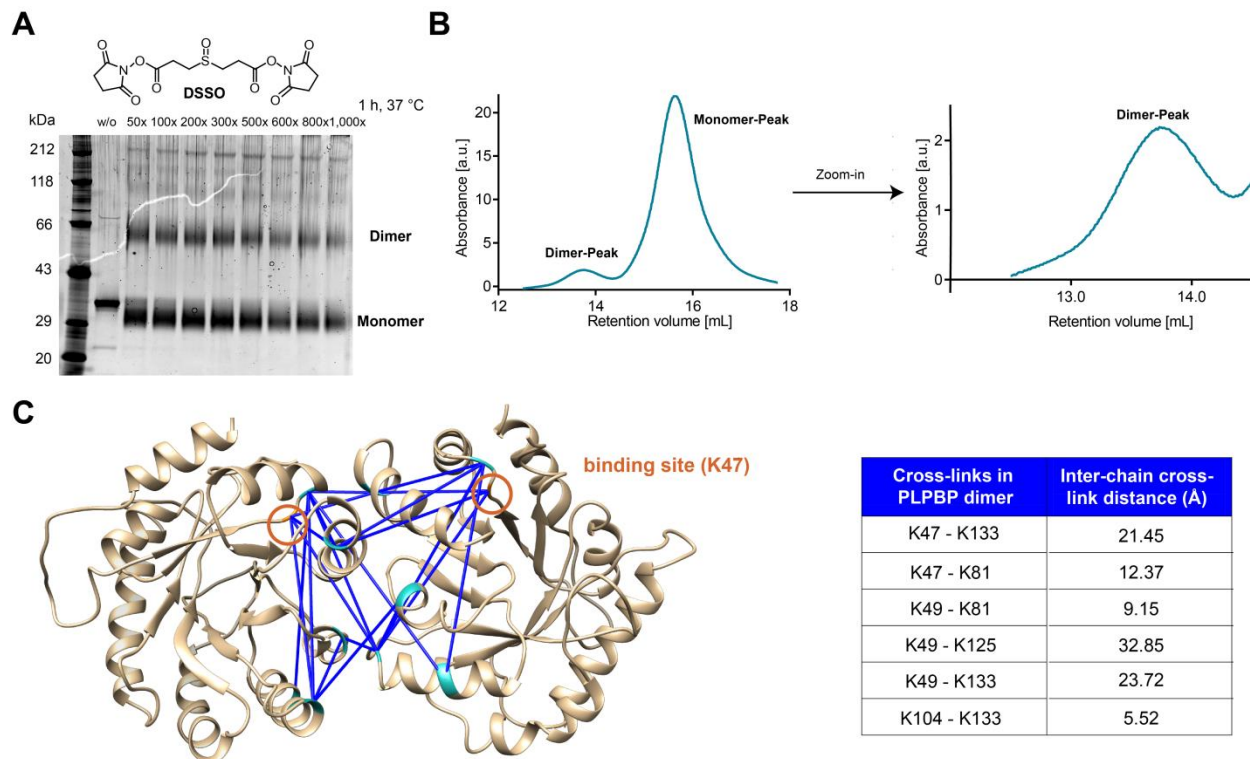
Figure II-33 Analysis of PLPBP dimerization interface with XL-MS. (**A**) Cross-linking of PLPBP dimer with DSSO revealed a sufficient covalent fixation of dimer with a 50-fold excess of linker and a reaction time of 1 h at 37 °C. (**B**) SEC after DSSO cross-linking of the PLPBP dimer (exemplarily presented for 50-fold excess) for 1 h at 37 °C. A sufficient amount of dimer could be purified (zoom-in) for MS sample preparation. The curve looked similar for incubation with 100-fold DSSO excess. Cross-links solely

occurring in the dimer samples (**C**, **right**) served as restraints for molecular docking of two PLPBP monomers (derived from SWISS-MODEL[191] using the yeast ortholog as template; PDB: 1B54[289]) using HADDOCK[19] to generate a dimer model (**C**, **left**). Cross-links displayed on the dimer were all within the maximum distance threshold of DSSO (35 Å). Lysines involved in cross-linking contacts are highlighted cyan, the **PLP** binding site in orange.

8.2 Structural properties of recombinant human PLPBP mutants

VitB6-EP was first described in 1954.[294] In classical VitB6-EP seizures are observed within the first months of life, even within pregnancy, including abnormal foetal movements, severly abnormal EEG, or hypoxic ischaemic encephalophaty.[295] Later on, characteristics as speech and motor delay, as well as learning difficulties may result.[285] Patients are mostly resistant to antiepileptic drugs, but seizures are usually controlled by oral uptake of vitamin B6 exceeding the daily demand of 1.2-1.4 mg for adults.[295,296] Today five genetic variants are known to cause VitB6-EP. Two of them result in the inactivation of **PLP**, including **PN**-dependent epilepsy and hyperprolinemia type II, and two have been attributed to the reduced synthesis/availability of **PLP** (**PLP**-dependent epilepsy and congenital hypophosphatasia).[296] **PN**-dependent epilepsy is caused by mutations in the antiquitin gene (*ALDH7A1*) leading to accumulation of L-Δ^1-piperideine 6-carboxylate (P6C).[297] Hyperprolinemia type II is the result of an inborn error in the *ALDH4A1* gene encoding L-Δ^1-pyrroline 5-carboxylate (P5C) dehydrogenase which leads to accumulation of P5C.[298] Both metabolites were shown to inactivate **PLP** by a Knoevenagel condensation.[297,298] On the other hand, **PLP**-dependent epilepsy is caused by a mutation in the *PNPO* gene, leading to impaired production of **PLP** out of its reduced precursors. Finally, mutations affecting tissue nonspecific alkaline phosphatase (*TNSALP*), which is involved in the uptake of **PLP** at the plasma membrane of cells, are causing congenital hypophosphatasia.[296] The fifth gene associated with VitB6-EP is *PLPBP* where -as mentioned above- the molecular reasons for the clinical phenotypes have not been elucidated.

To date, 20 pathogenic variants in a total of 27 patients have been identified.[285,290,299,300] Truncated PLPBP variants lead to the most severe syndroms of the disease due to complete absence of the protein from patient cells.[290] Six of the pathogenic PLPBP mutants have been recombinantly accessed and displayed altered thermal stability and solubility.[288] However, as our results concerning oligomerization behavior and **PLP** exposition of human wild-type PLPBP differed from previous results,[288] we attempted to re-analyze biochemical properties of two previously expressed mutants (Y69C and R241Q) as well as to examine novel variants (R41Q, V45D, and E67K) (**Figure II-34A**). Notably, all mutants except the E67K could be successfully

obtained from the *E. coli* expression system although with a markedly reduced yield than wild-type PLPBP (**Figure II-34B**). In the case of E67K the clinical phenotype was the most severe amongst all mutants examined in this thesis,[285,290,299,300] indicating major disturbances in protein function and folding in line with permitted bacterial expression. Moreover, E67 is highly conserved across PLPBP orthologs.[288,290] To evaluate expression efficiency of PLPBP mutants in human cells we transfected HEK293 with expression vectors coding for wild-type and pathogenic FLAG-tagged PLPBP under the control of the constitutive human cytomegalovirus (CMV) promotor. In line with observations from the bacterial expression system the E67K variant displayed lowest protein levels, whereas the mutant with the best expression was Y69C (**Figure II-34C**).

In line with UV/Vis spectra (**Figure II-34D**), MS analysis of $NaBH_4$-reduced proteins revealed that besides the wild-type also the Y69C, as well as the R41Q mutant are completely saturated with **PLP** after purification. In contrast, R241Q and V45D do not harbor any cofactor, which confirms their involvement in **PLP** binding suggested by modeling studies (**Figure II-34B**).[290] Notably, absorbance measurements with the V45D mutant were permitted due to low yield and final concentration of the protein (50 µL of a 6.5 µM solution from 2 L expression culture). Incubation with a 4-fold molar excess of **PLP** previous to reduction and MS analysis, revealed that both V45D, as well as R241Q are indeed able to still bind the cofactor (**Figure II-34B**). Interestingly, **PLP**-bound R241Q mutant was detected at a 2-fold excess over its apo-form (determined *via* relative abundance of MS signal), whereas V45D was still almost exclusively present in its cofactor-free state. These results indicate a more severe impact of the V45D substitution on **PLP** binding, most likely caused by the introduction of a negative charge in the phosphate binding pocket. Further, this might explain the worse clinical outcome of patients carrying the V45D mutation compared to the milder phenotype caused by the R241Q variant.

We further compared thermal stability of PLPBP mutants and wild-type in the presence and absence of **PLP** (**Figure II-34E**). In line with MS and UV/Vis spectra, mutants lacking **PLP** after the purification displayed overall lower metling temperatures in the range of 40 °C. R41Q has reduced stability (49.5 °C) compared to the wild-type (56.1 °C), whereas the Y69C substitution resulted in a slight increase in melting temperature (59.2 °C).[288] Stability of PLPBP variants matched corresponding expression levels in human cells (**Figure II-34C**).

A

B

PLPBP variant	MW (Da, not reduced)	MW (Da, reduced)	MW (Da) w/ 4x PLP incubated (reduced)	Yield [mg]	Severity of mutation
WT	33,452 (- **PLP**)	33,684 (+ **PLP**)	33,684 (+ **PLP**)	3.82	/
Y69C	33,393 (- **PLP**)	33,624 (+ **PLP**)	33,624 (+ **PLP**)	1.32	Moderate
R41Q	33,425 (- **PLP**)	33,656 (+**PLP**)	33,656 (+ **PLP**)	0.14	Mild
R241Q	33,425 (- **PLP**)	33,425 (- **PLP**)	33,425 and 33,656 (1:2 ratio; - and + **PLP**)	0.34	Mild
V45D	33,469 (- **PLP**)	33,469 (- **PLP**)	33,469 and 33,700 (20:1 ratio, - and + **PLP**)	0.05	Moderate

C **D** **E**
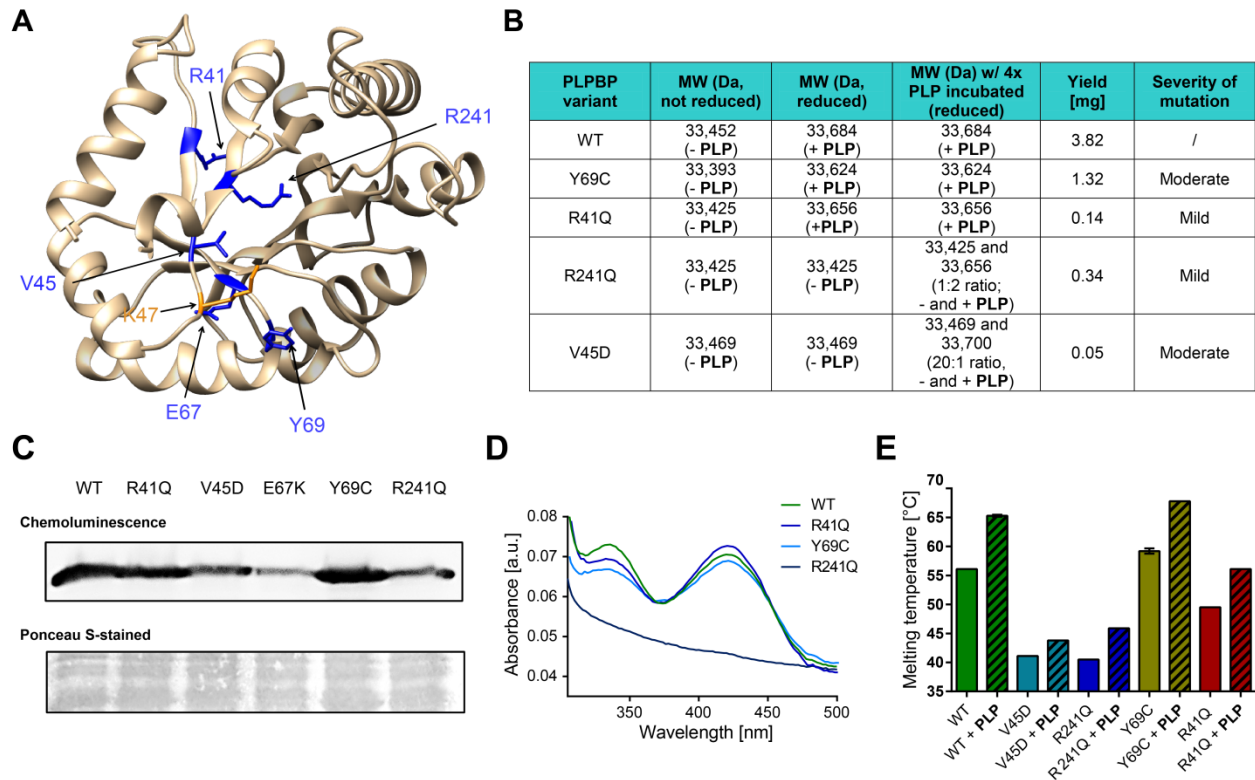
Figure II-34 Stability and PLP-binding capacity of PLPBP mutants. (**A**) Localization of mutated residues presented on the human wild-type PLPBP model structure (blue). The **PLP**-binding site (K47) is colored orange. (**B**) Intact protein MS of wild-type (WT) and mutants (Strep-tagged) in their not reduced form, in their NaBH₄-reduced forms, as well as in their reduced forms after incubation with a 4-fold molar excess of **PLP** for 30 min at r.t.. For R241Q and V45D the relative abundance of the species is given. In addition, protein yields from a 2 L epression culture, as well as clinical severity of mutations are given.[285,290,299,300] (**C**) Western blot analysis of transfection of WT and PLPBP mutants into HEK293 cells. (**D**) UV/Vis spectra of PLPBP WT and mutants. (**E**) Thermal stability of PLPBP WT and mutants in absence or presence of **PLP** (20-fold molar excess; *n*=3).

In the next step, we evaluated oligomerization behavior of the PLPBP mutants under reductive conditions (**Figure II-35**). Interestingly, all variants except the R41Q substitution are exclusively monomer, with a slight residual amount of dimer for the R241Q mutant. As dimeric R41Q as well as monomeric R241Q both display a milder outcome of VitB6-EP,[285,290,299,300] oligomerization behavior does not affect the overall clinical severity but might contribute to differences in individual symptoms. Notably, even under non-reductive conditions the Y69C mutant was not present as dimer contrarily to previous postulations (data not shown).[288] Alterations on oligomerization behavior might be explained by the relative localization of affected amino acids to regions involved in dimerization (**Figure II-35F**). Interestingly, R41 is located most distant towards lysine residues involved in cross-linking contacts within the PLPBP dimer (**Figure II-**

33C) and its mutation might be therefore not disrupting the oligomeric state. Moreover, the R41Q mutant was exclusively present as dimer, indicating that this substitution might even enhance oligomerization. However, cellular consequences behind different dimerization properties of the variants remain to be determined.
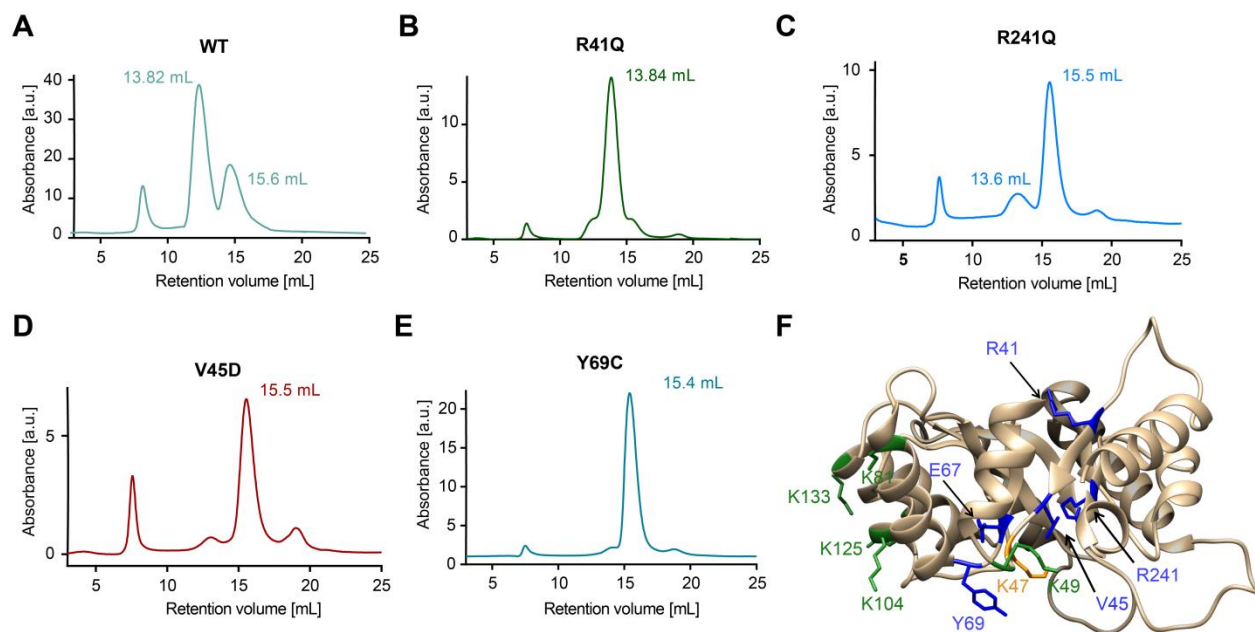
Figure II-35 Impact of pathogenic mutations on PLPBP oligomerization. Analytical size-exclusion chromatograms of wild-type (WT) (**A**), R41Q (**B**), R241Q (**C**), V45D (**D**), and Y69C (**E**) mutants under reductive conditions. (**F**) Position of pathogenic mutations (blue) with respect to binding site (orange), as well as lysine residues involved in cross-linking contacts (green) presented on the human WT PLPBP model structure.

8.3 Insights into the cellular role of human PLPBP

8.3.1 Localization of human PLPBP within the cell

We proceeded with investigation of the cellular localization of PLPBP. Thereby, we performed immunochemical staining of the target protein in a genetically unmodified hepatocellular carcinoma-derived cell line (HepG2) after fixation of cells. We deliberately selected a liver cancer-derived cell line as the liver is the most active organ in pyridoxine metabolism.[301–304] PLPBP turned out to be localized in both cytosol and mitochondria, as we observed colocalization with the mitochondrial marker TOM20 (**Figure II-36**). This was in line with previous observations performing immunohistochemical staining in cells overexpressing PLPBP.[290]

KEY: TOM20 = mitochondria   DAPI = nuclei   ▷ ⋯⋰⋱ regions of PLPBP and TOM20 colocalization
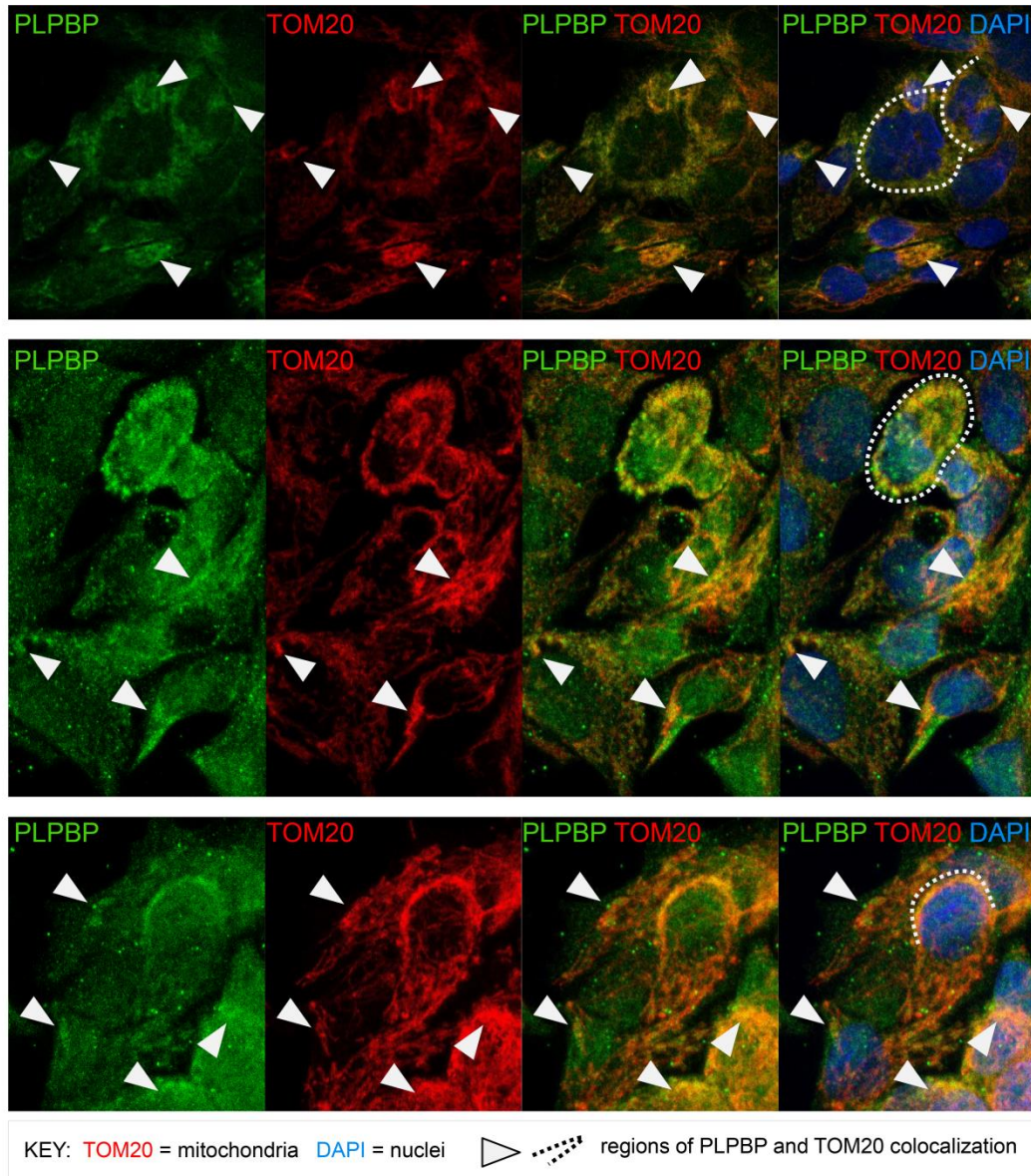
Figure II-36 PLPBP is localized in the cytosol and mitochondria. Immunohistochemistry for PLPBP (green) displays the protein to be localized on the one hand in the cytosol and on the other hand in mitochondria, as we observed colocalization with the mitochondrial marker (TOM20, red). Stainings were performed on cells fixed on coverslips and nuclei were counterstained with DAPI.

8.3.2 Cellular binding partners of human PLPBP

We further investigated the subset of cellular binding partners of PLPBP. Therefore, we treated HepG2 with cell-permeable DSSO cross-linker and performed co-IP against PLPBP (**Figure II-37A**).[110] More than half (32 out of 55) of the significantly enriched interactors of PLPBP are

mitochondrial proteins (GO term 0005739: mitochondrion[279]), supporting the results from our immunochemical staining.

A

B

Annotated pathway (KEGG)	Relevant genes
Tricarboxylic acid cycle (TCA cycle)	ACO2, MDH2, CS, PC, IDH1
Pyruvate metabolism	ALDH2, ACAT1, MDH2, ALDH1B1, PC
Glyoxylate and dicarboxylate metabolism	ACO2, ACAT1, MDH2, SHMT2, CS
Biosynthesis of amino acids	ACO2, SHMT2, CS, PC, PYCR1, IDH1
Valine, leucine, and isoleucine degradation	ALDH2, ACAT1, ACAA2, ECHS1, ALDH1B1, ABAT
Fatty acid degradation	ALDH2, ACAT1, ACAA2, ECHS1, ALDH1B1

Figure II-37 Identification of cellular binding partners of PLPBP. (**A**) A combined cross-link/co-IP against human PLPBP from HepG2 revealed significant enrichment of several mitochondrial proteins. Volcano plot representing *t*-test results of anti-PLPBP co-IP compared to the isotype control co-IP (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). Mitochondrial proteins are highlighted in green, **PLP**-DEs in blue, citric acid cycle associated proteins in red, and mitochondrial PYCR1 is colored orange. In the case that a mitochondrial protein was reported as a **PLP**-DE or citric acid cycle associated protein it was colored blue or red, respectively. (**B**) STRING analysis[280] of enriched proteins. Significantly overrepresented KEGG pathways[305] are presented together with involved genes.

Interestingly, we detected a direct interaction with human mitochondrial PYCR (*PYCR1*) and human mitochondrial SHMT (*SHMT2*). A gene clustering analysis revealed that *PLPBP* orthologs are in close proximity to L-Δ^1-pyrroline-5-carboxylate reductase (PYCR) genes with a high prevalence of 20% within the genomes analyzed (11,411 organisms).[283] PYCR orthologs reduce L-Δ^1-pyrroline-5-carboxylate to generate proline.[306] As genes involved in a common pathway or function are frequently found near each other on the chromosome,[307,308] gene clustering often reveals a functional link. Further, clustering of *PLPBP* was also observed with SHMT orthologs with a prevalence of 0.5%.[283] Taking into account that *PYCR1*, *PLPBP*, and *SHMT2* are located on three different chromosomes in the human genome, a physiological interaction could -in addition to previous gene clustering results- also point to a functional link. However, we did not -at least under our experimental conditions- find a prevalence of **PLP**-DEs or vitamin B6 metabolizing enzymes (GO term 0030170: **PLP**-binding[279]) amongst the interaction partners which could be expected if PLPBP would be involved in **PLP**-transfer to other **PLP**-DEs as postulated previously.[285] A STRING analysis revealed overrepresentation of proteins involved in pyruvate and glyoxylate metabolism, as well as degradation of fatty acids (**Figure II-37B**).[280]
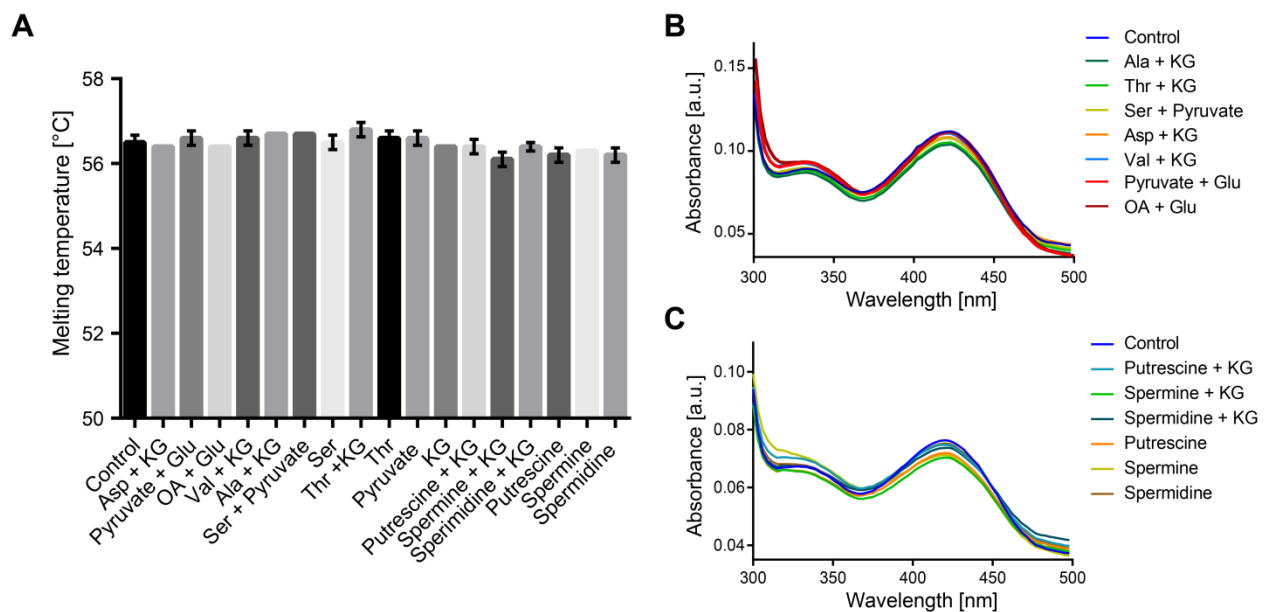
Figure II-38 Transaminase activity of PLPBP. (**A**) Thermal stability of monomeric PLPBP in the presence of several metabolites. UV/Vis spectra of monomeric PLPBP after 5 min incubation at 37 °C with 20-fold molar excess of amino acids (**B**) and polyamines (**C**; KG = α-ketoglutaric acid, OA = oxaloacetic acid).

In addition, we observed a high number of TCA cycle associated enzymes (highlighted in red in **Figure II-37A**). This was an interesting observation as yeast deficient of the PLPBP ortholog display a disturbed TCA cycle,[290] suggesting a connection of PLPBP to proper function of mitochondrial metabolism. As several intermediates of the TCA cycle are α-ketoacids like α-ketoglutaric acid we hypothesized that PLPBP might function as a transaminase that may act on metabolites of the cycle. However, incubation of PLPBP dimer and monomer with several common transaminase educts as well as polyamines like putrescine did not result in any spectral[309] or stability changes (**Figure II-38** and **S9**), suggesting that the protein is not acting as an aminotransferase.

Interestingly, highly enriched mitochondrial L-malate dehydrogenase (*MDH2*), catalyzing the NAD+-dependent oxidation of L-malate to oxaloacetic acid as part of the TCA cycle,[310] as well as cytosolic isocitrate dehydrogenase (*IDH1*), catalyzing the NADP+-dependent oxidation of isocitric acid to α-ketoglutaric acid,[311] were shown to be reversible inactivated by high molar excess of **PLP** (around 1,000-fold).[312–316] Therefore, we hypothesized that PLPBP might have an impact on the activity of these high-confident interactors for instance as a **PLP**-donor. However, neither IDH nor MDH showed altered turnover in the presence of PLPBP monomer and dimer (10-fold molar excess), suggesting that those are not directly regulated by PLPBP, at least not under the conditions examined (**Figure II-39**).

A

Isocitric acid + NAPD+ $\xrightarrow{\textbf{Isocitric acid DH}}$ α-ketoglutaric acid + CO_2 + NADPH/H+

B

Oxaloacetic acid + NADH/H+ $\xrightarrow{\textbf{L-Malate DH}}$ L-Malate + NAD+
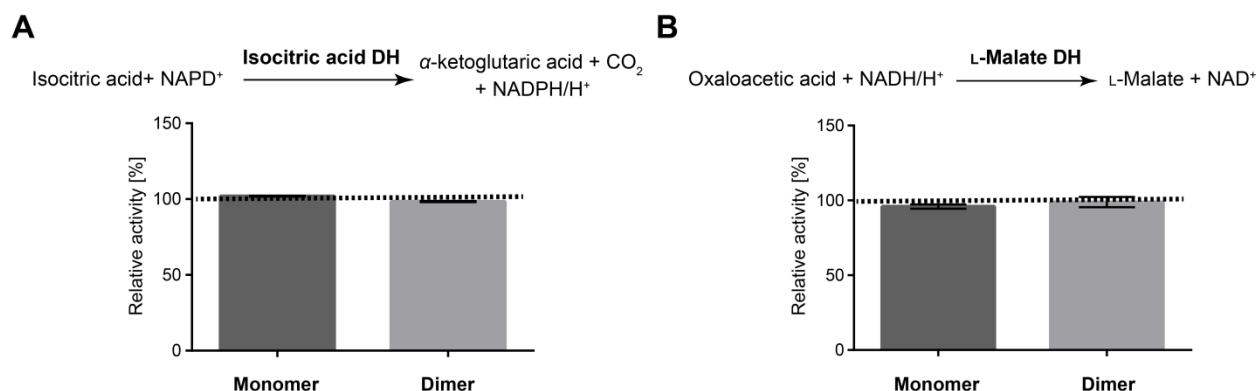
Figure II-39 Influence of PLPBP on activity of citric acid cycle associated proteins. Enzymatic activities of isocitric acid dehydrogenase (DH) (**A**) and L-malate DH (**B**) in the presence of 10-fold molar excess of monomer and dimer compared to the control (*n*=3). Proteins were preincubated with PLPBP 30 min before initiating the assay.

8.3.3 Proteomic changes upon knockdown of human PLPBP

Further, as individuals with VitB6-EP display reduced PLPBP levels within the cell,[285] we examined proteomic changes caused by PLPBP deficiency in HepG2 cells treated with PLPBP-targeting (+)-siRNA compared to a non-targeting (-)-siRNA control. Western blot analysis revealed a sufficient knockdown of PLPBP expression levels after 48 h (**Figure II-40A**). This was corroborated by a full-proteome analysis demonstrating a significant reduction in PLPBP protein levels of (+)-siRNA treated samples compared to the control (**Figure II-40B**).
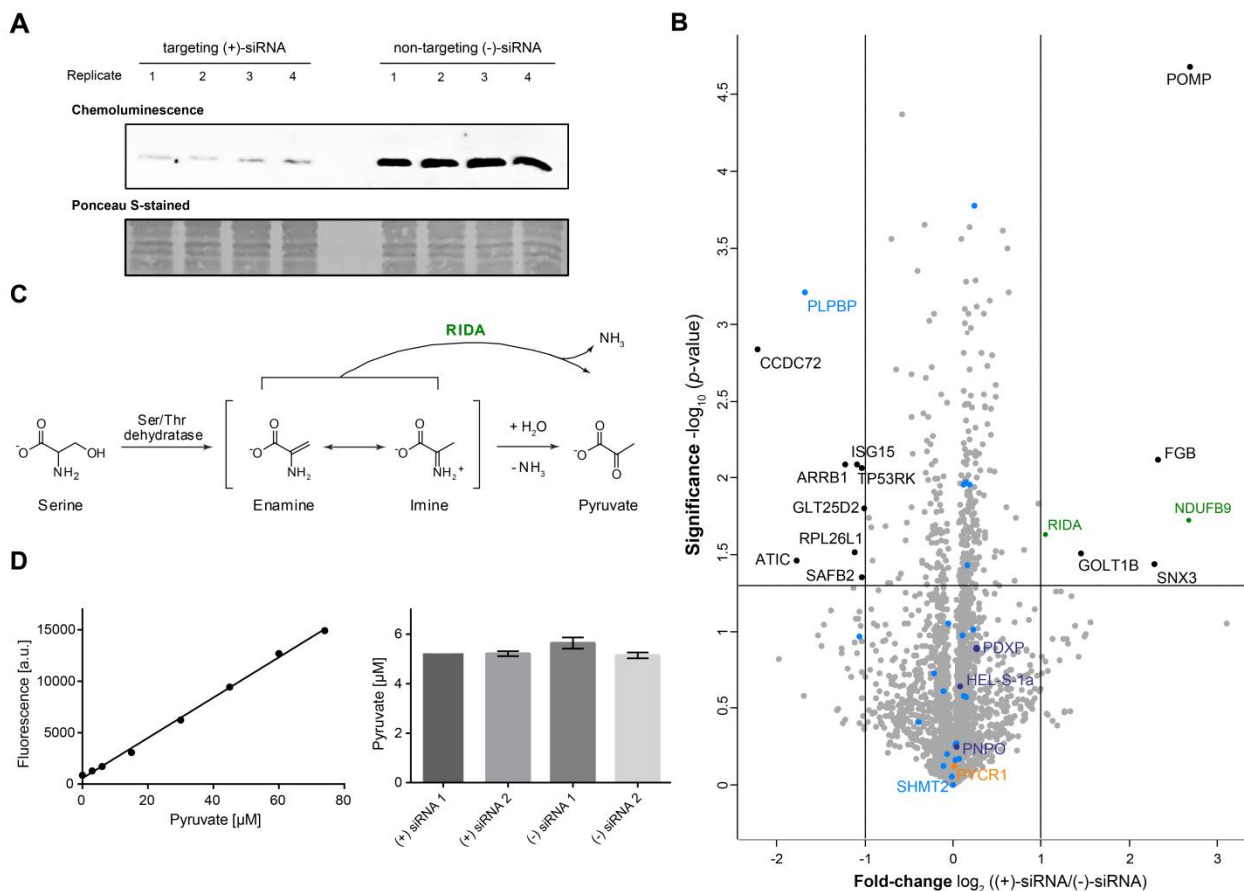
Figure II-40 Knockdown of PLPBP expression levels in HepG2. (**A**) Western blot analysis of siRNA-mediated knockdown of PLPBP in HepG2 reveals sufficient reduction of expression levels after 48 h upon incubation with targeting (+)-siRNA compared to a non-targeting (-)-siRNA control. Ponceau-S stained membrane revealed equal protein amounts for all conditions. (**B**) Full proteome analysis of siRNA-mediated knockdown of PLPBP in HepG2 for 48 h. Volcano plot representing t-test results of (+)-siRNA samples compared to the non-targeting control (n=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) or depletion factor of $\log_2 = -1$ (2-fold depletion), respectively, and -$\log_{10}$ (p-value) of 1.3 (solid lines). Dysregulated mitochondrial proteins are highlighted in green, all **PLP**-DEs in blue, vitamin B6-metabolizing enzymes in purple, and PYCR1 is colored orange.

(**C**) Reactions catalyzed by Ser/Thr dehydratase and 2-iminobutanoate/2-iminopropanoate deaminase (RIDA) exemplarily presented for the substrate serine. (**D**) Determination of intracellular pyruvate levels in HepG2 after incubation with targeting (+)- or non-targeting (-)-siRNA control (right; *n*=2 biological replicates), presented together with the calibration curve (left).

We did not observe any changes in expression levels of **PLP**-DEs and **PLP** salvage enzymes hPLK (*HEL-S-1a*), **PLP**-phosphatase (*PDXP*) or PNPO. Further, expression of the previously identified PLPBP interacting proteins like SHMT2 and PYCR1 remained unaffected as well. However, we detected several other dysregulated proteins which might display links to observed phenotypes of reduced PLPBP levels or to the cellular function of the protein. For example, TP53-regulating kinase (*TP53RK*) expression levels were reduced upon PLPBP knockdown. TP53RK was found to upregulate p53 activity by phosphorylation of p53 Ser-15,[317] which induces growth arrest or apoptosis, both preventing the proliferation or survival of damaged cells.[318] Dysregulation of TP53RK expression levels could provide a link to the proposed role of PLPBP as tumor-suppressor.[287] Further, proteasome maturation protein (*POMP*), an essential factor for proteasome biogenesis and regulation of cellular protein degradation,[319] was found to be significantly upregulated upon PLPBP knockdown. Increased protein degradation might also contribute to changes in cellular amino acid pools as observed for some of the individuals with VitB6-EP[285] and the *E. coli yggS* knockout.[282]

Interestingly, we found 2-iminobutanoate/2-iminopropanoate deaminase (*RIDA*) significantly overexpressed upon *PLPBP* knockdown. RIDA is broadly conserved through all domains of life and was shown to accelerate the release of ammonia from enamine/imine intermediates generated upon threonine and serine dehydration catalyzed by the corresponding **PLP**-DE Ser/Thr dehydratase (**Figure II-40C**).[320] Enhanced levels of RIDA therefore lead to an increased formation of pyruvate and *α*-ketobutyrat, the reaction products of threonine/serine dehydratase. This observation fits to previous reports, were an involvement of the *E. coli* PLPBP ortholog YggS in the regulation of pyruvate and *α*-ketobutyrat was postulated, as knockout strains accumulate the two metabolites.[282] Due to the high conservation of PLPBP and the ability of human PLPBP to completely reverse the *E. coli yggS* knockout phenotype,[282] a role of the human ortholog in regulation of pyruvate and *α*-ketobutyrate levels *via* influencing RIDA expression or activity levels might be possible as well. We therefore examined intracellular pyruvate levels in HepG2 cells after knockdown of PLPBP (**Figure II-40D**). However, although RIDA was overexpressed we did not observe a significant accumulation of pyruvate upon reduction of PLPBP levels. As a knockdown does not result in 100% elimination of the target protein from the cell, the detected changes are probably the most pronounced ones caused by a

reduction of PLPBP expression levels. It could have been possible that knockdown was not sufficient to cause significant changes in pyruvate levels as detected in the *yggS* knockout. On the other hand, it could also be possible that alterations of PLPBP levels do not have a similar impact on pyruvate levels as observed in the case of bacterial orthologs.

8.3.4 Proteomic changes upon overexpression of human PLPBP

To examine the proteomic changes caused by an overexpression of PLPBP, we transfected HepG2 cells with the plasmid encoding for FLAG-tagged PLPBP under the control of human CMV promotor. Western blot analysis revealed a sufficient overexpression of FLAG-tagged PLPBP after 48 h (**Figure II-41A**).
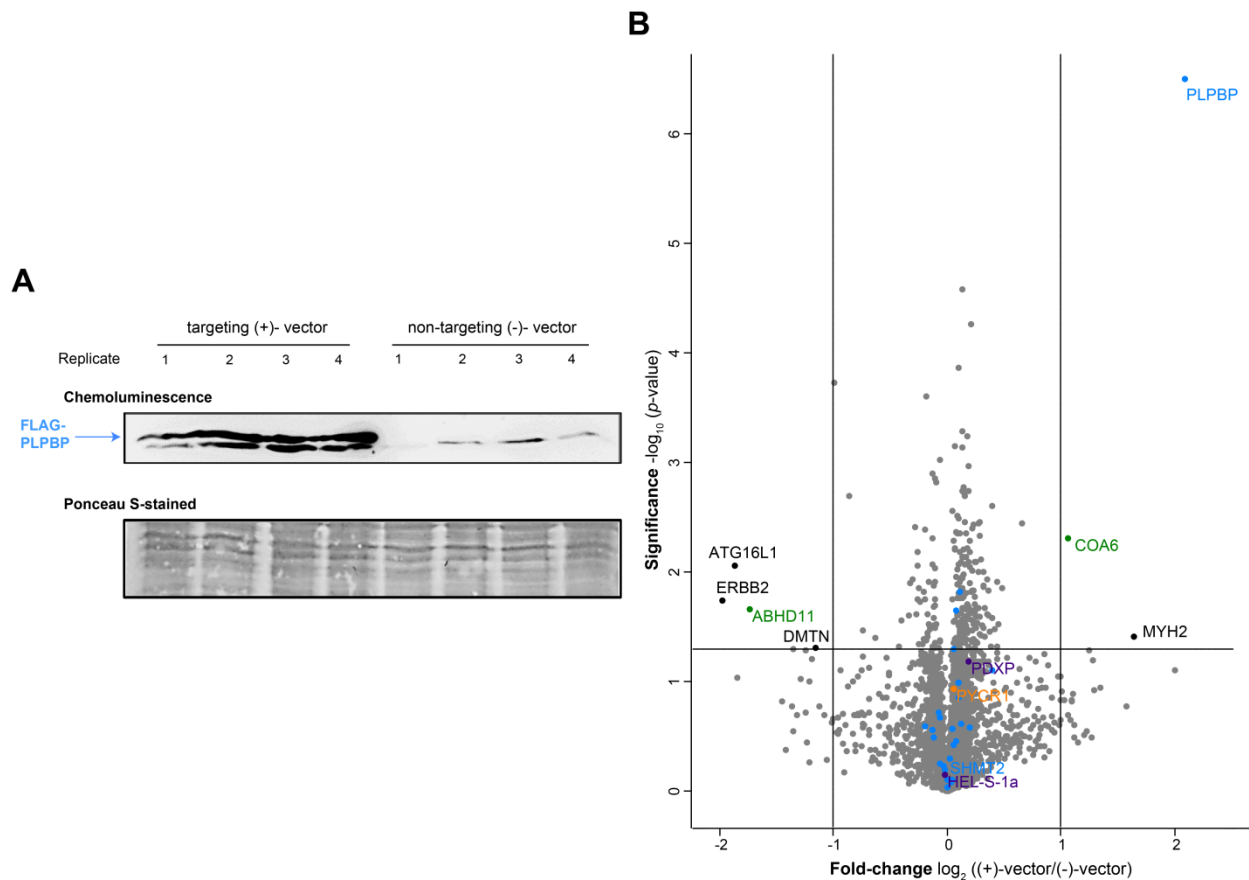
Figure II-41 Overexpression of FLAG-tagged PLPBP in HepG2. (**A**) Western blot analysis of overexpression of FLAG-tagged PLPBP in HepG2 reveals sufficient increase in expression levels after 48 h upon incubation with PLPBP-encoding (+)-vector compared to an empty (-)-vector as a control. Ponceau-S stained membrane revealed equal protein amounts for all conditions. (**B**) Full proteome analysis of PLPBP overexpression in HepG2 for 48 h. Volcano plot representing *t*-test results of (+)-vector samples compared to the empty plasmid control (*n*=4 biological replicates). Cut-off values were defined as

enrichment factor of $\log_2 = 1$ (2-fold enrichment) or depletion factor of $\log_2 = -1$ (2-fold depletion), respectively, and $-\log_{10}$ (p-value) of 1.3 (solid lines). Dysregulated mitochondrial proteins are highlighted in green, all **PLP**-DEs in blue, vitamin B6-metabolizing enzymes in purple, and PYCR1 is colored orange.

In line with western blot results, a subsequent full-proteome analysis demonstrated a significant overexpression of PLPBP compared to cells transfected with the control-vector (**Figure II-41B**). Global proteomic expression changes upon PLPBP overexpression were not that pronounced as upon PLPBP knockdown. However, we detected several proteins dysregulated which are related to cytoskeleton organization and cell morphology. For instance, the motor protein myosin 2 (MYH2) required for muscle contraction is involved in cytoskeleton formation *via* association with actin filaments.[321] Another example is receptor tyrosine-protein erbB-2 (ERBB2), which regulates the outgrowth and stabilization of peripheral microtubules.[322] Finally, dematin (DMTN) is a membrane cytoskeleton-associated protein important for F-actin bundle formation and stabilization.[323] An involvement of PLPBP in the organization of cytoskeleton and muscles could display a possible link to the epileptic phenotypes observed for mutations in the *PLBPP* gene. However, as for PLPBP knockdown analysis, we did not observe changes in expression levels of **PLP**-DEs, as well as **PLP** salvage enzymes.

8.4 Comparison with bacterial phenotypes

Due to the high degree of conservation of PLPBP orthologs and the ability of human PLPBP to completely reverse the *E. coli yggS* knockout phenotype,[282] we examined in a next step proteomic changes caused by a knockout of YggS. As no differences of growth rates for wild-type and $\Delta yggS$ *E. coli* strains were observed,[282] we cultivated both bacterial cultures to an OD_{600nm} of 0.7 previous to proteomic analysis. Several proteins displayed to be dysregulated upon knockout of *yggS*, including several **PLP**-DEs like Ser/Thr dehydratase which was more than 16-fold overexpressed (TdcB; **Figure II-42A**). This is an interesting observation as RIDA which enhances the reaction rates of Ser/Thr dehydratase displayed to be overexpressed upon human PLPBP knockdown (**Figure II-40B**). However, though some **PLP**-DEs were dysregulated upon *yggS* knockout, most **PLP**-DEs as well as vitamin B6 salvage enzymes likes PLK or PNPO were not affected.
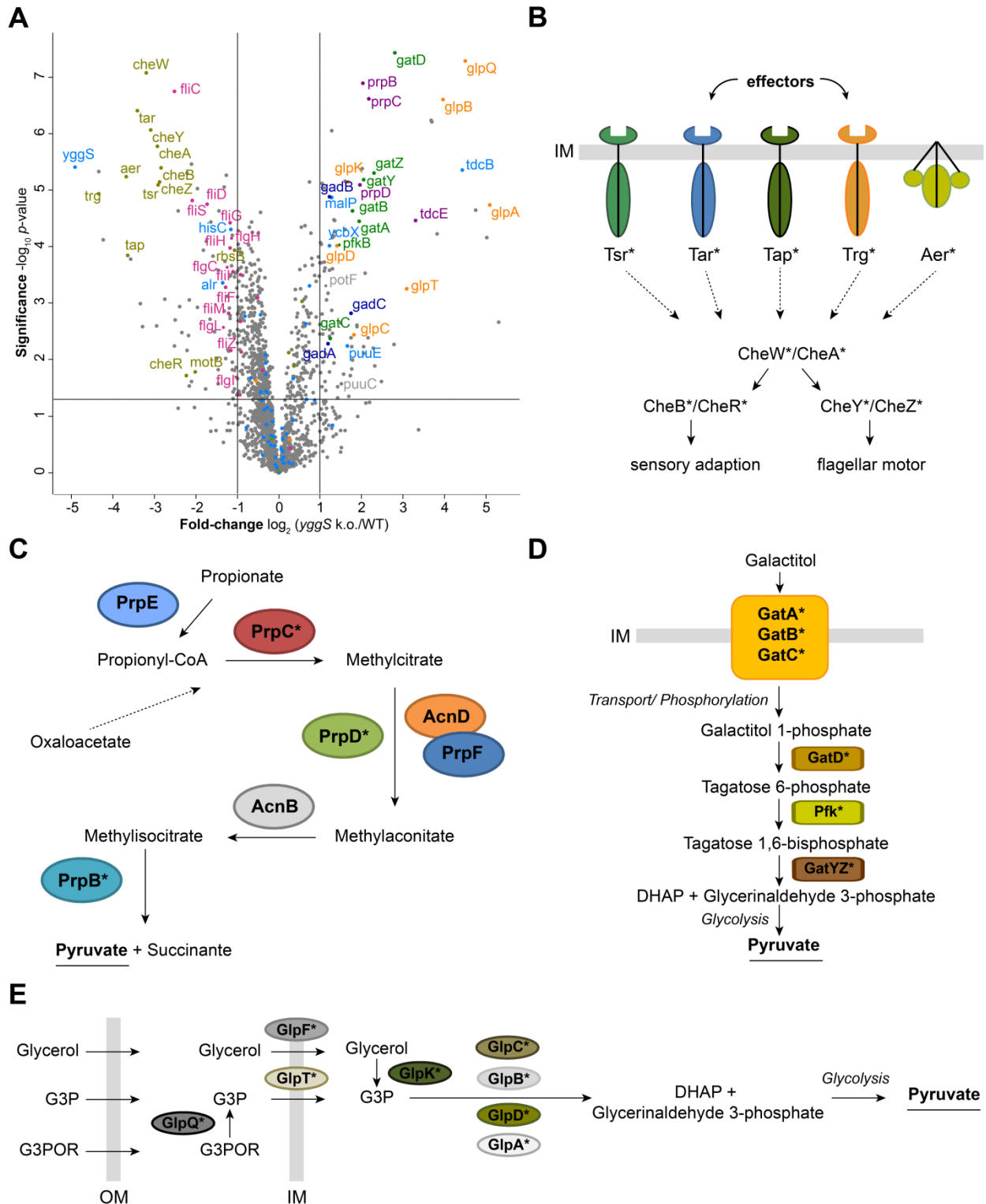
Figure II-42 Proteomic changes upon *yggS* knockout in *E. coli* BW25113. (**A**) Volcano plot representing *t*-test results of *yggS* knockout compared to the wild-type control (*n*=4 biological replicates). Cut-off values were defined as enrichment factor of $\log_2 = 1$ (2-fold enrichment) or depletion factor of $\log_2 = -1$ (2-fold depletion), respectively, and $-\log_{10}$ (*p*-value) of 1.3 (solid lines). **PLP**-DEs are colored blue,

proteins involved in flagellum formation and regulation pink, chemotaxis proteins ocher, proteins involved in transport and metabolism of galactitol green, of glycerol orange, of glutamate dark blue, and proteins of the methylisocitrate cycle purple. Chemotaxis signaling (**B**), methylcitrate cycle (**C**), galactitol uptake and metabolism (**D**), and glycerol uptake and metabolism in *E. coli* (IM = inner membrane, OM = outer membrane, G3P = glycerol 3-phosphate, G3POR = G3P with ether to rest R, DHAP = dyhydroxyacetone phosphate). Proteins that were dysregulated upon *yggS* knockout are marked with an aterisk (*).

We observed upregulation of proteins involved in putrescine metabolism and transport (PuuC, PuuE, PotF). This is in line with previous reports postulating increased levels of putrescine and putrescine-related metabolites upon *yggS* knockout.[283] Remarkably, we observed dysregulation of entire pathways for instance downregulation of *E. coli* chemotaxis signaling pathway including the methyl-accepting chemotaxis proteins (MCPs) Tsr, Tar, Tap, Trg, and Aer that sense a variety of amino acids, sugars, temperature, or pH-values. MCPs utilize a set of cytoplasmatic signaling proteins (CheW, CheA, CheB, CheR, CheY, and CheZ) to control flagellar rotation and sensory adapation (**Figure II-42B**).[324,325] In line with a reduction of chemotaxis-related proteins we observed downregulation of a huge number of flagellar proteins and motor proteins like FliC, FlgL, and MotB. Moreover, proteins involved in galactitol[326] and glycerol[327] uptake and metabolism together with enzymes involved in the methylcitrate cycle[328] were significantly overexpressed (**Figure II-42C-E**). Interestingly, all upregulated metabolic pathways result in the direct generation of pyruvate or metabolic precursors as dihydroxyacetone phosphate that can be converted to pyruvate *via* the glycolysis. This observations are in line with Δ*yggS E. coli* strains accumulating pyruvate,[282] as mentioned above. The data strongly support an involvement of PLPBP orthologs in central energy metabolic pathways already suggested by the co-IP data from human cells.

8.5 Conclusion

Taken together, we demonstrated -contrarily to previous postulations- that human PLPBP is mainly present as a dimer, which displays an unusual oligomerization behavior as bacterial and yeast orthologs are exclusively present in a monomeric state. This observation points to an additional adopted function of the human ortholog. Interestingly, the solvent exposure of **PLP** is altered upon dimerization, which shields the cofactor from the environment, unraveling important contributions of the oligomeric state and binding interfaces to **PLP** availability. We further analyzed several PLPBP mutants causing VitB6-EP and observed altered cofactor binding as well as oligomerization properties compared to the wild-type. As cellular PLPBP is present both in cytosol and mitochondria, it could take over a global role for instance in **PLP** housekeeping.

Although PLPBP was proposed to mediate transport of **PLP** to apo-vitamin B6-binding enzymes, we could not detect an overrepresented interaction with **PLP**-DEs upon cross-link/co-IP. However, analysis of the cellular binding network of PLPBP resulted in several energy-related proteins including enzymes associated with the TCA cycle. Though PLPBP may not be involved in cofactor transport it could be involved in **PLP** storage, supported by the fact that PLPBP can be enriched with the cofactor mimics even under short incubation times of 30 min with low concentrations of probe, as well as in the lysate (cf. Part II, Chapter 3) which points to a fast and efficient uptake of **PLP**. Interestingly, in HeLa cells, the median level of protein expression is 21,000 copies per cell, but roughly 200,000 copies of PLPBP are expressed in one cell.[329] Therefore, disruption of the protein may hamper cellular **PLP** storage and therefore increase the free **PLP** pools. This may further lead to increased activity of **PLP**-DEs as our experiments suggested that cofactor loading may be a means of enzyme regulation (cf. Part II, Chapter 6). Increased activity of B6-enzymes may lead to phenotypic effects observed upon PLPBP disruption as altered intracellular amino acid pools.[282,283,285,300,330] A knockdown, as well as an overexpression of human PLPBP did not result in changes of expression levels of **PLP**-DEs or **PLP** salvage enyzmes. However, our global proteome analysis revealed several dysregulated proteins upon PLPBP expression changes, given hints to possible cellular involvements of the **PLP**-DE. Moreover, a proteomic analysis of *E. coli* carrying a knockout in the *yggS* gene (the PLPBP ortholog) resulted in alterations of entire metabolic pathways, for instance downregulation of the chemotaxis response. Remarkably, several pathways resulting in the generation of pyruvate were upregulated including galactitol and glycerol metabolism and uptake, suggesting a central role of PLPBP orthologs in cellular energy housekeeping as already suggested by the co-IP data. As PLPBP is associated with pathophysiological phenotypes, as well as takes over a special **PLP**-related role supported by our labeling experiments, further investigations are urgently needed to fully understand the cellular role of this extraordinary **PLP**-DE.

9 Conclusion

Since the distribution of a cofactor among proteins and complexes within the cell determines its metabolic effects, as well as its role in health and disease, a deeper insight into the global distribution of **PLP** over the protein population is of urgent need. So far, attempts to study the **PLP**-ome have focused on the visualization of yeast mitochondrial **PLP**-binding proteins by western blot using an anti-**PL** antibody.[273] The basic limitation of a gel-based readout is the impossibility to identify and quantify the whole set of **PLP**-DEs, which could be addressed by a MS-based readout. Peptides modified by the reduced **PLP** analogue display mass spectrometric characteristics similar to those observed for phosphopeptides.[331] However, the suppressed ionization efficiency of **PLP**-modified peptides as well as their low abundance especially in complex cellular mixtures complicates their confident identification.[331,332] A possible solution displays an enrichment of **PLP**-modified peptides *via* immobilized metal ion affinity chromatography (IMAC). Nevertheless, though enriched by Ti^{4+}-IMAC, only few **PLP**-binding sites were identified from an *E. coli* lysate.[332] Our recently introduced **PL**-derived cofactor mimics enabled the confident identification and quantification of **PLP**-binding proteins from intact *S. aureus* cells. We were able to enrich 73% of the total annotated pool of **PLP**-DEs in *S. aureus* and assigned functions to previously unknown or uncharacterized proteins. The big advantage of our strategy is a fast and accurate quantification of **PLP**-binders under native cellular conditions. Further, our tool offers several application possibilities like the evaluation of inhibitor specificities *in situ*.

This part of the thesis describes the customization of our cofactor toolbox for higher organisms like human cells. Although all **PL**-probes were phosphorylated by hPLK, we observed strong differences in the turnover. The minimal modified **PL1** probe was phosphorylated with a catalytic efficiency in the range of the natural vitamer. However, longer 2'-substitued **PL2** and **PL3**

displayed catalytic efficiencies that were 50-fold reduced compared to **PL**. This observation was in contrast to studies with SaPLK, which phosphorylated probes at levels comparable to **PL**. Differently evolved substrate turnover mechanisms are most likely the reason for these divergences.

Though phosphorylated most efficiently, **PL1** displayed only weak *in situ* labeling. This stems from the increased intrinsic electrophilicity due to an extended conjugated system of the alkyne tag in direct neighborhood to the pyridine ring, which makes it prone to be attacked by cellular nucleophiles. However, a prolonged incubation of the cells for seven passages with **PL1** resulted in a large fraction (about one-third) of known **PLP**-DEs significantly enriched together with few non-vitamin B6-dependent background binders. **PL2** is efficiently entering the cell, suggested by an intense *in situ* labeling detected *via* in-gel fluorescence. A stronger labeling pattern underlines the higher electrophilicity of **PL1**, as the **PL2** alkyne moiety is isolated from the conjugated pyridine system by two CH_2-groups. Although we were able to enrich some **PLP**-DEs, **PL2** labeling suffered from a huge amount of background binders after incubation times up to 48 h especially at high probe concentrations. However, several significantly enriched **PLP**-DEs like OAT were not enriched upon previous labeling with **PL1**, suggesting different preferences for modifications at the **PL** scaffold. As metabolic labeling led to efficient enrichment of a large part of the human **PLP**-ome with **PL1**, we cultivated cells in the presence of **PL2** for seven passages. Although **PL2** supported the growth of the cells and the amount of non-vitamin B6-dependent background binders was dramatically reduced, the number of significantly enriched **PLP**-DEs could not be increased. If this observation is caused by the inefficient phosphorylation or the bulkier 2'-substituiton of **PL2**, making it unsuitable to fit in human **PLP**-DE active sites, remains to be determined. Finally, we examined the azide-functionalized **PL3**-probe for its labeling efficiency in human cells. We did neither observe any enrichment of **PLP**-DEs with **PL3** after 72 h, nor were cells able to grow in the presence of the probe, suggesting that the azide functionality prevented essential **PLP**-DEs from fulfilling their cellular roles sufficiently. This was in contrast to observations in bacteria, were **PL1** and **PL3** supported the growth, but **PL2** could not account for essential biological functions necessary for bacterial survival.[198] Taken all examined labeling conditions together we were able to cover for 46% of the total known pool of human **PLP**-DEs.[127,250]

Although we observed several non-vitamin B6-dependent proteins enriched upon labeling with **PL1** and **PL2**, we were able to demonstrate that **PLP**-DEs exclusively bind to the phosphorylated versions of the probes. Thereby, we exemplarily selected recombinant SHMT1 and PLPBP, which were both stabilized and labeled solely upon incubation with **PLP**-probes.

Moreover, labeling could be outcompeted by excess of natural **PLP**, suggesting that **PLP**-probes bind to the active sites of **PLP**-DEs.

As the development of selective **PLP**-DE inhibitors is challenging due to conserved active sites, an efficient cellular target screening method is of urgent need. We therefore applied human metabolic **PL1**-labeling to screen for *in situ* **PLP**-DE targets of the vitamin B6-antagonists AOAA and Pen. As for previously examined DCS in *S. aureus*,[198] we demonstrate an altered target scope insight the cell compared to *in vitro* experiments, opening new perspectives for therapeutic applications. These results serve as proof-of-principle for target screening of **PLP**-DE inhibitors in intact cells and may inspire exploration of more potent compounds.

Interestingly, comparison of labeling in different cell lines revealed an altered profile of enriched **PLP**-DEs with different fold-changes dependent on the type of cell. This was an intriguing finding as expression levels of corresponding **PLP**-DEs as well as vitamin B6 salvage enzymes were comparable throughout the examined cell lines. These results were corroborated by analysis with an anti-**PL** antibody, resulting in different saturation of **PLP**-DEs with cofactor in line with the chemoproteomic data. It is therefore tempting to speculate that **PLP** loading may be an additional regulator of enzyme activity. **PL** cofactor probes thus have the advantage to monitor and directly readout the degree of bound cofactor of a respective enzyme class rather than their abundance.

As **PLP** contains a very reactive aldehyde[246] making a tight regulation of free levels within the cell crucial, we wanted to address how **PLP** is trafficked to apo-vitamin B6-dependent enzymes inside the cell. As cofactor channeling by PLK and PNPO has been suggested as possible mechanism,[275–277] we investigated the subset of cellular binding partners of hPLK and PNPO by combined cross-link/co-IP. Neither under normal conditions, nor upon incubation with excess of B6 vitamers, a significant overrepresentation of **PLP**-DEs among the interactors of **PLP** salvage enzymes was detected. We therefore could not support a channeling hypothesis, at least not under the conditions examined.

Finally, we investigated structural and cellular properties of PLPBP more in detail. The so far mainly uncharacterized protein is conserved throughout evolution and was the most enriched hit under almost all labeling conditions. PLPBP orthologs are essential for some organisms like *S. aureus* or *Pseudomonas aeruginosa*,[283] whereas they are dispensable for others (*e.g. E. coli* or human cells).[282,290] Mutations in the human *PLBPP* gene cause severe epileptic attacks,[285] demonstrating a central role of the protein in health. Remarkably, PLPBP was efficiently enriched within a short *in situ* labeling time of 30 min. Further, PLPBP was the only **PLP**-DE

which was able to react with **PLP**-probes in lysate. A fast probe uptake, independent of cellular integrity, indicates a special role of PLPBP compared to other human **PLP**-DEs. This was underlined by examination of human recombinant PLPBP that displayed to be, independent of **PLP**, present as monomer and dimer, which makes up the major fraction. As **PLP**-DEs are dimers or of higher oligomeric composition, a monomeric state exerts an extraordinary feature of PLPBP and could be an explanation for the fast uptake of our cofactor probes even under perturbed conditions as resulted upon cell lysis. The observed dimeric state of PLPBP seems to be a special property of the human ortholog, as all described bacterial, as well as the yeast version of the protein are monomeric.[283,289] Interestingly, we found that cofactor accessibility is altered upon oligomerization. UV/Vis-studies as well as XL-MS analysis displayed **PLP** to be located at the dimerization interface and therefore shielded from the environment. This is an intriguing finding as oligomerization could be a tool for regulation of cofactor accessibility and exchange with the solvent. We further examined spectral and oligomeric properties of PLPBP mutants associated with VitB6-EP. Moreover, we observed altered expression efficiency in bacterial and human cells dependent on the type of mutation. Remarkably, mutants are only present either in the monomeric or dimeric state. As human PLPBP turned out to be localized in cytosol and mitochondria by immunohistochemical staining, a potential role in global **PLP** trafficking could be possible. Though proposed previously,[285] we could not find evidence for PLPBP transferring **PLP** to apo-B6-enzymes by a combined cross-link/co-IP. However, we detected several contacts to energy-related proteins in particular to enzymes associated with the TCA cycle. Though incubated with a huge number of common transaminase educts, we could not detect any spectral changes, suggesting that PLPBP does not act on for instance metabolites of the Krebs cycle. Further, we did not observe any changes in expression profiles of **PLP**-DEs or **PLP** salvage enzymes upon PLPBP knockdown or overexpression revealed by a global proteome analysis. However, we found several proteins to be dysregulated upon changing expression levels of PLPBP, which might give hints to its cellular function. Examination of proteomic changes upon knockout of the *E. coli* PLPBP ortholog YggS revealed several upregulated metabolic pathways that result in the generation of pyruvate. Taken together, PLPBP displays an exciting **PLP**-binding protein with a special vitamin B6-realted role insight the cell. Our data may serve as starting point for in-depth characterization of this conserved and unusual vitamin B6-dependent protein.

10 Outlook

The currently main limitation concerning the screening of the human **PLP**-ome with the **PL**-probes is that the only efficiently functioning probe in terms of the enrichment of **PLP**-DEs is **PL1**. Labeling with this probe is further restricted by its high intrinsic electrophilicity making it prone to be attacked by cellular nucleophiles that prevent downstream CuAAC and therefore protein enrichment. The inefficient labeling of **PL2** is most likely caused by the hampered bioactivation of the longer 2'-substituted probe, preventing it from efficient phosphorylation within the cell. However, as **PL2** labels some **PLP**-DEs and supports the growth of human cells, the question whether the low number of enriched **PLP**-DEs is caused by low levels of the probe within the cell or by steric hindrance of probe uptake due to unsuitable active site geometries remains to be answered. One possible strategy could be the application of a **PL2P** prodrug analogue. Such a prodrug ideally masks the negative charge of the phosphate group enabling the compound to cross the cellular membrane. Phosphoramidates display phosphate mimics which have been shown to deliver 5′-nucleoside monophosphates after simple esterase activation, spontaneous cyclization, followed by phosphoramidase-catalyzed cleavage of the amide.[333] Such a **PL2P**-analogue could directly release biologically active probe and could therefore act independently of the phosphorylation by hPLK. A successful application of other probes could overcome the restricted labeling capacity of **PL1** and therefore enable access to a larger fraction of the known human **PLP**-ome and may also result in so far uncharacterized proteins. An increased coverage would render the *in situ* target screening method suitable for vitamin B6-antagonists that target **PLP**-DEs like GABA-AT, which are currently not labeled by **PL1**.

As demonstrated, the degree of cofactor loading seems to be cell type dependent. Our probes therefore serve not only for the identification of **PLP**-DEs, but also for the readout of their cofactor binding state. This opens up new application possibilities especially in the clinical field. For instance, the cofactor loading degree in infected or diseased cells can be investigated and compared to uninfected or healthy cells as many **PLP**-DEs play crucial roles in cellular metabolism and function and are associated with diseases.

With an efficient labeling strategy in hand, the approach could not only be used to study different human cell types as already performed in the current thesis, but may also serve to examine even more complex systems like organoids. Another application could be the study of the human microbiome, as well as the investigation of cross-talks between host and pathogens that may affect the status of **PLP**-DEs.

One further possibility could be the investigation of intracellular cofactor trafficking, as well as *in organello* transport. For instance, while there is some redundancy in pathways for amino acid processing between the cytosol and mitochondria, the mitochondrial requirement for **PLP** to support Fe-S cluster assembly and heme biosynthesis seems to be absolute.[334] As our probes have the potential to be linked to a fluorophore previous to labeling, a fluorescence-based readout of subcellular cofactor localization and distribution under different cellular states or conditions is possible. Further, cofactor loading of comparably less investigated organelles like peroxisomes may also lead to interesting insights in **PLP** distribution.

Though the main entry of vitamin B6 into cells and tissues is mediated *via* simple diffusion of the unphosphorylated vitamers with subsequent re-phosphorylation by PLK,[197] there is some evidence for a direct uptake of **PLP**. Studies with isolated hepatocytes indicated that the liver can also release **PLP** into the circulation.[335] Further, there is evidence for **PLP** entering erythrocytes.[336] These studies support the existence of a human vitamin B6 plasma membrane transporter as observed for yeast.[337] Further, as **PL** is readily phosphorylated by cytosolic PLK, the intriguing question remains how **PLP** is entering the organelles within the cell. For instance, as many human mitochondrial **PLP**-DEs are known it is questionable whether those enzymes are entering mitochondria already fully saturated with cofactor, or if they are still supplied with **PLP** after the entry. As **PLP** loading may display a means of enzyme regulation supported by our data, there is a need for the cofactor to enter organelles even in the membrane-impermeable, phosphorylated form as mitochondria have virtually no ability to synthesize **PLP**. Indeed, rat liver mitochondria have been demonstrated to take up **PLP**.[338] Further, a mitochondrial **PLP**-transporter Mtm1p has been discovered in yeast. Knockdown, as well as

knockout of Mtm1p led to disruption of mitochondrial iron homeostasis, guided by the **PLP**-DE cysteine desulfurase, which is required for Fe-S cluster biosynthesis, as well as by the yeast ortholog of ALAS, catalyzing the first committed step in heme biosynthesis.[273] These phenotypes support the need of an active loading of mitochondria with **PLP** to fulfill their biological roles. Inspired by the work in bacteria, where diazirine and alkyne functionalized cofactor mimics (**Figure II-6**) led to discovery of microbial vitamin transporters,[227] synthetic introduction of a diazirine into **PL**-probes may also enable the identification of human plasma and organelle **PLP**-transporters and therefore lead to exiting insights into cellular cofactor trafficking mechanisms.

Finally, further investigations of conserved PLPBP are of urgent need to fully understand the role of this special **PLP**-binding protein. A knockout of human PLPBP would lead to interesting insights into proteomic changes and can surve as starting point for several experimental set-ups. For instance, intracellular pyruvate determination can be repeated using the knockout cells and transfection of knockout cells with the human expression plasmids carrying the VitB6-EP mutants may provide access to the analysis of interaction partners of PLPBP variants. The interactome of PLPBP mutants would give hints on their cellular localization, which may further be elucidated by immunohistochemical staining. Finally, proteomic changes in the presence of those pathological mutants can be investigated.

11 Experimental section

11.1 Chemical synthesis

Synthesis of **PL2** and **PL3** was performed as described previously.[198] Synthesis of **PL1** was performed as described with slight modifications.
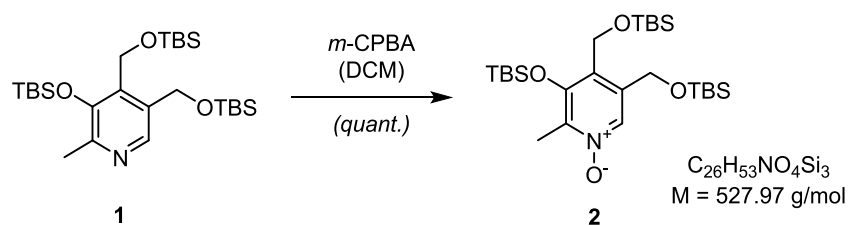
Scheme II-1 Synthesis of *tert*-butyldimethylsilyl (TBS)-protected PN *N*-oxide. *N*-oxidation of **1** by *meta*-chloro perbenzoic acid (*m*-CPBA) in dichloromethane (DCM) led to quantitative formation of **2**.

PN hydrochloride was protected with TBS-Cl to yield **1**. *Meta*-chloro perbenzoic acid (*m*-CPBA) (77%, 3.86 g, 17.2 mmol, 1.2 eq) was added to a solution of **1** (8.00 g, 15.7 mmol, 1.0 eq) in anhydrous dichloromethane (DCM, 150 mL) at 0 °C in five portions. After addition, the reaction was stirred vigorously at r.t for 2 h. Then, the reaction mixture was washed with 20% sodium sulfite solution (1×80 mL) and saturated sodium bicarbonate solution (1×80 mL), dried over sodium sulfate and concentrated to yield crude *N*-oxide **2** quantitatively as a white solid, which was used directly in the next step (**Scheme II-1**).
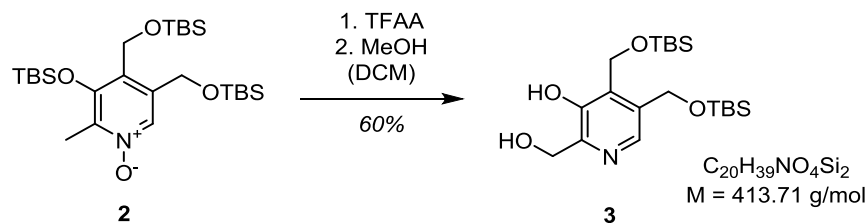
Scheme II-2 Synthesis of TBS-protected diol 3. BOEKELHEIDE-rearrangement of *N*-oxide **2** by trifluoroacetic anhydride (TFAA) in DCM and subsequent addition of MeOH gave diol **3**.

TFAA (10.3 mL, 15.5 g, 73.9 mmol, 5.0 eq) was slowly added to a solution of crude *N*-oxide **2** (7.80 g, 14.8 mmol, 1.0 eq) in anhydrous DCM (120 mL) *via* syringe at 0 °C. The ice bath was removed and the reaction was stirred overnight. Then, anhydrous MeOH (30 mL) was slowly added at 0 °C and the reaction was stirred for 20 min and subsequently allowed to warm to r.t. while stirring for another 20 min. The mixture was diluted with DCM (200 mL) and neutralized with saturated sodium bicarbonate solution (1×200 mL). After separation, the organic phase was dried over sodium sulfate and concentrated under reduced pressure. The residue obtained was purified by flash chromatography (EtOAc/hexanes, 10-50%) to yield diol **3** (3.64 g, 8.80 mmol, 60%) as a pale yellow solid (**Scheme II-2**). All further steps were carried out as described.[198]

11.2 Biological methods

11.2.1 Cloning and overexpression of recombinant proteins

GBA was purchased from ProSpec (Cat# enz-908), native porcine IDH (NADP) from Creative Enyzmes (Cat# NATE-0350), and L-MDH (pig heart) from Sigma (Cat# 10127914001). N-terminally Strep-II tagged hPLK, SHMT1, and PLPBP were cloned and expressed using the primers and conditions listed in **Table II-2**. Cloning was performed using the Invitrogen Gateway cloning system with pDONR201[Kan] as the donor vector and pDest007[Amp] as the destination vector. Proteins were expressed in *E. coli* Rosetta2 (DE3) (Merck, Cat# 71400) carrying a chloramphenicol resistance plasmid. Bacteria were grown in lysogeny broth (LB)-media containing ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37 °C to an OD_{600} of 0.6-0.8 and expression was induced through the addition of 0.2 µg/mL anhydrotetracycline (ATET). Bacteria were harvested and washed with PBS (6,000 x *g*, 4 °C) prior to cell lysis and protein purification.

Table II-2 Primer sequences, template DNA, and overexpression conditions for hPLK, SHMT1, and PLPBP, respectively.

Protein	forward primer (5'->3')	reverse primer (5'->3')	Over-expression	Template DNA
hPLK	ggggacaagtttgtacaaaaaagcaggctttgagaatctttattttcagggcgaggaggagtgccgggtg	ggggaccactttgtacaagaaagctgggtgtcacagcaccgtggcctgg	37 °C, 6 h	human cDNA library from HeLa (BioAcademia, Cat# 02-723)
SHMT1	ggggacaagtttgtacaaaaaagcaggctttgagaatctttattttcagggcacgatgccagtcaacgggg	ggggaccactttgtacaagaaagctgggtgtcagaagtcaggcaggccag	37 °C, 1 h	SHMT1 ORF clone in pcDNA3.1$^+$/C-(K)DYK, NM_004169 (GeneScript, Cat# OHu21374)
PLPBP	ggggacaagtttgtacaaaaaagcaggctttgagaatctttattttcagggctggagagctggcagcatgtcg	ggggaccactttgtacaagaaagctgggtgtcagtgctcctgtgccacctc	37 °C, 2 h	PLPBP ORF clone in pcDNA3.1$^+$/C-(K)DYK, NM_007198.3 (GeneScript, Cat# OHu09065)

hPLK mutants (ΔhPLK), as well as PLPBP mutants (ΔPLPBP) were generated *via* Quick Change site-directed mutagenesis[339] using the primers listed in **Table II-3** and **4**. As templates served pDest007^{Amp}-hPLK and pDest007^{Amp}-PLPBP for recombinant protein overexpression. For transfection into human cells the template was pcDNA3.1$^+$/C-(K)DYK-PLPBP.

Table II-3 Primer sequences for hPLK mutants (ΔhPLK).

ΔhPLK	forward primer (5'->3')	reverse primer (5'->3')
V14A	CAGAGCCACGCGATCCGCGGC	GCCGCGGATCGCGTGGCTCTG
V41A	GCGGTGAACTCTGCGCAGTTTTCAAAC	GTTTGAAAACTGCGCAGAGTTCACCGC
F43A	GTGAACTCTGTCCAGGCGTCAAACCACACAGGC	GCCTGTGTGGTTTGACGCCTGGACAGAGTTCAC
F43G	AACTCTGTCCAGGGCTCAAACCACACA	TGTGTGGTTTGAGCCCTGGACAGAGTT
T47A	TCAAACCACGCGGGCTATGCC	GGCATAGCCCGCGTGGTTTGA

Table II-4 Primer sequences for PLPBP mutants (ΔPLPBP).

ΔPLPBP	forward primer (5'->3')	reverse primer (5'->3')
Y69C	agcagttcctgaacgcagttctcgccaaaagtgc	gcacttttggcgagaactgcgttcaggaactgct
R241Q	ccaaaaatcgtgcttcctatttggacatttgtagatcctactt	aagtaggatctacaaatgtccaaataggaagcacgattttt

		gg
R41Q	accgccactagctggggctggatgg	ccatccagccccagctagtggcggt
V45D	ggtttggttttgctgtccgccactagccgg	ccggctagtggcggacagcaaaaccaaacc
E67K	tcctgaacgtagttcttgccaaaagtgcgctgc	gcagcgcactttggcaagaactacgttcagga

Polymerase chain reaction (PCR) products were digested with *Dpn 1* (New England BioLabs, Cat# R0176S) to remove wild-type template DNA. Plasmids carrying point mutated proteins were transformed into *E. coli* XL1 Blue for nick repair, previous to transformation into the expression strain Rosetta2 (DE3). Plasmids were transformed into *E. coli* Top10 (Thermo Fisher Scientific, Cat# C404010) for plasmid extraction for transfection. Expression and transfection (cf. Part II, Chapter 11.7.6) of point mutants carried out as described for wild-type hPLK and PLPBP.

11.2.2 Cell lysis, purification, and analytics of recombinant proteins

Buffer compositions, as well as columns used for purification are listed in **Table II-5**. The bacterial overexpression culture was resuspended in strep binding buffer and lysed by sonication. The lysate was clarified by centrifugation (36,000 x g, 30 min, 4 °C). Supernatant was loaded onto a StrepTrap column (5 mL, GE Healthcare, Cat# 28-9075) equilibrated with binding buffer using an Äkta purification system (GE Healthcare). After extensive washing, proteins were eluted in binding buffer containing 2.5 mM desthiobiotin. For preparation of apo-SHMT1, the column was washed with 20 mL binding buffer containing 25 mM hydroxylamine prior to elution. SEC was applied for further purification. Proteins were loaded onto columns equilibrated with SEC buffer. Oligomeric state of the proteins was determined according to calibration curves. Concentration of hPLK and its mutants was calculated according to Lambert-Beer's law from the absorbance at 280 nm recorded in an InfiniteM200 PRO reader (Tecan) and an extinction coefficient of 40,000 M^{-1}•cm^{-1} at 280 nm (calculated with ProtParam). Concentrations of holo- and apo-SHMT1 were determined with the Bradford assay (Roti-Nanoquant, Roth) and of PLPBP and its mutants with the BCA assay (Thermo Fisher Scientific). For PLPBP two purification strategies were applied. Either the protein was directly subjected to SEC (20 mM HEPES pH 7.6, 100 mM KCl, 1 mM DTT) after the Strep-purification (**Table II-5**, PLPBP strategy [a]; applied for studies with **PL**-probes; cf. Part II, Chapter 4), or was incubated with tobacco etch virus (TEV)-protease (**Table II-5**, PLPBP strategy [b]; applied for in-depth characterization of PLPBP; cf. Part II, Chapter 8). Therefore, concentration of PLPBP was determined after Strep-purification and DTT was added to a final amount of 1 mM (from a 500 mM stock prepared fresh in ddH$_2$O). The digest was initiated with the addition of a 40-fold molar excess of TEV-protease and incubation carried out for 16 h at 10 °C and 300 rpm.

Part II-11 Experimental section

Afterwards, buffer was exchanged to His-loading buffer (20 mM sodium dihydrogene phosphate pH 8.0, 150 mM NaCl, and 10 mM imidazole) using a reconcentrator (10 K MWCO). HisTrap column (5 mL, GE Healthcare, Cat# GE29-0510-21) was equilibrated with His-loading buffer previous to loading of PLPBP to remove His-tagged TEV protease. Flowthrough fractions containing PLPBP were subjected to thermofluor assays for determination of optimal SEC buffer composition. Finally, the protein was loaded onto the SEC column for purification.

Table II-5 Purification strategies including corresponding SEC buffer compositions for hPLK, holo- and apo-SHMT1, as well as PLPBP. Molecular weights (MW) were determined by intact-protein MS. Proteins were reduced with NaBH$_4$ previous to MS if indicated (*; TEV = tobacco etch virus).

Protein	Strep binding buffer	SEC buffer	SEC column	MW found (Da)
hPLK	100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA	20 mM HEPES pH 7.0, 100 mM KCl, 1 mM DTT	Superdex200 (GE Healthcare, Cat# 17517501)	38,209.5
Holo-SHMT1	50 mM HEPES pH 8.0, 150 mM NaCl, 1 mM EDTA	10 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA	Superdex200 (GE Healthcare, Cat# 17517501)	56,410.6*
Apo-SHMT1	100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA	10 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 10% glycerol	Superdex200 (GE Healthcare, Cat# 17517501)	56,178.6*
PLPBP [a]	50 mM HEPES pH 8.0, 150 mM NaCl, 1 mM EDTA	20 mM HEPES pH 7.6, 100 mM KCl, 1 mM DTT	Superdex75 (GE Healthcare, Cat# GE17-5174-01)	33,684.1* (not digested)
PLPBP [b]	50 mM HEPES pH 8.0, 150 mM NaCl, 1 mM EDTA	50 mM Bicine pH 9.0, 150 mM NaCl, 1 mM DTT	Superdex75 (GE Healthcare, Cat# GE17-5174-01)	30,480.1* (after TEV protease-digest)

Molecular weights of the proteins were confirmed by intact-protein MS (**Table II-5**). Proteins were stored at -80°C in small aliquots. ΔhPLK were purified as the wild-type protein. ΔPLPBP were purified *via* strategy [a] with an analytical Superdex200 column using 50 mM Bicine pH 9.0, 150 mM NaCl, 1 mM DTT as SEC buffer due to lower protein amounts. In order to monitor the impact of **PLP** on the oligomerization behavior of PLPBP, 10 µM dimer or monomer were incubated either 10 min at 37 °C in SEC buffer [b] alone or buffer containing 100 µM **PLP** and

separated using an analytical Superdex200 (GE Healthcare, Cat# 17517501) equilibrated with SEC buffer containing 100 µM **PLP** when indicated

11.2.3 UV/Vis spectra of proteins

UV/Vis spectra of protein samples (80-100 µM) were recorded in corresponding SEC buffers at 37 °C on an InfiniteM200 PRO reader (TECAN, Cat# IN-MNANO; 300-600 nm, 2 nm increments).

11.2.4 hPLK kinetic assay

PL-probes were phosphorylated using SaPLK as described previously.[198] UV/Vis-spectra of **PL** and **PL**-probes, as well as their phosphorylated counterparts were recorded in assay buffer (20 mM HEPES pH 7.0, 100 mM KCl, 1 mM MgCl$_2$) at 37 °C using an InfiniteM200 PRO reader (TECAN, Cat# IN-MNANO). The initial rate of product formation was measured by following the increase in absorbance at the maxima of **PLP** (392 nm) and **PLP**-probes (**PL1P**: 416 nm, **PL2P**: 396 nm, **PL3P**: 400 nm). A typical assay contained 0.25 µM hPLK or the corresponding mutant in assay buffer and activity was monitored at 37 °C using an InfiniteM200 PRO reader (TECAN, Cat# IN-MNANO) in presence of 1 mM ATP. **PL** and **PL**-probes were added from 100x stocks in dimethylsulfoxide (DMSO) to varying concentrations including no enzyme controls. Three biological replicates with three technical replicates each were performed. Data were processed using Prism 6 (GraphPad). Absorbance values were transformed into concentrations using molar extinction coefficients of **PLP** and **PLP**-probes at corresponding absorbance maxima derived from the recorded UV/Vis-spectra (ε_{392nm} (**PLP**) = 6,054.42 M^{-1}cm^{-1}, ε_{416nm} (**PL1P**) = 9,657.49 M^{-1}cm^{-1}, ε_{396nm} (**PL2P**) = 9,374.44 M^{-1}cm^{-1}, ε_{400nm} (**PL3P**) = 6,240.76 M^{-1}cm^{-1}). Initial slopes from linear regions were determined from three technical replicates (with corresponding error) and plotted against concentrations of **PL** or **PL**-probe, respectively. Kinetic values (k_{cat}, K_m, v_{max}) and corresponding errors were derived from non-linear regression analysis (k_{cat} and Michaelis-Menten) using default parameters. Means of kinetic values and corresponding errors from the three biological replicates were calculated using error propagation.

11.2.5 Reconstitution of apo-SHMT1 with PLP and PL1P

For reconstitution with **PLP** and **PL1P**, 50 µM apo-SHMT1 were incubated with 5-fold molar excess of cofactor in 10 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT, 0.5 mM ethylenediaminetetraacetic acid (EDTA), and 10% glycerol overnight at 4 °C under rotation. **PL1P** was generated as described previously.[198] Residual **PLP** or **PL1P** was removed by

desalting the proteins using a HiTrap Desalting column (5 mL, GE Healthcare, Cat# 17-1408-01) and an Äkta purification system (GE Healthcare) simultaneously placing them into 10 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT, and 0.5 mM EDTA.

11.2.6 Sample preparation for intact-protein MS

10 µM protein samples (25 µL, SHMT1 and PLPBP in corresponding SEC buffer) were treated with 10 mM $NaBH_4$ (2 µL of 250 mM stock prepared fresh in 0.1 M NaOH) at r. t. for 30 min. Residual $NaBH_4$ was quenched by acidification to pH 5-6 with HCl (5-10 µL of 0.5% FA) and neutralized to pH 7 with NaOH (5-10 µL of 0.1 M NaOH). Samples were diluted to 50 µL with PBS (5 µM final enzyme concentration). 2 µM hPLK, ∆hPLK or not reduced PLPBP or ∆PLPBP were prepared in SEC-buffer for measurement by intact-protein MS.

11.2.7 Intact-protein MS

Full-length protein measurements were performed as described previously.[198] Proteins were measured on a MassPREP On-Line Desalting Cartridge (Waters) on an Ultimate 3000 HPLC system (Dionex) coupled to a Finnigan LTQ-FT Ultra mass spectrometer (Thermo Fisher Scientific) with electrospray ionization (spray voltage 4.0 kV, tube lens 110 V, capillary voltage 48 V, sheath gas 60 arb, aux gas 10 arb, sweep gas 0.2 arb). Xcalibur Xtract Software (Thermo Fisher Scientific) was used for data analysis and deconvolution.

11.2.8 Gel-based labeling of recombinant proteins

Gel-based labeling of recombinant proteins and **PL**-probe phosphorylation was performed as described previously.[198] In brief, proteins were labeled with **PL**- or **PLP**-probes and reduced with $NaBH_4$ previous to CuAAC to rhodamine-azide. Samples were analyzed by SDS-PAGE and fluorescence scanning. For competition with **PLP**, apo-SHMT1 and PLPBP dimer and monomer were incubated with 5-fold molar excess of **PL1P** for 30 min at 25 °C (10 µM protein in SEC buffer, 50 µL per sample). Afterwards, **PLP** was added to indicated amounts from 2 and 100 mM stocks in water, respectively, and proteins were further incubated for 30 min at 25 °C previous to $NaBH_4$ reduction and CuAAC.

11.2.9 Thermal stability assay with recombinant proteins

2 µM of protein in SEC buffer or GBA in PBS were incubated with varying excess of **PL[P]** and **PL[P]**-probes (2 µL of corresponding stocks) or 20-fold molar excess of transaminase educts. **PLP** and **PLP**-probes were generated with the SaPLK as descrived previously.[198] Sypro orange

(Thermo Fisher Scientific, Cat# S6650) was added to a final concentration of 1x from a 5,000x stock. Temperature was increased from 20 to 89.6 °C with a heating rate of 0.3 °C per min, monitoring fluorescence at 569 nm in a CFX96 Real-Time System (Bio-Rad, Cat# 20421). Three independent replicates were performed for each condition. Denaturation curves were fitted according to a sigmoidal trace. Melting temperatures were calculated as mean value with corresponding standard deviation (SEM).

11.2.10 Cell-type specific protein expression analysis

Cell lines were grown as described under 11.3 and incubated for 15 min at 4 °C in lysis buffer (1% (w/v) sodium deoxycholate, 1% (v/v) NP-40 in PBS pH 7.4). Supernatant was clarified at 21,000 $x\,g$ for 20 min at 4 °C. Protein concentration was adjusted with a BCA-assay. Lysate was mixed with 2x gel loading buffer. Samples were separated on a 12.5% SDS-gel (80 µg of lysate per cell line) and plotted on a PVDF membrane (Roti-PVDF, 0.2 µm, Roth, Cat# 8989.1) using a semi-dry blotting station (Trans-Bot SD Semi-Dry Transfer Cell, Bio-Rad, Cat# 1703940). For PLPBP, SHMT2, PNPO, ALAS1, and SHMT1 blocking was performed using 5% milk-powder (w/v) in PBS-T (PBS supplemented with 0.5% Tween-20) at r.t. for 1 h. Antibodies (mouse monoclonal anti-PLPBP antibody (clone 1G2, OriGene, Cat# TA505162); rabbit polyclonal anti-SHMT2 antibody (abbexa, Cat# abx128462); rabbit polyclonal anti-PNPO antibody (Sigma, Cat# HPA023204); rabbit polyclonal anti-ALAS1 antibody (Thermo Fisher Scientific, Cat# PA5-57434); rabbit polyclonal anti-SHMT1 antibody (abcam, Cat# ab55736)) were diluted according to manufactures protocol in 5% milk-powder in PBS-T and added to the membranes. Immunobinding carried out overnight at 4 °C. After extensive washing the membranes were incubated with secondary antibodies (goat anti-rabbit antibody conjugated to horseradish peroxidase (0.5 mg/mL, Invitrogen, Cat# 32260) or goat anti-mouse antibody conjugated to horseradish peroxidase (0.5 mg/mL, Invitrogen, Cat# 32230)) diluted 1:10,000 in 5% milk-powder in PBS-T for 1 h at r.t. For PLK, blocking was performed using 3% bovine serum albumin (BSA; w/v) in PBS-T at r.t. for 1 h. Mouse monoclonal anti-PLK antibody (Santa Cruz Biotechnology, Inc., Cat# sc-365173, RRID: AB_10708566) was diluted 1:100 in 3% BSA (w/v) in PBS-T and added to the membrane. Immunobinding carried out overnight at 4 °C. After extensive washing the membranes were incubated with the secondary antibody (goat anti-mouse antibody conjugated to horseradish peroxidase (0.5 mg/mL, invitrogen)) diluted 1:2,500 in 3% BSA in PBS-T for 1 h at r.t. The membranes were washed and chemo-luminescence was detected after incubation with freshly prepared ECI western blotting substrate solution (Pierce, Cat# PIER80196) in a Luminescent LAS 4000 image analyzer (Fujifilm, ordered *via* GE

Healthcare, Cat# 28955810). Membranes were stained with ponceau S (Sigma, Cat# P3504) to inspect for equal protein loading amounts.

11.2.11 PLP-ome analysis with the anti-PL antibody

Cell lines were grown as described under 11.3 and incubated for 15 min at 4 °C in lysis buffer (1% (w/v) sodium deoxycholate, 1% (v/v) NP-40 in PBS pH 7.4) either without or with 20 mM NaBH$_4$ added from a 500 mM stock prepared fresh in cold 0.1 M NaOH. Supernatant was clarified at 21,000 x g for 20 min at 4 °C. Protein concentration was adjusted with a BCA-assay. For analysis of recombinant proteins, hPLK was diluted to 30 µM in SEC buffer and mixed with 2x gel loading buffer. PLPBP and SHMT1 at 30 µM were saturated with 10-fold molar excess of **PLP** for 30 min at 37 °C. Non reduced samples were diluted with 2x loading buffer. Reduced samples were incubated with 10 mM NaBH$_4$ for 30 min at r.t. (added from a 250 mM stock prepared fresh in cold 0.1 M NaOH) previous to quenching with 5 µL 1% FA in H$_2$O and subsequent neutralization with 5 µL 0.2 M NaOH. Samples were diluted by addition of 2x loading buffer, separated on a 12.5% SDS-gel (70 µg of lysate per cell line and 8 µL of the 15 µM protein samples) and plotted on a PVDF membrane (Roti-PVDF 2.0, 0.2 µm, Roth, Cat# 8989.1) using a semi-dry blotting station (Trans-Bot SD, Bio-Rad, Cat# 1703940). Blocking was performed using 5% milk-powder (w/v) in PBS-T at r.t. for 1 h. The rabbit polyclonal anti-**PL** antibody (reduced vitamer, conjugated to BSA; GeneTex, Cat# GTX12625) was diluted 1:500 in 5% milk-powder in PBS-T and added to the membranes. Immunobinding carried out overnight at 4 °C. After extensive washing the membranes were incubated with secondary antibody (goat anti-rabbit antibody conjugated to horseradish peroxidase (0.5 mg/mL, Invitrogen, Cat# 32260) diluted 1:5,000 in 5% milk-powder in PBS-T for 2 h at r.t. The membranes were washed and chemo-luminescence was detected after incubation with freshly prepared ECI western blotting substrate solution (Pierce, Cat# PIER80196) in a Luminescent LAS 4000 image analyzer (Fujifilm, ordered *via* GE Healthcare, Cat# 28955810). Membranes were stained with ponceau S (Sigma, Cat# P3504) to inspect for equal protein loading amounts.

11.2.12 PLPBP incubation with DCS, Pen, and transaminase educts

130 µM of TEV-digested PLPBP monomer and dimer were incubated with 10, 25, or 50 mM DCS (Roth, Cat# CN37.1), L-Pen (TCI, Cat# P1370), and D-Pen (Acros Organics *via* Fisher Scientific, Cat# 10224750) (added from 10x stocks in water, pH adjusted to neutral), respectively, for 20 min at 25 °C in a total volume of 10 µL. 100 µM of TEV-digested PLPBP monomer and dimer were incubated with 20-fold molar excesses of transaminase educts at

37 °C. Afterwards, UV/Vis-spectra were recorded as described under 11.2.3.

11.2.13 Analytical labeling of intact cells for in-gel fluorescence readout

Cells were grown as described under 11.3, seeded in 6-well plates and grown until a confluency of 80-90%. HepG2 were labeled in in-house prepared RPMI and HeLa in in-house prepared DMEM media lacking dialyzed FCS (cf. Part II, Chapter 11.7.1). Probes (1,000x stocks in DMSO) or corresponding DMSO controls were diluted in media and applied to the cells. Labeling carried out for indicated incubation times. Cells were washed once with cold PBS (4 °C) and scraped to detach. After harvesting at 600 x g for 5 min at 4 °C, cells were lysed (PBS with 1% (w/v) sodium deoxycholate and 1% (v/v) NP-40) in the presence of 20 mM $NaBH_4$ (added from a 500 mM stock prepared fresh in cold 0.1 M NaOH) for 30 min on ice. If indicated, cells were lysed in the presence of 1 mM maleimide (added from a 100 mM stock prepared fresh in H_2O) and 20 mM $NaBH_4$ for 30 min on ice. Afterwards, lysate was centrifuged for 15 min at 21,000 x g and 4 °C. Proteins were precipitated by adding 4-fold volume of ice-cold acetone (-80 °C) overnight at -20 °C. Precipitated proteins were pelleted by centrifugation (18,000 x g, 15 min, 4 °C) and washed with ice-cold MeOH (2x 0.5 mL) using sonication to resuspend the pellets between washes. Proteins were solubilized in 50 µL PBS containing 0.4% (w/v) SDS und subjected to CuAAC as described previously.[198] If indicated, samples were incubated with 20 or 40 mM IAA added from a 100x fold stock prepared fresh in H_2O for 30 min at r.t. in the dark previous to CuAAC. To screen different click-conditions for **PL1** labeling, $CuSO_4$ concentrations, as well as ligand concentrations were varied. In addition, we examined the water-soluble THPTA ligand. After CuAAC, samples were quenched with 2x gel loading buffer and analyzed by SDS-PAGE with fluorescence scanning. Protein loading was visualized by Coomassie-staining of the gels.

11.2.14 Analytical labeling in lysate for in-gel fluorescence readout

Cells were grown as described under 11.3, seeded in 6-well plates and grown until a confluency of 80-90%. Cells were washed once with cold PBS (4 °C) and scraped to detach. After harvesting at 600 x g for 5 min at 4 °C, cells were lysed (PBS with 1% (w/v) sodium deoxycholate and 1% (v/v) NP-40) in the presence of 1 mM maleimide (added from a 100 mM stock prepared freshly in H_2O) for 30 min on ice. After centrifugation for 15 min at 21,000 x g and 4 °C, **PL1P** was added to concentrations indicated from a 2 mM stock in buffer (generated with SaPLK as described previously)[198]. Labeling carried out for 1 h at 4 °C under slow rotation previous to reduction with 20 mM $NaBH_4$ for 30 min on ice. Protein precipitation and CuAAC

carried out as described under 11.2.13.

11.2.15 Activity assays with IDH and L-MDH

All assays were performed at 32 °C in an InfiniteM200 PRO reader (TECAN, Cat# IN-MNANO). For IDH activity measurements, 1 µM IDH (from a 19.5 µM stock, diluted with 50 mM HEPES pH 7.4 from the manufactor's stock), 6 mM isocitrate (from a 400 mM stock in ddH$_2$O), and 3.75 mM NADP$^+$ (from a 187.5 mM stock in ddH$_2$O) were incubated in 50 mM HEPES pH 7.4 and absorbance increase at 340 nm was monitored over the time. For L-MDH activity measurements, 0.695 µM L-MDH (from a 13.9 µM stock, diluted with 100 mM potassium dihydrogene phosphate pH 7.4 from the manufactor's stock), 6 mM oxaloacetic acid (from a 300 mM stock in ddH$_2$O), and 3.75 mM NADH (from a 187.5 mM stock in ddH$_2$O) were incubated in 100 mM potassium dihydrogene phosphate pH 7.4 and absorbance decrease at 340 nm was monitored over the time. To examine possible activity alterations caused by PLPBP monomer and dimer, 10-fold molar excesses of TEV-digested PLPBP over IDH and L-MDH were added in the presence of isocitrate or oxaloacetic acid, respectively, and the mixture was incubated for 30 min at 32 °C previous to initation of reaction with NADP$^+$ or NADH, respectively. Triplicates per condition were prepared. Linear regions of substrate turnover were fitted and slopes were normalized to the controls containing IDH and L-MDH mixed with PLPBP SEC buffer using Prism 6 (GraphPad).

11.2.16 Determination of intracellular pyruvate concentrations

Cells were cultured as described under 11.3 and PLPBP knockout was performed as described under 11.7.5. Cells were seeded in 152 cm² dishes and grown until 80-90% confluency. Volumes for siRNA transfection as reported under 11.7.5 were adjusted according to the increased surface area of culture dishes. Cells were washed once with PBS and then detached in PBS (600 x *g*, 4 °C, 5 min). Sample preparation was performed according to the manufactors protocol of the Pyruvate Assay Kit (cayman chemical, Cat# 700470). In brief, cells were lysed with 0.25 M metaphosphoric acid and cell debris was removed at 10,000 x *g* and 4 °C for 5 min. After neutralization of supernatant with potassium carbonate, samples were assayed for pyruvate concentration using a coupled assay and fluorescence readout in an InfiniteM200 PRO reader (TECAN, Cat# IN-MNANO). Absolute pyruvate concentrations were determined using a standard curve included with the kit. Two independent biological replicates were performed for each condition.

11.3 Cell culture

All cell lines were obtained from ECACC *via* Sigma Aldrich. HepG2 (male, Cat# 85011430), K562 (female, Cat# 89121407), HEK293 (female, Cat# 85120602), and HCT116 (male, Cat# 91091005) were cultivated in RPMI-1640, HeLa (female, Cat# 93021013) and HuH7 (male, Cat# 01042712) in DMEM media supplemented with 10% L-glutamine and 10% FCS at 37 °C and humidified 5% CO_2 atmosphere. The cells were routinely tested for mycoplasma contamination.

11.4 Microbe strains

E. coli K12 (DSMZ, Cat# 498) and BW25113 parent strain (Keio knockout library, CGSC# 7636) were grown in LB media and *E. coli* Δ*yggS* (Keio knockout library, CGSC# 11827, JW2918-1) in LB media with 25 µg/mL kanamycin A at 37 °C with shaking at 200 rpm until reaching the exponential phase (OD_{600nm} 0.7).

11.5 Immunochemical staining and imaging

HepG2 cells were cultured as described under 11.3 and plated on glass coverslips for 2 days. For fixation, cells were washed with PBS and incubated with 4% paraformaldehyde (PFA) in PBS for 15 min at r.t. Afterwards, cells were washed with PBS three times. Cells were permeabilized with 0.3% Triton X100 in PBS for 5 min. After permeabilization, cells were blocked for 1 h at r.t. with 10% Normal Goat Serum (#VEC-S-100, Biozol) and 3% BSA (#A4503-50G, Sigma-Aldrich) in 0.1% Tween in PBS. Cells were then incubated with the primary antibody in blocking solution overnight at 4 °C, using the following dilutions: 1:200 rabbit polyclonal anti-PLPBP (#PA5-32036, Thermo Scientific) and 1:500 mIgG2a-anti-TOM20 (#sc-17764, Santa Cruz Biotechnology) antibody. The next day, following several washes with 0.1% Tween in PBS, cells were incubated with 1:1,000 dilution of biotin-conjugated secondary goat anti-rabbit antibody (Life Technologies) in blocking solution for 1 h at r.t. to enhance the PLPBP signal. After additional washing steps, cells were incubated with Alexa Fluor546-conjugated goat anti-mIgG secondary antibody and AlexaFluor488-conjugated streptavidin (Life Technologies), both 1:1,000, in blocking solution for at least 1 h at r.t. Further, 0.1 µg/ml 4,6-diamidino-2-phenylindole (DAPI, Sigma Aldrich) was applied to counterstain nuclei. Finally, coverslips were inversely glued onto object slides using AquaPolymount (Polyscience). Immunohistochemically stained cells were imaged using a Leica SP8 confocal laser scanning microscope and the 40x water immersion objective.

11.6 XL-MS

11.6.1 Chemical cross-linking of PLPBP

DSSO was synthesized as described previously.[110] PLPBP dimer (TEV-digested, 5 µM in 50 µL 20 mM HEPES pH 7.6, 100 mM KCl) was incubated with varying molar excess of DSSO cross-linker (stock solutions dissolved in DMSO) and incubated for 1 h at 37 °C with shaking (300 rpm). Samples were analyzed *via* SDS-PAGE. The gel was stained with SYPRO Ruby Protein Gel Stain (Thermo Fisher Scientific, Cat# S12000). For MS analysis, PLPBP dimer (TEV-digested, 20 µM in 200 µL 20 mM HEPES pH 7.6, 100 mM KCl) was incubated with either 50 or 100-fold excess DSSO for 1 h at 37 °C and 300 rpm and subsequently separated on an analytical Superdex200 (GE Healthcare, Cat# 17517501) connected to an Äkta system (GE Healthcare) in 20 mM HEPES pH 7.6, 100 mM KCl. Fractions corresponding to monomer and dimer were unified and evaporated *in vacuo*.

11.6.2 MS preparation of cross-linked proteins

MS preparation of cross-linked PLPBP was performed as described previously.[110] In brief, proteins were denatured, reduced, alkylated and predigested with Lys-C (Wako, Cat# 125-05061) followed by overnight incubation with trypsin (Promega, Cat# V5111). Cross-linked peptides were enriched *via* cation-exchange chromatography (SUPELCO, Cat# 66889-U) and desalted with double layer C18 stage-tips (SUPELCO, Cat# 66883-U) previous to MS analysis.

11.6.3 MS analysis of cross-linked peptides

MS analysis of cross-linked PLPBP peptides was performed as described previously.[110] In brief, peptides were separated within 105 min on a UltiMate 3000 nano HPLC system (Thermo Fisher Scientific) equipped with an Acclaim C18 PepMap100 75 µm ID x 2 cm trap (Thermo Fisher Scientific, Cat# ES803A) and an Acclaim PepMap RSLC C18 separation column (75 µm ID x 50 cm; Thermo Fisher Scientific, Cat# 164535) coupled to an EASY-source equipped Thermo Fisher LTQ Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). Each sample was measured twice, one time using the EThcD-CID and the other time using the MS^2-MS^3 fragmentation strategy for DSSO cross-linked peptides.

11.6.4 Computational analysis of cross-linking data

Computational analysis of cross-linked PLPBP peptides was performed as described

previously.[110] In brief, MS raw data for DSSO cross-linked proteins were analyzed using Proteome Discoverer Version 2.2 (Thermo Fisher Scientific, Cat# OPTON-30795) and its XlinkX plug-in. The database contained the PLPBP sequence, as well as all common contaminants from the MaxQuant (version 1.6.0.1)[124] bin folder to receive statistically significant results. The detected cross-links resulting from the two fragmentation strategies, as well as DSSO excesses were combined. A PLPBP model crystal structure was generated based on the template crystal structure from yeast (PDB: 1B54)[291] using SWISS-MODEL.[190–193] The PLPBP crystal structure served as input for dimer modeling using HADDOCK software.[19] Cross-links solely occurring in the dimer fraction of PLPBP were defined as active residues using the Easy prediction interface and default parameters. Cross-links were displayed onto the dimer model (cluster 1_1) using the XLinkAnalyzer[15] (version 1.1) plug-in into Chimera[194] (version 1.11) applying a distance threshold of 35 Å for DSSO.

11.7 Proteomics

11.7.1 PL-probe labeling and competition in intact cells

All cell lines were cultivated as described under 11.3. HeLa and HuH7 were transferred into in-house prepared DMEM media (**Table II-6**) lacking pyridoxine. HepG2, K562, HEK293, and HCT116 were transferred into in-house prepared RPMI media (**Table II-7**) lacking pyridoxine. **PL**-probes were added fresh from 1,000x stocks in DMSO to the concentrations indicated. For labeling times up to 24 h incubation carried out without dialyzed FCS, for incubation times longer than 24 h, as well as for metabolic labeling, 10% dialyzed FCS (Sigma, Cat# F0392) were added to in-house prepared media. For all labeling experiments cell lines were additionally transferred into corresponding in-house prepared media containing 5 µM **PL** added fresh from a corresponding stock in DMSO to ensure for equal amounts of organic solvent.

For metabolic labeling, cells were cultivated for indicated time points in corresponding in-house prepared media supplemented with 10 µM **PL1** (added fresh from a 100 mM stock in DMSO), 20 µM **PL2** (added fresh from a 100 mM stock in DMSO), or 5 µM **PL** (added from a corresponding stock in DMSO to ensure for equal amounts of organic solvent), respectively, and afterwards seeded onto plates (152 cm²) and grown until 80-90% confluency. We tried to culture HeLa and HepG2 in the presence of 20 and 10 µM **PL3**. For labeling times up to 48 h, cells were seeded onto plates (152 cm²) and grown until 80-90% confluency previous to labeling. For competition experiments AOAA (MedChemExpress, Cat# HY-107994), L-, and D-Pen were

diluted in in-house prepared media lacking dialyzed FCS to a final concentration of 1 mM (AOAA) or 10 mM (L- and D-Pen) and pH values were adjusted. **PL1** labeled HEK293 cells (passaged seven times in media lacking **PN**) were incubated with B6-antagonists for 2 h at 37 °C. One separate plate was seeded per biological replicate. Cells were washed once with cold PBS (4 °C) and then detached in cold PBS. After harvesting at 600 x g and r.t., cells were resuspended in 1 mL lysis buffer (1% (w/v) sodium deoxycholate, 1% (v/v) NP-40 in PBS pH 7.4) containing 20 mM $NaBH_4$ (added from a 500 mM stock prepared fresh in 0.1 M NaOH) and incubated for 30 min at 4 °C under rotation. Cell debris was removed at 4 °C for 20 min (21,000 x g). Soluble proteins were precipitated by adding 4x volume of cold acetone (-80 °C) overnight at -20°C.

Table II-6 Ingredients of in-house prepared DMEM media. The pH value of the media was adjusted to 7.88 (at r.t.).

Ingredient	Concentration [g/L]	Ingredient	Concentration [g/L]
$CaCl_2 \cdot 2\ H_2O$	0.265	L-Lysine • HCl	0.146
$MgSO_4$	0.09767	L-Methionine	0.03
KCl	0.4	L-Phenylalanine	0.066
$NaHCO_3$	3.7	L-Serine	0.042
NaCl	6.4	L-Threonine	0.095
$NaH_2PO_4 \cdot H_2O$	0.125	L-Tryptophan	0.016
D-Glucose	4.5	L-Tyrosine	0.07145
Phenol Red • Na	0.0159	L-Valine	0.094
$Fe(NO_3)_3 \cdot 9\ H_2O$	0.0001	Choline Chloride	0.004
L-Arginine	0.069	Folic Acid	0.004
L-Cystine	0.0542	*myo*-Inositol	0.0072
L-Glutamine	0.293	Niacinamide	0.004
Glycine	0.03	D-Pantothenic Acid • ½ Ca	0.004
L-Histidine	0.031	Riboflavin	0.0004
L-Isoleucine	0.105	Thiamin • HCl	0.004
L-Leucine	0.105		

Table II-7 Ingredients of in-house prepared RPMI media. The pH value of the media was adjusted to 7.33 (at r.t.).

Ingredient	Concentration [g/L]	Ingredient	Concentration [g/L]
$Ca(NO_3)_2 \cdot 4\ H_2O$	0.1	L-Asparagine	0.05
$MgSO_4 \cdot 6\ H_2O$	0.1	L-Aspartic Acid	0.02
KCl	0.4	L-Cystine	0.0652
$NaHCO_3$	2.0	L-Glutamic Acid	0.02
NaCl	6.0	L-Glutamine	0.293
$Na_2HPO_4 \cdot 2\ H_2O$	1.0	Glycine	0.01
D-Glucose	2.0	L-Histinde	0.015
Phenol Red • Na	0.0159	L-Hydroxyproline	0.02
D-Biotin	0.0002	L-Isoleucine	0.05
Choline Chloride	0.003	L-Leucine	0.05
Folic Acid	0.001	L-Lysine • HCl	0.04
myo-Inositol	0.035	L-Methionine	0.015
Niacinamide	0.001	L-Phenylalanine	0.015
p-Amino Benzoic Acid	0.001	L-Proline	0.02
D-Pantothenic Acid•½Ca	0.00025	L-Serine	0.03
Riboflavin	0.0002	L-Threonine	0.02
Thiamin • HCl	0.001	L-Tryptophan	0.005
Vitamin B12	0.000005	L-Tyrosine	0.02
L-Arginine	0.2	L-Valine	0.02

11.7.2 PLP-probe labeling in lysate

Cells were cultured as described under 11.3, seeded onto plates (152 cm²) and grown until 80-90% confluency. Cells were lysed in 1 mL lysis buffer (1% (w/v) sodium deoxycholate, 1% (v/v) NP-40 in PBS pH 7.4) containing protease/phosphatase inhibitors (applied at 1:100 (v/v), Halt cocktail, Thermo Fisher Scientific, Cat# 78446) for 30 min on ice. Afterwards, 1 mM maleimide (added from a 100 mM stock prepared fresh in H_2O) was added and lysate was incubated another 30 min on ice for **PL1P** samples. Cell debris was removed at 4 °C for 20 min (21,000 x *g*). After addition of 1 µM **PL1P** or **PL2P**, respectively, added from a 2 mM stock, prepared as described previously,[198] lysate was incubated for 1 h under rotation at 4 °C. 20 mM $NaBH_4$ were added from a 500 mM stock prepared fresh in 0.1 M NaOH and reduction carried out for 30 min at 4 °C. Cell debris was removed at 4 °C for 20 min (21,000 x *g*). Soluble proteins were precipitated by adding 4x volume of cold acetone (-80 °C) overnight at -20°C.

11.7.3 Click chemistry and sample preparation for MS

Precipitated proteins were pelleted by centrifugation (21,000 x g, 15 min, 4 °C) and washed with ice-cold MeOH (2x 1 mL), using sonication to resuspend the pellets between washes. Proteins were solubilized in 0.5 mL PBS containing 0.4% (w/v) SDS and protein concentrations were adjusted to 2 mg/mL (= 1 mg total per sample). Samples were subjected to CuAAC by adding 0.1 mM biotin-azide (5 µL of 10 mM stock in DMSO), 0.5 mM BTTAA ligand (25 µL of 10 mM stock in ddH$_2$O; Jena Bioscience, Cat# CLK-067-25), 1 mM CuSO$_4$ (10 µL of 50 mM stock in ddH$_2$O) and 2 mM sodium ascorbate (10 µL of 100 mM prepared fresh in ddH$_2$O) to each sample and incubating for 1 h at r. t. in the dark. For **PL1** labeling in HuH7, proteins were additionally incubated with 20 mM IAA (added from a 100x stock prepared fresh in H$_2$O) for 30 min in the dark previous to CuAAC. **PL3**-labeled samples were incubated with 0.2 mM EZ-Link™ Phosphine-PEG$_3$-Biotin (10 mM stock in DMSO; Thermo Fisher Scientific, Cat# 88901) for 4 h at 37 °C, followed by 20 h at 25 °C. Upon precipitation and washing (as described above), the protein pellets were re-solubilized in 0.5 mL PBS containing 0.4% SDS (w/v) and centrifuged (18,000 x g, 5 min, r.t.) prior to avidin bead enrichment. Enrichment, reduction, digestion, desalting, and sample preparation for MS was performed as described previously[198] with the exception that enrichment was performed for 1.5 h. In brief, proteins were added to pre-equilibrated beads, enriched under rotation and washed extensively to remove background binders. Samples were reduced and alkylated previous to pre-digest with Lys-C (Wako, Cat# 125-05061) for 2 h at 37 °C. Trypsin (Promega, Cat# V5111) digestion took place overnight at 37 °C and samples were desalted on C18 columns (Waters, Cat# WAT054960) afterwards. Peptides were evaporated *in vacuo* and re-dissolved in water containing 1% FA previous to LC-MS/MS.

11.7.4 Cross-link/co-IP

Cross-link/co-IP was performed using the DSSO cross-linker as described previously.[110] In brief, intact cells were cross-linked with 2 mM DSSO in PBS at 37 °C for 1 h. After quenching of the cross-linker cells were lysed, protein concentration was adjusted for all replicates, and pull-down was performed for 3 h at 4 °C under rotation. For PLPBP pull-down a rabbit-polyclonal anti-PLPBP antibody (1.02 mg/mL, Thermo Scientific, Cat# PA5-32036) was applied 1:100 and compared to equal amounts of isotype control to account for non-specific binding to antibody constant regions. After extensive washing, proteins were digested on-bead at 37 °C and overnight using trypsin (Promega, Cat# V5111). Desalting was performed using C18 stage-tips (SUPELCO, Cat# 66883-U), peptides were evacuated *in vacuo* and re-dissolved in water

containing 1% FA previous to LC-MS/MS analysis.

11.7.5 HepG2 siRNA transfection

HepG2 cells were seeded in 58 cm^2 dishes and grown until 80-90% confluency. A volume of 58 µL of a 5 µM targeting (ON-TARGETplus human PLPBP siRNA, SMARTpool, Dharmacon, Cat# L-008449-01-0005) or non-targeting siRNA (ON-TARGETplus non-targeting pool, Dharmacon, Cat# D-001810-10-05) stock were mixed with 1.102 mL of reduced-serum media (Opti-MEM, Thermo Scientific, Cat# 31985062) and incubated for 5 min at r.t. Further, 46.4 µL DharmaFECT 1 transfection reagent (Dharmacon, Cat# T-2001-01) were mixed with 1.114 mL reduced-serum media and incubated for 5 min at r.t. siRNA and transfection reagent mixes were combined and incubated for another 20 min at r.t. The mixture was filled up to 10 mL using full-media (RPMI with 10% FCS). Growth media was removed from the cells and 10 mL mixture were added per plate. 4 biological replicates were prepared for targeting and non-targeting siRNA, respectively. Transfection carried out for 48 h at 37 °C. Cells were washed once with cold PBS (4 °C) and were scraped to detach in 500 µL lysis buffer (PBS pH 7.4 supplemented with 1% (w/v) sodium deoxycholate and 1% (v/v) NP-40) per plate. Cells were incubated for 20 min on ice previous to centrifugation for 20 min at 21,000 x g and 4 °C. Protein concentrations were adjusted to 200 µg per replicate in 500 µL lysis buffer using a BCA assay and precipitated adding 4x volume of ice-cold acetone (-80 °C) and incubation at -20 °C overnight. Samples for western blot analysis (cf. Part II, Chapter 11.2.10) were taken from residual lysate.

11.7.6 PLPBP vector transfection

For whole proteome analysis, three million HepG2 cells were seeded per 58 cm^2 dish per replicate and let settle down overnight. 24 µg PLPBP-containing transfection vector (pcDNA3.1$^+$/C-(K)DYK; NM_007198 ORF clone, GenScript, Cat# OHu09065D) or 24 µg empty transfection vector (pcDNA3.1$^+$/C-(K)DYK; GeneScript, custom-made) were mixed with 1.26 mL Opti-MEM (Thermo Fisher Scientifc, Cat# 31985062) and incubated for 5 min at r.t. 60 µg Lipofectamin 2000 (Thermo Fisher Scientific, Cat# 11668030) were mixed with 1.44 mL Opti-MEM (Thermo Fisher Scientific, Cat# 31985062) and were incubated for 5 min at r.t. Afterwards both mixtures were unified and incubated another 20 min at r.t. Afterwards 3 mL were added directly to the cells covered with growth media per plate. Four biological replicates were prepared for PLPBP-containing vector and the empty vector control, respectively. Transfection carried out for 48 h at 37 °C. Cells were washed once with cold PBS (4 °C) and were scraped to

detach in 500 µL lysis buffer (PBS pH 7.4 supplemented with 1% (w/v) sodium deoxycholate and 1% (v/v) NP-40) per plate. Cells were incubated for 20 min on ice previous to centrifugation for 20 min at 21,000 x g and 4 °C. Protein concentrations were adjusted to 200 µg per replicate in 500 µL lysis buffer using a BCA assay and precipitated adding 4x volume of ice-cold acetone (-80 °C) and incubation at -20 °C overnight. Samples for western blot analysis (cf. Part II, Chapter 11.2.10) were taken from residual lysate.

For expression analysis of PLPBP WT and mutants *via* western blot, one million HEK293 cells were seeded per well in a 6-well plate. Volumes for transfection were adjusted according to reduced culture surface area. Thereby, 4 µg DNA were diluted in 250 µL Opti-MEM and 10 µg Lipofectamine 2000 were diluted in 250 µL Opti-MEM. Transfection carried out for 48 h at 37 °C. Afterwards, cells were washed once with cold PBS and then scraped to detach in PBS. Cells were lysed (PBS pH 7.4 supplemented with 1% (w/v) sodium deoxycholate and 1% (v/v) NP-40) for 20 min on ice previous to centrifugation for 20 min at 21,000 x g and 4 °C. Supernatant was diluted with 2x gel loading buffer previous to SDS-PAGE with subsequent western blot analysis (cf. Part II, Chapter 11.2.10).

11.7.7 Analysis of *E. coli yggS* knockout strain

E. coli BW25113 wild-type and yggS knockout were cultured as described under 11.4. Bacteria were harvested and washed with cold PBS (4 °C). Cells were resuspended to 1 mL OD_{600nm} of 20 in PBS and lysed by sonication. Cell debris was removed for 20 min at 21,000 x g and 4 °C. Protein concentrations were adjusted to 200 µg per replicate in 500 µL PBS using a BCA assay and precipitated adding 4x volume of ice-cold acetone (-80 °C) and incubation at -20 °C overnight. 4 replicates per condition were conducted.

11.7.8 Sample preparation for full-proteome analysis

Acetone precipitated proteins were pelletized (4 °C, 10 min, 21,000 x g) and washed twice with ice-cold MeOH (-80 °C), using sonication to resuspend the pellets between washes. Proteins were dissolved in 300 µL denaturation buffer (6 M urea, 2 M thiourea in 20 mM HEPES pH 7.5) and reduced for 1 h at 37 °C with 1 mM DTT added from a 500 mM stock prepared freshly in 20 mM HEPES pH 7.5. Proteins were alkylated with 5 mM IAA added from a 500 mM stock prepared freshly in 20 mM HEPES pH 7.5 for 30 min at 25 °C. IAA was quenched by adding 5 mM DTT from a 500 mM stock prepared freshly in 20 mM HEPES pH 7.5 and incubating another 30 min at 25 °C. Pre-digestion took place by adding Lys-C (1:80 (w/w), Wako, Cat# 125-05061) for 4 h at 37 °C. Samples were diluted 4x with 50 mM TEAB (prepared fresh in water

from a 1 M stock solution; Sigma, Cat# T7408) and incubated with trypsin (1:80 (w/w), sequencing grade, modified; Promega, Cat# V5111) overnight at 37 °C. In order to quench the reaction, FA was added to a final amount of 1% (v/v). Samples were desalted using C18 columns (Waters, Cat# WAT054960) and prepared for LC-MS/MS analysis as described previously.[198]

11.7.9 LC-MS/MS analysis of proteomic samples

Samples were analyzed *via* LC-MS/MS using a UltiMate 3000 nano HPLC system (Thermo Fisher Scientific) equipped with an Acclaim C18 PepMap100 75 μm ID x 2 cm trap (Thermo Fisher Scientific, Cat# ES803A) and an Acclaim PepMap RSLC C18 separation column (75 μm ID x 50 cm; Thermo Fisher Scientific, Cat# 164535) coupled to an EASY-source equipped Thermo Fisher LTQ Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). Samples were loaded onto the trap column at a flow rate of 5 μL/min with aqueous 0.1% trifluoroacetic acid (TFA) and then transferred onto the separation column at 0.3 μL/min. Buffers for the nano-chromatography pump were aqueous 0.1% FA (buffer A) and 0.1% FA in ACN (buffer B). Samples were separated using a gradient raising buffer B from 5 to 22% in 112 min, followed by a buffer B increase to 32% within 10 min. Buffer B content was further raised to 90% within the next 10 min and held another 10 min at 90%. Subsequently buffer B was decreased to 5% and held until the end of the run (total: 152 min). During sample separation MS full scans were performed at 120,000 resolution in the orbitrap with quadrupole isolation. The MS instrument was operated in a 3 s top speed data dependent mode. The scan range was set from 300 to 1,500 m/z with 60% RF lens amplitude. The AGC target was set to 2.0e5, the maximum ion injection time was 50 ms (co-IP and labeling), or 35 ms (global proteome analysis) and internal calibration was performed using the lock mass option. Peptides with intensity higher than 5.0e3 and charge states 2-7 were fragmented with HCD (30%). MS^2 scans were recorded in the ion trap operating in rapid mode. The isolation window was set to 1.6 m/z and the AGC target to 1.0e4 with maximum injection time of 100 ms. Ions were injected for all available parallelizable time. Dynamic exclusion time was 60 s with 10 ppm low and high mass tolerance. Samples for HCT116 **PL1** labeling were analyzed on a UltiMate 3000 nano HPLC system (Thermo Fisher Scientific) equipped with an Acclaim C18 PepMap100 75 μm ID x 2 cm trap (Thermo Fisher Scientific, Cat# ES803A) and an Acclaim PepMap RSLC C18 separation column (75 μm ID x 50 cm; Thermo Fisher Scientific, Cat# 164535) coupled to an EASY-source equipped Q Exactive Plus mass spectrometer (Thermo Fisher Scientific). Sample loading and separation was performed as for measurements on Fusion instrument. Full scans were performed at 140,000 resolution over a scan range of 300-1,500 m/z using an AGC target of 3e6 and a maximum ion

injection time of 80 ms. For data-dependent MS^2 scans the AGC target was set to 1e5 at 17,500 resolution and 1.6 m/z isolation window. Fragmentation was performed at a normalized collision energy of 27%. The maximum injection time was set to 100 ms. Q Exactive Plus was operating in a Top$N$ = 10 mode. Dynamic exclusion time was set to 60 s.

11.7.10 Statistical analysis of proteomics data

MS raw data were searched with MaxQuant software[124] (version 1.6.0.1) and default settings (with the exceptions that label-free quantification and match between runs were activated). All replicates from one condition (*e.g.* **PL1** treated samples) were defined as one fraction. Searches were performed against a Uniprot database of *Homo sapiens* proteome (taxon identifier: 9606, reference reviewed and unreviewed, downloaded on 25.12.2018) and *E. coli* K12 (taxon identifier: 83333, reference reviewed and unreviewed, downloaded on 25.06.2019). Resulting data were further analyzed using Perseus software version 1.6.0.0.[125] The rows were filtered (only identified by site, potential contaminant, reverse) and $\log_2$ transformed. Biological replicates were grouped, filtered for 100% valid values (enriched samples and co-IP), or for 3 out 4 valid values (full-proteome analysis) in at least one group and missing values were imputed for total matrix using default settings. A both sided, two-sample Student's t-test was performed and derived *p*-values were corrected for multiple testing by the method of Benjamini and Hochberg with a significance level of $p = 0.05$. Volcano plots were generated by plotting $\log_2$ (fold change of different conditions) against $-\log_{10}$ (*p*-value). **PLP**-dependent proteins (gene ontology (GO)-term: 0030170; **PLP**-binding) and mitochondrial proteins (GO term 0005739: mitochondrion) were identified with the help of a GO annotation file for *H. sapiens* or *E. coli*, respectively, submitted on 26.06.2018 or 25.06.2019, respectively.[127] To evaluate overrepresentation of KEGG terms,[305] a STRING analysis[280] with significantly overexpressed proteins was performed.

11.8 Metabolomics

11.8.1 Metabolomics sample preparation

HEK293 cells were cultured as described under 11.3. For time-point zero measurements, cells were seeded into 148 cm² dishes in chemically defined media complemented with 5 µM **PL** (added fresh from a 50 mM stock in DMSO). For later time points, cells were transferred into chemically defined media complemented with 10 µM **PL1** (added fresh from a 100 mM stock in DMSO) and cultured till the desired passaging numbers previous to seeding them onto 148 cm²

dishes. Cells were grown to 90% confluence. Quadruplicates per condition were prepared. After that, media was removed and cells were scraped to detach in 10 mL cold (4 °C) 0.9% NaCl (m/v) in water per plate. Cells were harvested at 600 x g for 2 min. Supernatant was removed and cells were lysed in 1 mL ice-cold MeOH (-80 °C, containing 1 µM trimethoprim as internal standard) overnight at -20 °C. Samples were centrifuged for 20 min at 21,000 x g and 4 °C. Supernatant was dried *in vacuo* and metabolites were re-dissolved in 90 µL 50/50 Vol-% MeOH/H_2O containing 1% FA for LC-MS/MS measurement.

11.8.2 LC-MS/MS analysis of metabolomics samples

Metabolic profiling *via* LC-MS/MS analysis was carried out on an UltiMate 3000 RSLC system (Thermo Fisher Scientific) coupled to an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific). Chromatographic separation was performed using an Accucore HILIC column (150 x 2.1 mm, 2.6 µm, 80 Å) (Thermo Fisher Scientific) at 40 °C and 80 mM NH_4OAc in 100% H_2O pH = 6.8 (A) and 100% ACN (B). After 5 min pre-equilibration with 95% B, samples were eluted with a linear gradient from 95% to 0% B over 30 min at a flow rate of 250 µL/min followed by 8 min re-equilibration with 95% B. Mass spectrometric measurements were accomplished in positive ion mode (HESI-II source, Thermo Fisher Scientific) with following source parameters: capillary voltage 4.5 kV, tube lens 40 V, vaporizer temperature 43 °C, sheath gas 50 L/h, aux gas 10 L/h, capillary voltage 12 V and capillary temperature 320 °C. Full scan measurements were recorded between 50 – 1000 m/z in profile mode at 30,000 resolution in the orbitrap. Parent ions of interest (**PL** m/z 168.06 and **PLP** m/z 248.03) were isolated (isolation width 1.0 m/z) subjected to collision-induced dissociation (normalized collision energy 35 V) in the multiple reaction monitoring (MRM) mode and most intense daughter ions (**PL/PLP** m/z 150.05) were used for quantification (selected ion monitoring (SIM) width 10 m/z) at a resolution of 7,500 in the orbitrap. Prior to measurement 8 pooled quality control (QC) samples were injected to equilibrate the column. In order to take metabolic degradation over time into account, sample order was randomized. Overall instrument performance was monitored using trimethoprim.

11.8.3 Statistical analysis of metabolomics data

Raw data were processed with Xcalibur Quan Browser (Thermo Fisher Scientific) using Genesis algorithm and manual integration mode. Intensities of the corresponding daughter ions were extracted and further processed using Excel. Finally, daughter ion intensities were plotted for **PL** and **PLP** for respective time points using Prism 6 (GraphPad).

11.9 Binding site identification of PLPBP

11.9.1 Sample preparation and LC-MS/MS analysis

Recombinant PLPBP (4 µg, 50 µL) was incubated with 5 equivalents of **PLP** in 50 mM HEPES pH 7.4, 100 mM KCl for 30 min at r.t., and was subsequently reduced with 10 mM NaBH$_4$ (2 µL of 250 mM stock prepared fresh in 0.1 M NaOH) for 30 min at r.t. The protein was precipitated adding ice-cold acetone (4x volume) and incubated at -20 °C overnight. Precipitated protein was pelleted by centrifugation (18,000 x g, 15 min, 4°C) and washed with ice-cold MeOH (2x 0.2 mL), using sonication to resuspend the pellets between washes. Protein was resuspended in 100 µL denaturation buffer (6 M urea, 2 M thiourea in 20 mM HEPES buffer pH 7.5) and reduced with 5 mM TCEP (1 µL of 500 mM stock in 20 mM HEPES buffer pH 7.5) for 1 h at 37 °C. Proteins were alkylated using 10 mM IAA (2 µL of 500 mM stock in 20 mM HEPES buffer pH 7.5) for 30 min at 25 °C, followed by quenching with 10 mM DTT (2 µL of 500 mM stock in 20 mM HEPES buffer pH 7.5) for 30 min at 25 °C. Samples were diluted with 300 µL 50 mM TEAB (prepared fresh in water from a 1 M stock solution; Sigma, Cat# T7408) previous to addition of trypsin (0.4 µL from 0.5 µg/µL stock, Sequencing Grade, Promega, Cat# V5111) and overnight digestion at 37 °C. Desalting with C18 cartridges (Waters, Cat# WAT054960) and preparation for LC-MS/MS analysis was performed as described previously.[198] LC-MS/MS analysis on Fusion mass spectrometer was performed as for proteomic samples (probe labeling and co-IPs) with the exception that the HPLC gradient was shortened to 62 min to account for reduced peptide amount. Samples were separated using a gradient raising buffer B from 5 to 28% in 37 min, followed by a buffer B increase to 35% within 5 min. Buffer B content was further raised to 90% within the next 0.1 min and held another 10 min at 90%. Subsequently buffer B was decreased to 5% and held until the end of the run.

11.9.2 Statistical analysis of MS binding site data

MS raw files were analyzed with MaxQuant software (version 1.6.0.1) as described in the proteomics section. For **PLP** binding site identification, the **PLP** moiety (+ 231.02966) at lysine was additionally set as a variable modification. As we expect one **PLP** modification site in PLPBP, we selected the site with the highest confidence based on best PEP, score, and manual evaluation of MS/MS spectra from MaxQuant.

III - REFERENCES

1. Leitner, A., Faini, M., Stengel, F. & Aebersold, R. Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends in Biochemical Sciences* **41**, 20–32 (2016).

2. Leitner, A. Cross-linking and other structural proteomics techniques: How chemistry is enabling mass spectrometry applications in structural biology. *Chem. Sci.* **7**, 4792–4803 (2016).

3. Leitner, A. *et al.* Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry, and Bioinformatics. *Mol. Cell. Proteomics* **9**, 1634–1649 (2010).

4. Yılmaz, Ş. *et al.* Cross-linked peptide identification: A computational forest of algorithms. *Mass Spectrometry Reviews* **37**, 738–749 (2018).

5. Leitner, A., Walzthoeni, T. & Aebersold, R. Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat. Protoc.* **9**, 120–37 (2014).

6. Leitner, A. *et al.* Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9455–60 (2014).

7. Iacobucci, C. *et al.* Carboxyl-Photo-Reactive MS-Cleavable Cross-Linkers: Unveiling a Hidden Aspect of Diazirine-Based Reagents. *Anal. Chem.* **90**, 2805–2809 (2018).

8. Giese, S. H., Belsom, A. & Rappsilber, J. Optimized fragmentation regime for diazirine photo-cross-linked peptides. *Analytical Chemistry* **88**, 8239–8247 (2016).

9. Yang, B. *et al.* Proximity-enhanced SuFEx chemical cross-linker for specific and multitargeting cross-linking mass spectrometry. *Proc. Natl. Acad. Sci.* **115**, 11162–11167 (2018).

10. Rappsilber, J. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173**, 530–540 (2011).

11. Schilling, B., Row, R. H., Gibson, B. W., Guo, X. & Young, M. M. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. *J. Am. Soc. Mass Spectrom.* **14**, 834–850 (2003).

12. Tosi, A. *et al.* XStructure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* **154**, 1207 (2013).

13. Bui, K. H. *et al.* Integrated structural analysis of the human nuclear pore complex scaffold. *Cell* **155**, 1233–1243 (2013).

14. Kahraman, A. *et al.* Cross-Link Guided Molecular Modeling with ROSETTA. *PLoS One* **8**, (2013).

15. Kosinski, J. *et al.* Xlink analyzer: Software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *J. Struct. Biol.* **189**, 177–183 (2015).

16. Chen, Z. A. *et al.* Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29**, 717–726 (2010).

17. Goetze, M. *et al.* StavroX-A software for analyzing crosslinked products in protein

interaction studies. *J. Am. Soc. Mass Spectrom.* **23**, 76–87 (2012).

18. Liu, F., Rijkers, D. T. S., Post, H. & Heck, A. J. R. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **12**, 1179–1184 (2015).

19. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737 (2003).

20. Herzog, F. *et al.* Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science (80-.).* **337**, 1348–1352 (2012).

21. Leitner, A. *et al.* The molecular architecture of the eukaryotic chaperonin TRiC/CCT. *Structure* **20**, 814–825 (2012).

22. Zhang, H. *et al.* Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry. *Mol. Cell. Proteomics* **8**, 409–420 (2009).

23. Chavez, J. D. *et al.* A general method for targeted quantitative cross-linking mass spectrometry. *PLoS One* **11**, (2016).

24. Navare, A. T. *et al.* Probing the protein interaction network of Pseudomonas aeruginosa cells by chemical cross-linking mass spectrometry. *Structure* **23**, 762–73 (2015).

25. Wu, X. *et al.* In vivo protein interaction network analysis reveals porin-localized antibiotic inactivation in Acinetobacter baumannii strain AB5075. *Nat. Commun.* **7**, (2016).

26. Zhang, H. *et al.* In vivo identification of the outer membrane protein OmcA-MtrC interaction network in shewanella oneidensis MR-1 cells using novel hydrophobic chemical cross-linkers. *J. Proteome Res.* **7**, 1712–1720 (2008).

27. Zheng, C. *et al.* Cross-linking measurements of in vivo protein complex topologies. *Mol Cell Proteomics* **10**, M110.006841 (2011).

28. De Jong, L. *et al.* In-Culture Cross-Linking of Bacterial Cells Reveals Large-Scale Dynamic Protein-Protein Interactions at the Peptide Level. *J. Proteome Res.* **16**, 2457–2471 (2017).

29. Schweppe, D. K. *et al.* Mitochondrial protein interactome elucidated by chemical cross-linking mass spectrometry. *Proc. Natl. Acad. Sci.* **114**, 1732–1737 (2017).

30. Lössl, P., Liu, F., Balaban, R. S., Heck, A. J. R. & Rabbitts, B. M. The interactome of intact mitochondria by cross-linking mass spectrometry provides evidence for coexisting respiratory supercomplexes. *Mol. Cell. Proteomics* **17**, 216–232 (2017).

31. Schweppe, D. K. *et al.* Host-Microbe Protein Interactions during Bacterial Infection. *Chem. Biol.* **22**, 1521–1530 (2015).

32. Chavez, J. D. *et al.* Chemical Crosslinking Mass Spectrometry Analysis of Protein Conformations and Supercomplexes in Heart Tissue. *Cell Syst.* **6**, 136-141.e5 (2018).

33. O'Reilly, F. J. & Rappsilber, J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nature Structural and Molecular Biology* **25**, 1000–1008 (2018).

34.   Selbach, M. & Mann, M. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nat. Methods* **3**, 981–983 (2006).

35.   Falsone, S. F., Gesslbauer, B. & Kungl, A. J. Coimmunoprecipation and proteomic analyses. in *Methods in Molecular Biology* **439**, 291–308 (2008).

36.   Ceriani, M. F. Coimmunoprecipitation on Drosophila cells in culture. *Methods Mol Biol* **362**, 423–427 (2007).

37.   Keilhauer, E. C., Hein, M. Y. & Mann, M. Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS). *Mol. Cell. Proteomics* **14**, 120–135 (2014).

38.   Yu, C. & Huang, L. Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Analytical Chemistry* **90**, 144–165 (2018).

39.   Wang, X. *et al.* Molecular Details Underlying Dynamic Structures and Regulation of the Human 26S Proteasome. *Mol. Cell. Proteomics* **16**, 840–854 (2017).

40.   Yu, C. *et al.* Characterization of Dynamic UbR-Proteasome Subcomplexes by *In vivo* Cross-linking (X) Assisted Bimolecular Tandem Affinity Purification (XBAP) and Label-free Quantitation. *Mol. Cell. Proteomics* **15**, 2279–2292 (2016).

41.   Guerrero, C., Tagwerker, C., Kaiser, P. & Huang, L. An Integrated Mass Spectrometry-based Proteomic Approach Quantitative Analysis of Tandem Affinity-purified in vivo Cross-linked Protein Complexes (qtax) to Decipher the 26 s Proteasome-interacting Network. *Mol. Cell. Proteomics* **5**, 366–378 (2006).

42.   Nittis, T. *et al.* Revealing Novel Telomere Proteins Using *in Vivo* Cross-linking, Tandem Affinity Purification, and Label-free Quantitative LC-FTICR-MS. *Mol. Cell. Proteomics* **9**, 1144–1156 (2010).

43.   Bernhard, O. K., Sheil, M. M. & Cunningham, A. L. Lateral Membrane Protein Associations of CD4 in Lymphoid Cells Detected by Cross-Linking and Mass Spectrometry [†]. *Biochemistry* **43**, 256–264 (2004).

44.   Wu, X. *et al.* Dynamic Proteome Response of Pseudomonas aeruginosa to Tobramycin Antibiotic Treatment . *Mol. Cell. Proteomics* **14**, 2126–2137 (2015).

45.   Chen, Z. A., Fischer, L., Cox, J. & Rappsilber, J. Quantitative cross-linking/mass spectrometry using isotope-labeled cross-linkers and MaxQuant. **23**, 145–150 (2016).

46.   Walzthoeni, T. *et al.* XTract: Software for characterizing conformational changes of protein complexes by quantitative cross-linking mass spectrometry. *Nat. Methods* **12**, 1185–1190 (2015).

47.   Tang, X. & Bruce, J. E. A new cross-linking strategy: protein interaction reporter (PIR) technology for protein–protein interaction studies. *Mol. Biosyst.* **6**, 939 (2010).

48.   Tang, X., Munske, G. R., Siems, W. F. & Bruce, J. E. Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. *Anal. Chem.* **77**, 311–318 (2005).

49.   Zhong, X. *et al.* Large-Scale and Targeted Quantitative Cross-Linking MS Using Isotope-Labeled Protein Interaction Reporter (PIR) Cross-Linkers. *J. Proteome Res.* **16**, 720–727

(2017).

50.   Leitner, A. *et al.* Expanding the Chemical Cross-Linking Toolbox by the Use of Multiple Proteases and Enrichment by Size Exclusion Chromatography. *Mol. Cell. Proteomics* **11**, M111.014126 (2012).

51.   Fritzsche, R., Ihling, C. H., Götze, M. & Sinz, A. Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun. Mass Spectrom.* **26**, 653–658 (2012).

52.   Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-Triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **67**, 3057–3064 (2002).

53.   Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)- Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chemie Int. Ed.* **41**, 2596–2599 (2002).

54.   Burke, A. M. *et al.* Synthesis of two new enrichable and MS-cleavable cross-linkers to define protein-protein interactions by mass spectrometry. *Org. Biomol. Chem.* **13**, 5030–5037 (2015).

55.   Huang, L. *et al.* A New in Vivo Cross-linking Mass Spectrometry Platform to Define Protein–Protein Interactions in Living Cells . *Mol. Cell. Proteomics* **13**, 3533–3543 (2014).

56.   Giese, S. H., Fischer, L. & Rappsilber, J. A Study into the Collision-induced Dissociation (CID) Behavior of Cross-Linked Peptides. *Mol. Cell. Proteomics* **15**, 1094–1104 (2015).

57.   Iacobucci, C. & Sinz, A. To Be or Not to Be? Five Guidelines to Avoid Misassignments in Cross-Linking/Mass Spectrometry. *Anal. Chem.* **89**, 7832–7835 (2017).

58.   Kao, A. *et al.* Development of a Novel Cross-linking Strategy for Fast and Accurate Identification of Cross-linked Peptides of Protein Complexes. *Mol. Cell. Proteomics* **10**, M110.002212 (2011).

59.   Gutierrez, C. B. *et al.* Developing an acidic residue reactive and sulfoxide-containing MS-cleavable homobifunctional cross-linker for probing protein-protein interactions. *Anal. Chem.* **88**, 8315–8322 (2016).

60.   Müller, M. Q., Dreiocker, F., Ihling, C. H., Schäfer, M. & Sinz, A. Cleavable cross-linker for protein structure analysis: Reliable identification of cross-linking products by tandem MS. *Anal. Chem.* **82**, 6958–6968 (2010).

61.   Back, J. W. *et al.* A new crosslinker for mass spectrometric analysis of the quaternary structure of protein complexes. *J. Am. Soc. Mass Spectrom.* **12**, 222–227 (2001).

62.   Soderblom, E. J. & Goshe, M. B. Collision-induced dissociative chemical cross-linking reagents and methodology: Applications to protein structural characterization using tandem mass spectrometry analysis. *Anal. Chem.* **78**, 8059–8068 (2006).

63.   Liu, F., Lössl, P., Scheltema, R., Viner, R. & Heck, A. J. R. Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. *Nat. Commun.* **8**, (2017).

64.   Yu, A. Y. H. & Houry, W. A. ClpP: A distinctive family of cylindrical energy-dependent

serine proteases. *FEBS Letters* **581**, 3749–3757 (2007).

65. Katayama-Fujimura, Y., Gottesman, S. & Maurizi, M. R. A multiple-component, ATP-dependent protease from Escherichia coli. *J. Biol. Chem.* **262**, 4477–4485 (1987).

66. Bewley, M. C., Graziano, V., Griffin, K. & Flanagan, J. M. The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *J. Struct. Biol.* **153**, 113–128 (2006).

67. Baker, T. A. & Sauer, R. T. ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta - Molecular Cell Research* **1823**, 15–28 (2012).

68. Gillet, R. & Felden, B. Emerging views on tmRNA-mediated protein tagging and ribosome rescue. *Molecular Microbiology* **42**, 879–885 (2001).

69. Gottesman, S., Roche, E., Zhou, Y. N. & Sauer, R. T. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev.* **12**, 1338–1347 (1998).

70. Weber-Ban, E. U., Reid, B. G., Miranker, A. D. & Horwich, A. L. Global unfolding of a substrate protein by the HSP100 chaperone ClpA. *Nature* **401**, 90–93 (1999).

71. Truscott, K. N., Lowth, B. R., Strack, P. R. & Dougan, D. A. Diverse functions of mitochondrial AAA+ proteins: protein activation, disaggregation, and degradation. *Biochem Cell Biol* **88**, 97–108 (2010).

72. Kang, S. G., Dimitrova, M. N., Ortega, J., Ginsburg, A. & Maurizi, M. R. Human mitochondrial ClpP is a stable heptamer that assembles into a tetradecamer in the presence of ClpX. *J. Biol. Chem.* **280**, 35424–35432 (2005).

73. Goard, C. A. & Schimmer, A. D. Mitochondrial matrix proteases as novel therapeutic targets in malignancy. *Oncogene* **33**, 2690–2699 (2014).

74. Puccio, H. & Kœnig, M. Friedreich ataxia: A paradigm for mitochondrial diseases. *Current Opinion in Genetics and Development* **12**, 272–277 (2002).

75. Jenkinson, E. M. *et al.* Perrault syndrome is caused by recessive mutations in CLPP, encoding a mitochondrial ATP-dependent chambered protease. *Am. J. Hum. Genet.* **92**, 605–613 (2013).

76. Gispert, S. *et al.* Loss of mitochondrial peptidase clpp leads to infertility, hearing loss plus growth retardation via accumulation of CLPX, mtDNA and inflammatory factors. *Hum. Mol. Genet.* **22**, 4871–4887 (2013).

77. Cole, A. *et al.* Inhibition of the Mitochondrial Protease ClpP as a Therapeutic Strategy for Human Acute Myeloid Leukemia. *Cancer Cell* **27**, 864–876 (2015).

78. Deepa, S. S. *et al.* Down-regulation of the mitochondrial matrix peptidase ClpP in muscle cells causes mitochondrial dysfunction and decreases cell proliferation. *Free Radic. Biol. Med.* **91**, 281–292 (2016).

79. Seo, J. H. *et al.* The Mitochondrial Unfoldase-Peptidase Complex ClpXP Controls Bioenergetics Stress and Metastasis. *PLoS Biol.* **14**, (2016).

80. Wojtyra, U. A., Thibault, G., Tuite, A. & Houry, W. A. The N-terminal zinc binding domain

of ClpX is a dimerization domain that modulates the chaperone function. *J. Biol. Chem.* **278**, 48981–48990 (2003).

81.　Zeiler, E. *et al.* Vibralactone as a tool to study the activity and structure of the ClpP1P2 complex from listeria monocytogenes. *Angew. Chemie - Int. Ed.* **50**, 11001–11004 (2011).

82.　Zeiler, E. *et al.* Structural and functional insights into caseinolytic proteases reveal an unprecedented regulation principle of their catalytic triad. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11302–7 (2013).

83.　Dahmen, M., Vielberg, M. T., Groll, M. & Sieber, S. A. Structure and mechanism of the caseinolytic protease ClpP1/2 heterocomplex from listeria monocytogenes. *Angew. Chemie - Int. Ed.* **54**, 3598–3602 (2015).

84.　Frees, D., Savijoki, K., Varmanen, P. & Ingmer, H. Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Molecular Microbiology* **63**, 1285–1295 (2007).

85.　Tomoyasu, T. *et al.* The ClpXP ATP-dependent protease regulates flagellum synthesis in Salmonella enterica serovar Typhimurium. *J. Bacteriol.* **184**, 645–653 (2002).

86.　Porankiewlcz, J., Wang, J. & Clarke, A. K. New insights into the ATP-dependent Clp protease: Escherichia coli and beyond. *Molecular Microbiology* **32**, 449–458 (1999).

87.　Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. & Berche, P. The ClpP serine protease is essential for the intracellular parasitism and virulence of Listeria monocytogenes. *Mol. Microbiol.* **35**, 1286–1294 (2000).

88.　Böttcher, T. & Sieber, S. A. β-lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of Staphylococcus aureus. *J. Am. Chem. Soc.* **130**, 14400–14401 (2008).

89.　Krysiak, J. *et al.* Quantitative Map of β-Lactone-Induced Virulence Regulation. *J. Proteome Res.* **16**, 1180–1192 (2017).

90.　Böttcher, T. & Sieber, S. A. Structurally Refined β-Lactones as Potent Inhibitors of Devastating Bacterial Virulence Factors. *ChemBioChem* **10**, 663–666 (2009).

91.　Balogh, D. *et al.* Insights into ClpXP proteolysis: heterooligomerization and partial deactivation enhance chaperone affinity and substrate turnover in Listeria monocytogenes. *Chem. Sci.* **8**, 1592–1600 (2017).

92.　Hackl, M. W. *et al.* Phenyl Esters Are Potent Inhibitors of Caseinolytic Protease P and Reveal a Stereogenic Switch for Deoligomerization. *J. Am. Chem. Soc.* **137**, 8475–8483 (2015).

93.　Pahl, A. *et al.* Reversible Inhibitors Arrest ClpP in a Defined Conformational State that Can Be Revoked by ClpX Association. *Angew. Chemie - Int. Ed.* **54**, 15892–15896 (2015).

94.　Zeiler, E., Korotkov, V. S., Lorenz-Baath, K., Böttcher, T. & Sieber, S. A. Development and characterization of improved β-lactone-based anti-virulence drugs targeting ClpP. *Bioorganic Med. Chem.* **20**, 583–591 (2012).

95.　Wang, J., Hartling, J. A. & Flanagan, J. M. The structure of ClpP at 2.3 Å resolution

suggests a model for ATP- dependent proteolysis. *Cell* **91**, 447–456 (1997).

96.   Zhang, J. *et al.* Structural switching of Staphylococcus aureus Clp protease: A key to understanding protease dynamics. *J. Biol. Chem.* **286**, 37590–37601 (2011).

97.   Ingvarsson, H. *et al.* Insights into the inter-ring plasticity of caseinolytic proteases from the X-ray structure of Mycobacterium tuberculosis ClpP1. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 249–259 (2007).

98.   Kang, S. G., Maurizi, M. R., Thompson, M., Mueser, T. & Ahvazi, B. Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *J. Struct. Biol.* **148**, 338–352 (2004).

99.   Kim, D. Y. & Kim, K. K. Crystal Structure of ClpX Molecular Chaperone from Helicobacter pylori. *J. Biol. Chem.* **278**, 50664–50670 (2003).

100.   Glynn, S. E., Martin, A., Nager, A. R., Baker, T. A. & Sauer, R. T. Structures of Asymmetric ClpX Hexamers Reveal Nucleotide-Dependent Motions in a AAA+ Protein-Unfolding Machine. *Cell* **139**, 744–756 (2009).

101.   Kim, Y. I. *et al.* Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nat. Struct. Biol.* **8**, 230–233 (2001).

102.   Martin, A., Baker, T. A. & Sauer, R. T. Distinct Static and Dynamic Interactions Control ATPase-Peptidase Communication in a AAA+ Protease. *Mol. Cell* **27**, 41–52 (2007).

103.   Frees, D. *et al.* New insights into staphylococcus aureus stress tolerance and virulence regulation from an analysis of the role of the ClpP protease in the strains Newman, COL, and SA564. *J. Proteome Res.* **11**, 95–108 (2012).

104.   Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T. & Baker, T. A. Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell* **11**, 671–683 (2003).

105.   Neher, S. B. *et al.* Proteomic Profiling of ClpXP Substrates after DNA Damage Reveals Extensive Instability within SOS Regulon. *Mol. Cell* **22**, 193–204 (2006).

106.   Bhat, N. H., Vass, R. H., Stoddard, P. R., Shin, D. K. & Chien, P. Identification of ClpP substrates in Caulobacter crescentus reveals a role for regulated proteolysis in bacterial development. *Mol. Microbiol.* **88**, 1083–1092 (2013).

107.   Feng, J. *et al.* Trapping and proteomic identification of cellular substrates of the ClpP protease in staphylococcus aureus. *J. Proteome Res.* **12**, 547–558 (2013).

108.   Szczepanowska, K. *et al.* CLPP coordinates mitoribosomal assembly through the regulation of ERAL1 levels. *EMBO J.* **35**, 2566–2583 (2016).

109.   Fischer, F., Langer, J. D. & Osiewacz, H. D. Identification of potential mitochondrial CLPXP protease interactors and substrates suggests its central role in energy metabolism. *Sci. Rep.* **5**, (2015).

110.   Fux, A., Korotkov, V. S., Schneider, M., Antes, I. & Sieber, S. A. Chemical Cross-Linking Enables Drafting ClpXP Proximity Maps and Taking Snapshots of In Situ Interaction Networks. *Cell Chem. Biol.* (2018). doi:10.1016/j.chembiol.2018.10.007

111. Cravatt, B. F. & Sorensen, E. J. Chemical strategies for the global analysis of protein function. *Current Opinion in Chemical Biology* **4**, 663–668 (2000).

112. Adam, G. C., Sorensen, E. J. & Cravatt, B. F. Chemical Strategies for Functional Proteomics. *Mol. Cell. Proteomics* **1**, 781–790 (2003).

113. Hoopmann, M. R. *et al.* Kojak: Efficient analysis of chemically cross-linked protein complexes. *J. Proteome Res.* **14**, 2190–2198 (2015).

114. Gersch, M., List, A., Groll, M. & Sieber, S. A. Insights into structural network responsible for oligomerization and activity of bacterial virulence regulator caseinolytic protease P (ClpP) protein. *J. Biol. Chem.* **287**, 9484–9494 (2012).

115. Gribun, A. *et al.* The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation. *J. Biol. Chem.* **280**, 16185–16196 (2005).

116. Sprangers, R., Gribun, A., Hwang, P. M., Houry, W. A. & Kay, L. E. Quantitative NMR spectroscopy of supramolecular complexes: Dynamic side pores in ClpP are important for product release. *PNAS* **102**, 16678–16683 (2005).

117. Lowth, B. R. *et al.* Substrate recognition and processing by a Walker B mutant of the human mitochondrial AAA+ protein CLPX. *J. Struct. Biol.* **179**, 193–201 (2012).

118. Stinson, B. M. *et al.* Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. *Cell* **153**, 628–639 (2013).

119. Malik, I. T. & Brötz-Oesterhelt, H. Conformational control of the bacterial Clp protease by natural product antibiotics. *Nat. Prod. Rep.* **34**, 815–831 (2017).

120. Yu, C. *et al.* Characterization of Dynamic UbR-Proteasome Subcomplexes by *In vivo* Cross-linking (X) Assisted Bimolecular Tandem Affinity Purification (XBAP) and Label-free Quantitation. *Mol. Cell. Proteomics* **15**, 2279–2292 (2016).

121. Huang, B. X. & Kim, H. Y. Effective Identification of Akt Interacting Proteins by Two-Step Chemical Crosslinking, Co-Immunoprecipition and Mass Spectrometry. *PLoS One* **8**, (2013).

122. Roux, K. J., Kim, D. I. & Burke, B. BioID: A screen for protein-protein interactions. *Curr. Protoc. Protein Sci.* (2013). doi:10.1002/0471140864.ps1923s74

123. Cox, J. *et al.* Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).

124. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

125. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731–740 (2016).

126. Carbon, S. *et al.* The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* (2018). doi:10.1093/nar/gky1055

127. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. *Nature Genetics* **25**, 25–29 (2000).

128. Matsushima, Y. *et al.* Drosophila protease ClpXP specifically degrades DmLRPPRC1 controlling mitochondrial mRNA and translation. *Sci. Rep.* **7**, (2017).

129. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

130. Quirós, P. M. *et al.* Multi-omics analysis identifies ATF4 as a key regulator of the mitochondrial stress response in mammals. *J. Cell Biol.* **216**, 2027–2045 (2017).

131. Falkevall, A. *et al.* Degradation of the amyloid β-protein by the novel mitochondrial peptidasome, PreP. *J. Biol. Chem.* **281**, 29096–29104 (2006).

132. Jobling, R. K. *et al.* PMPCA mutations cause abnormal mitochondrial protein processing in patients with non-progressive cerebellar ataxia. *Brain* **138**, 1505–1517 (2015).

133. Pryde, K. R., Taanman, J. W. & Schapira, A. H. A LON-ClpP Proteolytic Axis Degrades Complex I to Extinguish ROS Production in Depolarized Mitochondria. *Cell Rep.* **17**, 2522–2531 (2016).

134. Altieri, D. C., Stein, G. S., Lian, J. B. & Languino, L. R. TRAP-1, the mitochondrial Hsp90. *Biochimica et Biophysica Acta - Molecular Cell Research* **1823**, 767–773 (2012).

135. Iyer, R. K., Kim, H. K., Tsoa, R. W., Grody, W. W. & Cederbaum, S. D. Cloning and characterization of human agmatinase. *Mol Genet Metab* **75**, 209–18 (2002).

136. Kim, H. I. *et al.* Ornithine decarboxylase as a therapeutic target for endometrial cancer. *PLoS One* 1–18 (2017).

137. Free, R. B., Hazelwood, L. A. & Sibley, D. R. Identifying novel protein-protein interactions using co-immunoprecipitation and mass spectroscopy. *Current Protocols in Neuroscience* (2009). doi:10.1002/0471142301.ns0528s46

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

139. Wang, N., Gottesman, S., Willingham, M. C., Gottesman, M. M. & Maurizi, M. R. A human mitochondrial ATP-dependent protease that is highly homologous to bacterial Lon protease. *Proc. Natl. Acad. Sci.* **90**, 11247–11251 (2006).

140. Rep, M. *et al.* Promotion of mitochondrial membrane complex assembly by a proteolytically inactive yeast Lon. *Science (80-.).* **274**, 103–106 (1996).

141. Cheng, X. *et al.* PDIP38 associates with proteins constituting the mitochondrial DNA nucleoid. *J. Biochem.* **138**, 673–678 (2005).

142. Kao, T. Y. *et al.* Mitochondrial Lon regulates apoptosis through the association with Hsp60-mtHsp70 complex. *Cell Death Dis.* **6**, (2015).

143. Watabe, S. *et al.* Purification and characterization of a substrate protein for mitochondrial ATP-dependent protease in bovine adrenal cortex. *J. Biochem.* **115**, 648–654 (1994).

144. Bota, D. A. & Davies, K. J. A. Lon protease preferentially degrades oxidized mitochondrial

aconitase by an ATP-stimulated mechanism. *Nat. Cell Biol.* **4**, 674–680 (2002).

145. Kita, K., Suzuki, T. & Ochi, T. Diphenylarsinic acid promotes degradation of glutaminase C by mitochondrial lon protease. *J. Biol. Chem.* **287**, 18163–18172 (2012).

146. Lu, B. *et al.* Phosphorylation of Human TFAM in Mitochondria Impairs DNA Binding and Promotes Degradation by the AAA+ Lon Protease. *Mol. Cell* **49**, 121–132 (2013).

147. Fukuda, R. *et al.* HIF-1 Regulates Cytochrome Oxidase Subunits to Optimize Efficiency of Respiration in Hypoxic Cells. *Cell* **129**, 111–122 (2007).

148. Sharma, R. P. & Edwards, I. R. cis-platinum: Subcellular distribution and binding to cytosolic ligands. *Biochem. Pharmacol.* **32**, 2665–2669 (1983).

149. Zhang, Y. & Maurizi, M. R. Mitochondrial ClpP activity is required for cisplatin resistance in human cells. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1862**, 252–264 (2016).

150. Van Meerloo, J., Kaspers, G. J. L. & Cloos, J. Cell sensitivity assays: The MTT assay. *Methods Mol. Biol.* **731**, 237–245 (2011).

151. Uhlen, M. *et al.* Towards a knowledge-based Human Protein Atlas. *Nature Biotechnology* **28**, 1248–1250 (2010).

152. Uhlen, M. *et al.* Tissue-based map of the human proteome. *Science (80-.).* **347**, 1260419–1260419 (2015).

153. Paredes, F. *et al.* Poldip2 is an oxygen-sensitive protein that controls PDH and αKGDH lipoylation and activation to support metabolic adaptation in hypoxia and cancer. *Proc. Natl. Acad. Sci.* **115**, 1789–1794 (2018).

154. Wang, Z. *et al.* Cyclin B1/Cdk1 coordinates mitochondrial respiration for Cell-Cycle G2/M progression. *Dev. Cell* **29**, 217–232 (2014).

155. Goldberg, A. L. The mechanism and functions of ATP- dependent proteases in bacterial and animal cells. *Eur. J. Biochem.* **203**, 9–23 (1992).

156. Maurizi, M. R. *et al.* Sequence and structure of Clp P, the proteolytic component of the ATP-dependent Clp protease of Escherichia coli. *J. Biol. Chem.* **265**, 12536–12545 (1990).

157. Aizenman, E., Engelberg-Kulka, H. & Glaser, G. An Escherichia coli chromosomal 'addiction module' regulated by guanosine [corrected] 3',5'-bispyrophosphate: a model for programmed bacterial cell death. *Proc. Natl. Acad. Sci.* **93**, 6059–6063 (1996).

158. Wojtkowiak, D., Georgopoulos, C. & Zylicz, M. Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of Escherichia coli. *J. Biol. Chem.* **268**, 22609–22617 (1993).

159. Shapiro, J. A. A role for the Clp protease in activating Mu-mediated DNA rearrangements. *J. Bacteriol.* **175**, 2625–2631 (1993).

160. Schweder, T., Lee, K. H. O., Lomovskaya, O. & Matin, A. Regulation of Escherichia coli starvation sigma factor (σs) by ClpXP protease. *J. Bacteriol.* **178**, 470–476 (1996).

161. Iyoda, S. & Watanabe, H. ClpXP protease controls expression of the type III protein secretion system through regulation of RpoS and GrlR levels in enterohemorrhagic

Escherichia coli. *J. Bacteriol.* **187**, 4086–4094 (2005).

162. Damerau, K. & St. John, A. C. Role of Clp protease subunits in degradation of carbon starvation proteins in Escherichia coli. *J. Bacteriol.* **175**, 53–63 (1993).

163. Zhou, Y. & Gottesman, S. Regulation of proteolysis of the stationary-phase sigma factor RpoS. *J. Bacteriol.* **180**, 1154–1158 (1998).

164. Hengge-Aronis, R. Survival of hunger and stress: The role of rpoS in early stationary phase gene regulation in E. coli. *Cell* **72**, 165–168 (1993).

165. Kandror, O., Busconi, L., Sherman, M. & Goldberg, A. L. Rapid degradation of an abnormal protein in Escherichia coli involves the chaperones GroEL and GroES. *J. Biol. Chem.* **269**, 23575–23582 (1994).

166. Bressanin, D. *et al.* Proteolysis of the proofreading subunit controls the assembly of Escherichia coli DNA polymerase III catalytic core. *Biochim. Biophys. Acta - Proteins Proteomics* **1794**, 1606–1615 (2009).

167. Camberg, J. L., Hoskins, J. R. & Wickner, S. The Interplay of ClpXP with the cell division machinery in Escherichia coli. *J. Bacteriol.* **193**, 1911–1918 (2011).

168. Kitagawa, R., Takaya, A. & Yamamoto, T. Dual regulatory pathways of flagellar gene expression by ClpXP protease in enterohaemorrhagic Escherichia coli. *Microbiology* **157**, 3094–3103 (2011).

169. Potocka, I., Thein, M., Østerås, M., Jenal, U. & Alley, M. R. K. Degradation of a Caulobacter soluble cytoplasmic chemoreceptor is ClpX dependent. *J. Bacteriol.* **184**, 6635–6641 (2002).

170. Tsai, J. W. & Alley, M. R. K. Proteolysis of the Caulobacter McpA chemoreceptor is cell cycle regulated by a ClpX-dependent pathway. *J. Bacteriol.* **183**, 5001–5007 (2001).

171. Jessani, N. & Cravatt, B. F. The development and application of methods for activity-based protein profiling. *Current Opinion in Chemical Biology* **8**, 54–59 (2004).

172. Böttcher, T. & Sieber, S. A. β-lactones as privileged structures for the active-site labeling of versatile bacterial enzyme classes. *Angew. Chemie - Int. Ed.* **47**, 4600–4603 (2008).

173. Gersch, M. *et al.* The mechanism of caseinolytic protease (ClpP) inhibition. *Angew. Chemie - Int. Ed.* **52**, 3009–3014 (2013).

174. Gersch, M., Kolb, R., Alte, F., Groll, M. & Sieber, S. A. Disruption of oligomerization and dehydroalanine formation as mechanisms for ClpP protease inhibition. *J. Am. Chem. Soc.* **136**, 1360–1366 (2014).

175. Hsu, J. L., Huang, S. Y., Chow, N. H. & Chen, S. H. Stable-Isotope Dimethyl Labeling for Quantitative Proteomics. *Anal. Chem.* **75**, 6843–6852 (2003).

176. Bruce, J. E. In vivo protein complex topologies: Sights through a cross-linking lens. *Proteomics* **12**, 1565–1575 (2012).

177. Weisbrod, C. R. *et al.* In vivo protein interaction network identified with a novel real-time cross-linked peptide identification strategy. *J. Proteome Res.* **12**, 1569–1579 (2013).

178. Balogh, D. *et al.* Insights into ClpXP proteolysis: heterooligomerization and partial

deactivation enhance chaperone affinity and substrate turnover in Listeria monocytogenes. *Chem. Sci.* **8**, 1592–1600 (2017).

179. Gersch, M. *et al.* AAA+ chaperones and acyldepsipeptides activate the ClpP protease via conformational control. *Nat. Commun.* **6**, 6320 (2015).

180. Gersch, M. *et al.* Barrel-shaped ClpP Proteases Display Attenuated Cleavage Specificities. *ACS Chem. Biol.* **11**, 389–399 (2016).

181. Amor, A. J., Schmitz, K. R., Sello, J. K., Baker, T. A. & Sauer, R. T. Highly Dynamic Interactions Maintain Kinetic Stability of the ClpXP Protease during the ATP-Fueled Mechanical Cycle. *ACS Chem. Biol.* **11**, 1552–1560 (2016).

182. DeLano, W. L. The PyMOL Molecular Graphics System, Version 0.99 Schrödinger, LLC. *Schrödinger LLC* **Version 1.**, http://www.pymol.org (2002).

183. Yan, Y., Zhang, D., Zhou, P., Li, B. & Huang, S. Y. HDOCK: A web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* **45**, W365–W373 (2017).

184. D.A. Case and R.M. Betz and W. Botello-Smith and D.S. Cerutti and T.E. Cheatham and III and T.A. Darden and R.E. Duke and T.J. Giese and H. Gohlke and A.W. Goetz and N. Homeyer and S. Izadi and P. Janowski and J. Kaus and A. Kovalenko and T.S. Lee and S. AMBER 2016. *University of California, San Francisco* (2016).

185. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

186. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

187. Meagher, K. L., Redman, L. T. & Carlson, H. A. Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.* **24**, 1016–1025 (2003).

188. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).

189. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).

190. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L. & Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **37**, (2009).

191. Biasini, M. *et al.* SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, (2014).

192. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201 (2006).

193. Guex, N., Peitsch, M. C. & Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective.

Electrophoresis **30**, (2009).

194. Pettersen, E. F. *et al.* UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem* **25**, 1605–1612 (2004).

195. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* **45**, D331–D338 (2017).

196. Percudani, R. & Peracchi, A. A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Reports* (2003). doi:10.1038/sj.embor.embor914

197. Di Salvo, M. L., Contestabile, R. & Safo, M. K. Vitamin B 6 salvage enzymes: Mechanism, structure and regulation. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2011). doi:10.1016/j.bbapap.2010.12.006

198. Hoegl, A. *et al.* Mining the cellular inventory of pyridoxal phosphate-dependent enzymes with functionalized cofactor mimics. *Nature Chemistry* (2018). doi:10.1038/s41557-018-0144-2

199. Jang, Y. M. *et al.* Human pyridoxal phosphatase: Molecular cloning, functional expression, and tissue distribution. *J. Biol. Chem.* (2003). doi:10.1074/jbc.M309619200

200. Merrill, A. H., Henderson, J. M., Wang, E., McDonald, B. W. & Millikan, W. J. Metabolism of vitamin B-6 by human liver. *J. Nutr.* (1984). doi:10.1093/jn/114.9.1664

201. Midttun, O., Hustad, S. & Ueland, P. M. Quantitative profiling of biomarkers related to B-vitamin status, tryptophan metabolism and inflammation in human plasma by liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* (2009). doi:10.1002/rcm.4013

202. Helmreich, E. J. How pyridoxal 5'-phosphate could function in glycogen phosphorylase catalysis. *Biofactors* **3**, 159–72 (1992).

203. El-Sayed, A. S. & Shindia, A. A. PLP-Dependent Enzymes: a Potent Therapeutic Approach for Cancer and Cardiovascular Diseases. in *Targets in Gene Therapy* (2012). doi:10.5772/17449

204. Alexander, F. W., Sandmeier, E., Mehta, P. K. & Christen, P. Evolutionary relationships enzymes Regio-specific a , p and y families. *Eur. J. Biochem.* **219**, 953–960 (1994).

205. Eliot, A. C. & Kirsch, J. F. Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **73**, 383–415 (2004).

206. Parra, M., Stahl, S. & Hellmann, H. Vitamin B6 and Its Role in Cell Metabolism and Physiology. *Cells* (2018). doi:10.3390/cells7070084

207. Cellini, B., Montioli, R., Oppici, E., Astegno, A. & Borri Voltattorni, C. The chaperone role of the pyridoxal 5'-phosphate and its implications for rare diseases involving B6-dependent enzymes. *Clinical Biochemistry* **47**, 158–165 (2014).

208. Amadasi, A. *et al.* Pyridoxal 5 '-phosphate enzymes as targets for therapeutic agents. *Curr. Med. Chem.* (2007).

209. Bakay, R. A. E. & Harris, A. B. Neurotransmitter, receptor and biochemical changes in monkey cortical epileptic foci. *Brain Res.* **206**, 387–404 (1981).

210. Nishino, N., Fujiwara, H., Noguchi-Kuno, S. A. & Tanaka, C. GABAA receptor but not muscarinic receptor density was decreased in the brain of patients with Parkinson's disease. *Jpn. J. Pharmacol.* **48**, 331–339 (2008).

211. Butterworth, J., Yates, C. M. & Simpson, J. Phosphate- Activated Glutaminase in Relation to Huntington's Disease and Agonal State. *J. Neurochem.* **41**, 440–447 (1983).

212. Aoyagi, T. *et al.* Increased γ-Aminobutyrate Aminotransferase Activity in Brain of Patients with Alzheimer's Disease. *Chem. Pharm. Bull.* **38**, 1748–1749 (1990).

213. Renwick, S. B., Snell, K. & Baumann, U. The crystal structure of human cytosolic serine hydroxymethyltransferase: A target for cancer chemotherapy. *Structure* **6**, 1105–1116 (1998).

214. Zigmond, E. *et al.* Suppression of Hepatocellular Carcinoma by Inhibition of Overexpressed Ornithine Aminotransferase. *ACS Med. Chem. Lett.* (2015). doi:10.1021/acsmedchemlett.5b00153

215. Ducker, G. S. *et al.* Human SHMT inhibitors reveal defective glycine import as a targetable metabolic vulnerability of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci.* 201706617 (2017). doi:10.1073/pnas.1706617114

216. Snell, K., Natsumeda, Y., Eble, J. N., Glover, J. L. & Weber, G. Enzymic imbalance in serine metabolism in human colon carcinoma and rat sarcoma. *Br. J. Cancer* **57**, 87–90 (1988).

217. Paone, A. *et al.* SHMT1 knockdown induces apoptosis in lung cancer cells by causing uracil misincorporation. *Cell Death Dis.* **5**, (2014).

218. Wang, G., Shang, L., Burgett, A. W. G., Harran, P. G. & Wang, X. Diazonamide toxins reveal an unexpected function for ornithine δ-amino transferase in mitotic cell division. *Proc. Natl. Acad. Sci.* **104**, 2068–2073 (2007).

219. Choi, S. Y., Jang, J. H. & Kim, K. R. Analysis of differentially expressed genes in human rectal carcinoma using suppression subtractive hybridization. *Clin. Exp. Med.* (2011). doi:10.1007/s10238-010-0130-5

220. Storici, P., Qiu, J., Schirmer, T. & Silverman, R. B. Mechanistic crystallography, mechanism of inactivation of γ-aminobutyric acid aminotransferase by (1R,3S,4S)-3-amino-4- fluorocyclopentane-1-carboxylic acid as elucidated by crystallography. *Biochemistry* **43**, 14057–14063 (2004).

221. Wu, F., Yu, J. & Gehring, H. Inhibitory and structural studies of novel coenzyme-substrate analogs of human histidine decarboxylase. *FASEB J.* **22**, 890–897 (2007).

222. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).

223. Speers, A. E. & Cravatt, B. F. Profiling enzyme activities in vivo using click chemistry methods. *Chem. Biol.* **11**, 535–546 (2004).

224. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci.* **96**, 14694–14699 (2002).

225. Greenbaum, D. *et al.* Chemical Approaches for Functionally Probing the Proteome. *Mol.*

Cell. Proteomics **1**, 60–68 (2003).

226. Ong, S.-E. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Mol. Cell. Proteomics* **1**, 376–386 (2002).

227. Anderson, L. N. *et al.* Live Cell Discovery of Microbial Vitamin Transport and Enzyme-Cofactor Interactions. *ACS Chem. Biol.* **11**, 345–354 (2016).

228. Lowther, J. *et al.* Inhibition of the PLP-dependent enzyme serine palmitoyltransferase by cycloserine: Evidence for a novel decarboxylative mechanism of inactivation. *Mol. Biosyst.* **6**, 1682–1693 (2010).

229. Lee, H. S., Moon, B. J., Choi, S. Y. & Kwon, O. S. Human pyridoxal kinase: overexpression and properties of the recombinant enzyme. *Mol. Cells* **10**, 452–459 (2000).

230. Kimura, T. *et al.* The antimalarial drugs chloroquine and primaquine inhibit pyridoxal kinase, an essential enzyme for vitamin B6 production. *FEBS Lett.* **588**, 3673–3676 (2014).

231. Nodwell, M. B., Koch, M. F., Alte, F., Schneider, S. & Sieber, S. A. A subfamily of bacterial ribokinases utilizes a hemithioacetal for pyridoxal phosphate salvage. *J. Am. Chem. Soc.* **136**, 4992–4999 (2014).

232. Cao, P., Gong, Y., Tang, L., Leung, Y. C. & Jiang, T. Crystal structure of human pyridoxal kinase. *J. Struct. Biol.* **154**, 327–332 (2006).

233. Gandhi, A. K. *et al.* Kinetic and structural studies of the role of the active site residue Asp235 of human pyridoxal kinase. *Biochem. Biophys. Res. Commun.* **381**, 12–15 (2009).

234. Islam, K. The Bump-and-Hole Tactic: Expanding the Scope of Chemical Genetics. *Cell Chemical Biology* **25**, 1171–1184 (2018).

235. Musayev, F. N. *et al.* Crystal Structure of human pyridoxal kinase : Structural basis of M + and M 2 + activation. *Protein Sci.* **1**, 2184–2194 (2007).

236. Basu, S. S., Mesaros, C., Gelhaus, S. L. & Blair, I. A. Stable isotope labeling by essential nutrients in cell culture for preparation of labeled coenzyme a and its thioesters. *Anal. Chem.* (2011). doi:10.1021/ac1027353

237. Lehmann, J., Wright, M. H. & Sieber, S. A. Making a Long Journey Short: Alkyne Functionalization of Natural Product Scaffolds. *Chem. - A Eur. J.* **22**, 4666–4678 (2016).

238. Van Geel, R., Pruijn, G. J. M., Van Delft, F. L. & Boelens, W. C. Preventing thiol-yne addition improves the specificity of strain-promoted azide-alkyne cycloaddition. *Bioconjug. Chem.* **23**, 392–398 (2012).

239. Fux, A., Pfanzelt, M., Kirsch, V. C., Hoegl, A. & Sieber, S. A. Customizing Functionalized Cofactor Mimics to Study the Human Pyridoxal 5'-Phosphate-Binding Proteome. *Cell Chem. Biol.* **26**, 1–8 (2019).

240. Thornburg, J. M. *et al.* Targeting aspartate aminotransferase in breast cancer. *Breast Cancer Res.* **10**, (2008).

241. Futerman, A. H. & Hannun, Y. A. The complex life of simple sphingolipids. *EMBO Reports* **5**, 777–782 (2004).

242. Bras, J., Singleton, A., Cookson, M. R. & Hardy, J. Emerging pathways in genetic Parkinson's disease: Potential role of ceramide metabolism in Lewy body disease. *FEBS Journal* **275**, 5767–5773 (2008).

243. Park, T. S., Rosebury, W., Kindt, E. K., Kowala, M. C. & Panek, R. L. Serine palmitoyltransferase inhibitor myriocin induces the regression of atherosclerotic plaques in hyperlipidemic ApoE-deficient mice. *Pharmacol. Res.* **58**, 45–51 (2008).

244. Brancaleone, V. *et al.* D -Penicillamine modulates hydrogen sulfide (H2S) pathway through selective inhibition of cystathionine-γ-lyase. *Br. J. Pharmacol.* **173**, 1556–1565 (2016).

245. Chao, C. & Zatarain, J. R. Cystathionine-β-Synthase Inhibition for Colon Cancer: Enhancement of the Efficacy of Aminooxyacetic Acid via the Prodrug Approach. *Mol. Med.* **22**, 1 (2016).

246. Phillips, R. S. Chemistry and diversity of pyridoxal-5′-phosphate dependent enzymes. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1854**, 1167–1174 (2015).

247. Linster, C. L., Van Schaftingen, E. & Hanson, A. D. Metabolite damage and its repair or pre-emption. *Nature Chemical Biology* (2013). doi:10.1038/nchembio.1141

248. Windebank, A. J. Neurotoxicity of pyridoxine analogs is related to coenzyme structure. *Neurochem. Pathol.* (1985). doi:10.1007/BF02834268

249. Nilsson, B. L., Kiessling, L. L. & Raines, R. T. Staudinger Ligation: A Peptide from a Thioester and Azide. *Org. Lett.* **2**, 1939–1941 (2002).

250. Percudani, R. & Peracchi, A. The B6 database: A tool for the description and classification of vitamin B6-dependent enzymatic activities and of the corresponding protein families. *BMC Bioinformatics* (2009). doi:10.1186/1471-2105-10-273

251. Demoss, J. A. Studies on the mechanism of the tryptophan synthetase reaction. *BBA - Biochim. Biophys. Acta* **62**, 279–293 (1962).

252. Lambert, M. P. & Neuhaus, F. C. Mechanism of D-cycloserine action: alanine racemase from Escherichia coli W. *J. Bacteriol.* **110**, 978–987 (1972).

253. Wang, R. Physiological implications of hydrogen sulfide: A whiff exploration that blossomed. *Physiol. Rev.* **92**, 791–896 (2012).

254. Whiteman, M. & Winyard, P. G. Hydrogen sulfide and inflammation: The good, the bad, the ugly and the promising. *Expert Review of Clinical Pharmacology* **4**, 13–32 (2011).

255. Szabo, C. *et al.* Tumor-derived hydrogen sulfide, produced by cystathionine- -synthase, stimulates bioenergetics, cell proliferation, and angiogenesis in colon cancer. *Proc. Natl. Acad. Sci.* **110**, 12474–12479 (2013).

256. Liu, W. *et al.* Crystal structures of unbound and aminooxyacetate-bound Escherichia coli γ-aminobutyrate aminotransferase. *Biochemistry* **43**, 10896–10905 (2004).

257. Nadvi, N. A. *et al.* High resolution crystal structures of human kynurenine

aminotransferase-I bound to PLP cofactor, and in complex with aminooxyacetate. *Protein Sci.* **26**, 727–736 (2017).

258. Asimakopoulou, A. *et al.* Selectivity of commonly used pharmacological inhibitors for cystathionine β synthase (CBS) and cystathionine γ lyase (CSE). *Br. J. Pharmacol.* **169**, 922–932 (2013).

259. Matthews, M. L. *et al.* Chemoproteomic profiling and discovery of protein electrophiles in human cells. *Nat. Chem.* **9**, 234–243 (2017).

260. Weigert, W. M., Offermanns, H. & Degussa, P. S. D- Penicillamine—Production and Properties. *Angew. Chemie Int. Ed. English* **14**, 330–336 (1975).

261. Jaffe, I. A. Rheumatoid Arthritis with Arteritis. *Ann. Intern. Med.* **61**, 556 (1964).

262. Walshe, J. M. Penicillamine, a new oral therapy for Wilson's disease. *Am. J. Med.* **21**, 487–495 (1956).

263. Lyle, W. H. Penicillamine in metal poisoning. *J. Rheumatol.* **8**, 96–99 (1981).

264. Kuchinskas, E. J., Horvath, A. & du Vigneaud, V. An anti-vitamin B6 action of l-penicillamine. *Arch. Biochem. Biophys.* **68**, 69–75 (1957).

265. Rumsby, P. C. & Shepherd, D. M. The effect of penicillamine on vitamin B6 function in man. *Biochem. Pharmacol.* **30**, 3051–3053 (1981).

266. Takahashi, Y. & Matsuda, M. Effects of penicillamine on the contents of B6 vitamers of the mouse brain. *J. Nutr. Sci. Vitaminol. (Tokyo).* **22**, 375–379 (1976).

267. Sukanya, N., Vijaya, M., Radhakrishnan, A. N., Rao, N. A. & Savithri, H. S. Serine Hydroxymethyltransferase from Mung Bean (Vigna radiata) Is Not a Pyridoxal-5'-Phosphate-Dependent Enzyme. *PLANT Physiol.* **95**, 351–357 (2008).

268. du Vigneaud, V., Kuchinskas, E. J. & Horvath, A. l-Penicillamine and rat liver transaminase activity. *Arch. Biochem. Biophys.* **69**, 130–137 (1957).

269. Lowther, J., Beattie, A. E., Langridge-Smith, P. R. R., Clarke, D. J. & Campopiano, D. J. L-Penicillamine is a mechanism-based inhibitor of serine palmitoyltransferase by forming a pyridoxal-5′-phosphate-thiazolidine adduct. *Medchemcomm* **3**, 1003–1008 (2012).

270. Heyl, D., Harris, S. A. & Folkers, K. The Chemistry of Vitamin B6. VI. Pyridoxylamino Acids. *J. Am. Chem. Soc.* **70**, 3429–3431 (1948).

271. Hacker, S. M. *et al.* Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).

272. Schmidt, T. *et al.* ProteomicsDB. *Nucleic Acids Res.* **46**, D1271–D1281 (2018).

273. Whittaker, M. M., Penmatsa, A. & Whittaker, J. W. The Mtm1p carrier and pyridoxal 5′-phosphate cofactor trafficking in yeast mitochondria. *Arch. Biochem. Biophys.* **568**, 64–70 (2015).

274. Li, T. K., Lumeng, L. & Veitch, R. L. Regulation of pyridoxal 5′-phosphate metabolism in liver. *Biochem. Biophys. Res. Commun.* **61**, 677–684 (1974).

275. Kim, Y. T., Kwok, F. & Churchich, J. E. Interactions of pyridoxal kinase and aspartate

aminotransferase emission anisotropy and compartmentation studies. *J. Biol. Chem.* **263**, 13712–13717 (1988).

276. Cheung, P. Y. *et al.* Interaction between Pyridoxal Kinase and Pyridoxal-5-phosphate-Dependent Enzymes. *J. Biochem.* **134**, 731–738 (2003).

277. Yang, E. S. & Schirch, V. Tight binding of pyridoxal 5'-phosphate to recombinant Escherichia coli pyridoxine 5'-phosphate oxidase. *Arch. Biochem. Biophys.* **377**, 109–114 (2000).

278. Maere, S., Heymans, K. & Kuiper, M. BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* **21**, 3448–3449 (2005).

279. Ashburner, M. *et al.* The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2011).

280. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).

281. Ikegawa, S., Isomura, M., Koshizuka, Y. & Nakamura, Y. Cloning and characterization of human and mouse PROSC (proline synthetase co-transcribed) genes. *J. Hum. Genet.* **44**, 337–342 (1999).

282. Ito, T. *et al.* Conserved pyridoxal protein that regulates ile and val metabolism. *J. Bacteriol.* **195**, 5439–5449 (2013).

283. Prunetti, L. *et al.* Evidence that COG0325 proteins are involved in PLP homeostasis. *Microbiol. (United Kingdom)* **162**, 694–706 (2016).

284. Labella, J. I. *et al.* PipY, a member of the conserved COG0325 family of PLP-binding proteins, expands the cyanobacterial nitrogen regulatory network. *Front. Microbiol.* **8**, (2017).

285. Darin, N. *et al.* Mutations in PROSC Disrupt Cellular Pyridoxal Phosphate Homeostasis and Cause Vitamin-B6-Dependent Epilepsy. *J. Inherit. Metab. Dis.* **33**, 1–10 (2016).

286. Jaume, C. & Plecko, B. Treatable newborn and infant seizures due to inborn errors of metabolism. *Epileptic Disord.* **17**, 229–242 (2015).

287. Roessler, S. *et al.* Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. *Gastroenterology* **142**, (2012).

288. Tremiño, L., Forcada-Nadal, A. & Rubio, V. Insight into vitamin B6-dependent epilepsy due to PLPBP (previously PROSC) missense mutations. *Hum. Mutat.* (2018). doi:10.1002/humu.23540

289. Eswaramoorthy, S. *et al.* Structure of a yeast hypothetical protein selected by a structural genomics approach. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **59**, 127–135 (2003).

290. Johnstone, D. L. *et al.* PLPHP deficiency: Clinical, genetic, biochemical, and mechanistic insights. *Brain* **142**, 542–559 (2019).

291. Tremiño, L., Forcada-Nadal, A., Contreras, A. & Rubio, V. Studies on cyanobacterial

protein PipY shed light on structure, potential functions, and vitamin B6-dependent epilepsy. *FEBS Letters* **591**, 3431–3442 (2017).

292. Amorim Franco, T. M., Favrot, L., Vergnolle, O. & Blanchard, J. S. Mechanism-Based Inhibition of the Mycobacterium tuberculosis Branched-Chain Aminotransferase by d - And l -Cycloserine. *ACS Chem. Biol.* **12**, 1235–1244 (2017).

293. Ahmed, S. A., McPhie, P. & Miles, E. W. A thermally induced reversible conformational transition of the tryptophan synthase β2subunit probed by the spectroscopic properties of pyridoxal phosphate and by enzymatic activity. *J. Biol. Chem.* **271**, 8612–8617 (1996).

294. Hunt, A. D., Stokes, J., McCrory, W. W. & Stroud, H. H. Pyridoxine dependency: report of a case of intractable convulsions in an infant controlled by pyridoxine. *Pediatrics* **13**, 140–5 (1954).

295. Mills, P. B. *et al.* Genotypic and phenotypic spectrum of pyridoxine-dependent epilepsy (ALDH7A1 deficiency). *Brain* **133**, 2148–2159 (2010).

296. Plecko, B. Pyridoxine and pyridoxalphosphate-dependent epilepsies. in *Handbook of Clinical Neurology* **113**, 1811–1817 (2013).

297. Mills, P. B. *et al.* Mutations in antiquitin in individuals with pyridoxine-dependent seizures. *Nat. Med.* **12**, 307–309 (2006).

298. Farrant, R. D., Walker, V., Mills, G. A., Mellor, J. M. & Langley, G. J. Pyridoxal Phosphate De-activation by Pyrroline-5-carboxylic Acid. *J. Biol. Chem.* **276**, 15107–15116 (2001).

299. Shiraku, H. *et al.* PLPBP mutations cause variable phenotypes of developmental and epileptic encephalopathy. *Epilepsia Open* **3**, 495–502 (2018).

300. Plecko, B. *et al.* Confirmation of mutations in PROSC as a novel cause of Vitamin B 6-dependent epilepsy. *J. Med. Genet.* **54**, 809–814 (2017).

301. Johansson, S., Lindstedt, S. & Tiselius, H. G. Metabolism of [3H8]Pyridoxine in Mice. *Biochemistry* (1968). doi:10.1021/bi00846a039

302. Johansson, S., Lindstedt, S. & Tiselius, H. G. Metabolic interconversions of different forms of vitamin B6. *J. Biol. Chem.* (1974).

303. Contractor, S. F. & Shane, B. Metabolism of [14C]pyridoxol in the pregnant rat. *BBA - Gen. Subj.* (1971). doi:10.1016/0304-4165(71)90060-2

304. Colombini, C. E. & McCoy, E. E. Vitamin B6 Metabolism. the Utilization of [14C]Pyridoxine by the Normal Mouse. *Biochemistry* (1970). doi:10.1021/bi00805a012

305. Kanehisa, M. *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* (2008). doi:10.1093/nar/gkm882

306. Meng, Z. *et al.* Purification, characterization, and crystallization of human pyrroline-5-carboxylate reductase. *Protein Expr. Purif.* **49**, 83–87 (2006).

307. Ballouz, S., Francis, A. R., Lan, R. & Tanaka, M. M. Conditions for the evolution of gene clusters in bacterial genomes. *PLoS Comput. Biol.* **6**, (2010).

308. Tarailo-Graovac, M. & Chen, N. Gene Clustering in Eukaryotes. in *eLS* (2013). doi:10.1002/9780470015902.a0006117.pub3

309. Griswold, W. R. & Toney, M. D. Role of the pyridine nitrogen in pyridoxal 5'-phosphate catalysis: Activity of three classes of PLP enzymes reconstituted with deazapyridoxal 5'-phosphate. *J. Am. Chem. Soc.* **133**, 14823–14830 (2011).

310. Akram, M. Citric Acid Cycle and Role of its Intermediates in Metabolism. *Cell Biochemistry and Biophysics* **68**, 475–478 (2014).

311. Xu, X. *et al.* Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J. Biol. Chem.* (2004). doi:10.1074/jbc.M404298200

312. Olano, J., Soler, J., Busto, F. & De Arriaga, D. Chemical modification of NADP-isocitrate dehydrogenase from Cephalosporium acremonium. Evidence of essential histidine and lysine groups at the active site. *Eur. J. Biochem.* (1999). doi:10.1046/j.1432-1327.1999.00297.x

313. Fan, C. C. & Plaut, G. W. E. Functional Groups of Diphosphopyridine Nucleotide Linked Isocitrate Dehydrogenase from Bovine Heart. II. Studies of an Active Amino Group by Reaction with Aldehydes. *Biochemistry* (1974). doi:10.1021/bi00698a009

314. Chen, S. S. & Engel, P. C. Reversible modification of pig heart mitochondrial malate dehydrogenase by pyridoxal 5'-phosphate. *Biochem. J.* (1975). doi:10.1042/bj1510297

315. Wimmer, M. J., Mo, T., Sawyers, D. L. & Harrison, J. H. Biphasic inactivation of porcine heart mitochondrial malate dehydrogenase by pyridoxal 5' phosphate. *J. Biol. Chem.* (1975).

316. Bleile, D. M., Jameson, J. L. & Harrison, J. H. Inactivation of porcine heart cytoplasmic malate dehydrogenase by pyridoxal 5' phosphate. *J. Biol. Chem.* (1976).

317. Abe, Y. *et al.* Cloning and Characterization of a p53-related Protein Kinase Expressed in Interleukin-2-activated Cytotoxic T-cells, Epithelial Tumor Cell Lines, and the Testes. *J. Biol. Chem.* **276**, 44003–44011 (2001).

318. Dumaz, N. Serine15 phosphorylation stimulates p53 transactivation but does not directly influence interaction with HDM2. *EMBO J.* **18**, 7002–7010 (1999).

319. Fricke, B., Heink, S., Steffen, J., Kloetzel, P. M. & Krüger, E. The proteasome maturation protein POMP facilitates major steps of 20S proteasome formation at the endoplasmic reticulum. *EMBO Rep.* **8**, 1170–1175 (2007).

320. Lambrecht, J. A., Flynn, J. M. & Downs, D. M. Conserved Yjgf protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. *J. Biol. Chem.* **287**, 3454–3461 (2012).

321. Ivanov, A. I. Role for Actin Filament Turnover and a Myosin II Motor in Cytoskeleton-driven Disassembly of the Epithelial Apical Junctional Complex. *Mol. Biol. Cell* **15**, 2639–2651 (2004).

322. Salaun, D., Benseddik, K., Zaoui, K., Badache, A. & Daou, P. ErbB2 receptor controls microtubule capture by recruiting ACF7 to the plasma membrane of migrating cells. *Proc. Natl. Acad. Sci.* **107**, 18517–18522 (2010).

323. Koshino, I., Mohandas, N. & Takakuwa, Y. Identification of a novel role for dematin in regulating red cell membrane function by modulating spectrin-actin interaction. *J. Biol.*

Chem. **287**, 35244–35250 (2012).

324. Montefusco, D. J., Asinas, A. E. & Weis, R. M. Liposome-Mediated Assembly of Receptor Signaling Complexes. in *Methods in Enzymology* (2007). doi:10.1016/S0076-6879(07)23012-5

325. Wuichet, K., Alexander, R. P. & Zhulin, I. B. Comparative Genomic and Protein Sequence Analyses of a Complex System Controlling Bacterial Chemotaxis. in *Methods in Enzymology* (2007). doi:10.1016/S0076-6879(06)22001-9

326. Van Der Heiden, E. *et al.* A pathway closely related to the D-tagatose pathway of gram-negative enterobacteria identified in the gram-positive bacterium Bacillus licheniformis. *Appl. Environ. Microbiol.* (2013). doi:10.1128/AEM.03918-12

327. Iuchi, S., Cole, S. T. & Lin, E. C. C. Multiple regulatory elements for the glpA operon encoding anaerobic glycerol-3-phosphate dehydrogenase and the glpD operon encoding aerobic glycerol-3-phosphate dehydrogenase in Escherichia coli: Further characterization of respiratory control. *J. Bacteriol.* (1990). doi:10.1128/jb.172.1.179-184.1990

328. Dolan, S. K. *et al.* Loving the poison: The methylcitrate cycle and bacterial pathogenesis. *Microbiology (United Kingdom)* (2018). doi:10.1099/mic.0.000604

329. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* (2014). doi:10.1038/nmeth.2834

330. Ito, T. *et al.* Conserved pyridoxal 5'-phosphate-binding protein YggS impacts amino acid metabolism through pyridoxine 5'-phosphate in Escherichia coli. *Appl. Environ. Microbiol.* (2019). doi:10.1128/AEM.00430-19

331. Simon, E. S. & Allison, J. Determination of pyridoxal-5′-phosphate (PLP)-bonding sites in proteins: A peptide mass fingerprinting approach based on diagnostic tandem mass spectral features of PLP-modified peptides. *Rapid Commun. Mass Spectrom.* **23**, 3401–3408 (2009).

332. Wu, Y., Chen, J., Liu, Z. & Wang, F. Identification of pyridoxal phosphate-modified proteins using mass spectrometry. *Rapid Commun. Mass Spectrom.* **32**, 195–200 (2018).

333. Beltran, T. *et al.* Rational design of a new series of pronucleotide. *Bioorganic Med. Chem. Lett.* **11**, 1775–1777 (2001).

334. Whittaker, J. W. Intracellular trafficking of the pyridoxal cofactor. Implications for health and metabolic disease. *Archives of Biochemistry and Biophysics* **592**, 20–26 (2016).

335. Lumeng, L., Lui, A. & Li, T. K. Plasma content of B6 vitamers and its relationship to hepatic vitamin B6 metabolism. *J. Clin. Invest.* **66**, 688–695 (1980).

336. Anderson, B. B., Fulford-Jones, C. E., Child, J. A., Beard, M. E. & Bateman, C. J. Conversion of vitamin B 6 compounds to active forms in the red blood cell. *J. Clin. Invest.* **50**, 1901–1909 (1971).

337. Stolz, J. & Vielreicher, M. Tpn1p, the plasma membrane vitamin B6 transporter of Saccharomyces cerevisiae. *J. Biol. Chem.* **278**, 18990–18996 (2003).

338. Lui, A., Lumeng, L. & Li, T. K. Metabolism of vitamin B6 in rat liver mitochondria. *J. Biol.*

Chem. **256**, 6041–6046 (1981).

339. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).

IV - ABBREVIATIONS

ΔhPLK	PLK mutants
ΔPLPBP	PLPBP mutants
°C	Degree Celsius
Å	Angström
AAA+	ATPases associated with various cellular activities
ABPP	Activity-based protein-profiling
ACN	Acetonitrile
AGC	Automatic gain control
AMC	7-Amino-4-methylcoumarine
AOAA	Aminooxyacetic acid
ATET	Anhydrotetracycline
ATP	Adenosine 5'-triphosphate
BCA	Bicinchoninic acid
BHB	Bushnell Haas Broth
BS3	Bis(sulfosuccinimidyl)suberate
BSA	Bovine serum albumin
BTTAA	2-(4-((Bis((1-(tert-butyl)-1H-1,2,3-triazol-4-yl)methyl)amino)methyl)-1H-1,2,3-triazol-1-yl)acetic acid
C18	Octadecyl carbon chain bonded silica
CDM	Chemically defined media
cf.	Conferatur, "compare"
CID	Collision-induced dissociation
Clp	Caseinolytic protease
CMV	Cytomegalovirus

co-IP	Co-immunoprecipitation
Cross-link/ABPP	*In situ* chemical cross-linking combined with ABPP
Cross-link/co-IP	*In situ* chemical cross-linking with combined co-IP
CuAAC	Copper-catalyzed azide-alkyne 1,3-dipolar cycloaddition
Da	Dalton
DAPI	4′,6-Diamidin-2-phenylindol
DCC	*N,N′*-dicyclohexylcarbodiimide
DCM	Dichloromethane
DCS	ᴅ-Cycloserine
DHSO	Dihydrazide sulfoxide
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DMTMM	4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride
DNA	Deoxyribonucleic acid
DSBSO	Disuccinimidyl bissulfoxide
DSBU	Disuccinimidyl dibutyric urea
DSSO	Disuccinimidyl sulfoxide
DTT	Dithiothreitol
e.g.	Exempli gratia = for example
EcClpP	*E. coli* ClpP
EDTA	Ethylenediaminetetraacetic acid
EHEC	Enterohemorrhagic *E. coli*
EM	Electron microscopy
eq.	Equivalents

ETD	Electron-transfer-dissociation
EThcD	Electron-transfer/higher-energy collision dissociation
EtOAc	Ethyl acetate
FA	Formic acid
FCP	(1*R*,3*S*,4*S*)-3-amino-4-fluorocyclopentane-1-carboxylic acid
FCS	Fetal calf serum
FDR	False discovery rate
GABA	*γ*-Aminobutyric acid
GABA-AT	GABA-aminotransferase
GO	Gene ontology
h	Hour
HCC	Hepatocellular carcinoma
HCD	Higher-energy collision induced dissociation
hClpP	human ClpP
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High-performance liquid chromatography
IAA	Iodoacetamide
IC50 value	Half-maximal inhibitory concentration
IDH	Isocitrate dehydrogenase
IMAC	Immobilized metal affinity chromatography
k.o.	knockout
KG	α-Ketoglutaric acid
L-MDH	L-Malate dehydrogenase
LB	Lysogeny broth

Part IV-Abbreviations

LC-MS	Liquid chromatography combined with mass spectrometry
LC-MS/MS	Liquid chromatography combined with tandem mass spectrometry
LFQ	Label-free quantification
LmClpP1/2	*L. monocytogenes* ClpP1/2
LTQ	Linear trap quadrupole
m/z	Mass to charge ratio
m-AAA	Matrix-oriented AAA
m-CPA	*Meta*-chloro perbenzoic acid
MCP	Methyl-accepting chemotaxis protein
MD	Molecular dynamics
MeOH	Methanol
min	Minute
MOPS	3-(*N*-morpholino)propanesulfonic acid
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS^n	Multistage MS/MS
MTT	3-[4,5-dimethylthiazol-2-yl]- 2,5-diphenyl tetrazolium bromide
MW	Molecular weight
N/A	Not applicable
$NaBH_4$	Sodium borohydride
NHS	*N*-hydroxysuccinimide
NMR	Nuclaer magnetic resonance
NP-40	Nonidet P-40

OA	Oxaloacetic acid
OD	Optical density
PBS	Phosphate buffered saline
PBS-T	PBS supplemented with 0.5% Tween-20
PCR	Polymerase chain reaction
PDB	Protein data bank
Pen	Penicillamine
PEP	Posterior error probability
PFA	Paraformaldehyde
pH	Potentia hydrogenii
PIR	Protein Interaction Reporter
PL	Pyridoxal
PLK	Pyridoxal kinase
hPLK	Human PLK
PLP	Pyridoxal 5'-phosphate
PLP-DEs	**PLP**-dependent enzymes
PLP-ome	**PLP**-binding proteome
PM	Pyridoxamine
PMP	Pyridoxamine phosphate
PN	Pyridoxine
PNP	Pyridoxine phosphate
PNPO	Pyridoxine/pyridoxamine 5'-phosphate oxidase
PPI	Protein-protein interactions

ppm	Parts per million
QC	Quality control
QXL-MS	Quantitative XL-MS
r.t.	Room temperature
ROS	Reactive oxygene species
rpm	Rounds per minute
RPMI	Roswell Park Memorial Institute
s	Second
SaClpP	*S. aureus* ClpP
SCX	Strong cation exchange
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SEM	Standard error of the mean
SILAC	Stable isotope labeling by amino acids in cell culture
SIM	Selected ion monitoring
SuFEx	Sulfur-fluoride exchange
TBS	*Tert*-butyldimethylsilyl
TBTA	Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine
TCEP	Tris(2-carboxyethyl)phosphine
TEAB	Triethylammonium bicarbonate
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
TFAA	Trifluoroacetic anhydride

THF   Tetrahydrofolate

THPTA   Tris((1-hydroxy-propyl-1H-1,2,3-triazol-4-yl)methyl)amine

Tris   Tris(hydroxymethyl)aminomethane

UPR^{mt} Mitochondrial unfolded protein response

UV/Vis   Ultraviolet/Visible region

WT   Wild-type

VitB6-EP   Vitamin B6-dependent epilepsy

XL-MS   Chemical cross-linking in combination with mass spectrometry

V - APPENDIX
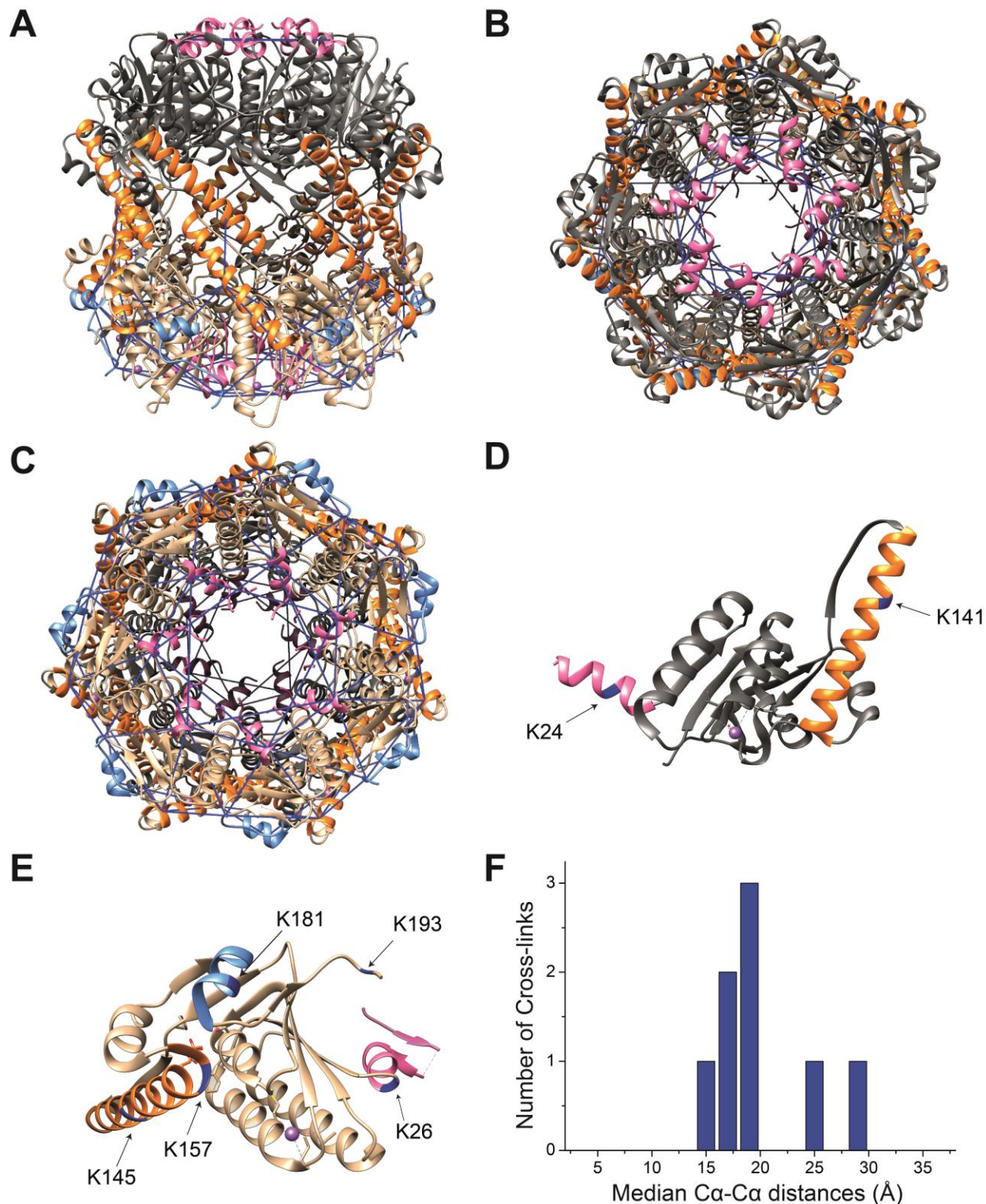
Figure S1 Cross-linking of LmClpP1/2 with the MS-cleavable linker DSSO. Identified cross-links (FDR<0.01) of the heterocomplex were mapped onto the crystal structure of LmClpP1/2 (PDB: 4RYF) in the side-view (**A**), the top-view of LmClpP1 (**B**), and the top-view of LmClpP2 (**C**). (**D**) Residues involved in cross-linking contacts are highlighted in blue in one LmClpP1 monomer. (**E**) Residues involved in cross-

linking contacts are highlighted in blue in one LmClpP2 monomer. ClpP1 is colored dark grey, ClpP2 is colored beige, E-helices are colored orange, G-helices blue, and N-terminal residues pink. (**F**) The histogram of median Cα-Cα distances of LmClpP1/2 reveals that all cross-links are within the maximum distance span of DSSO (35 Å).[110]
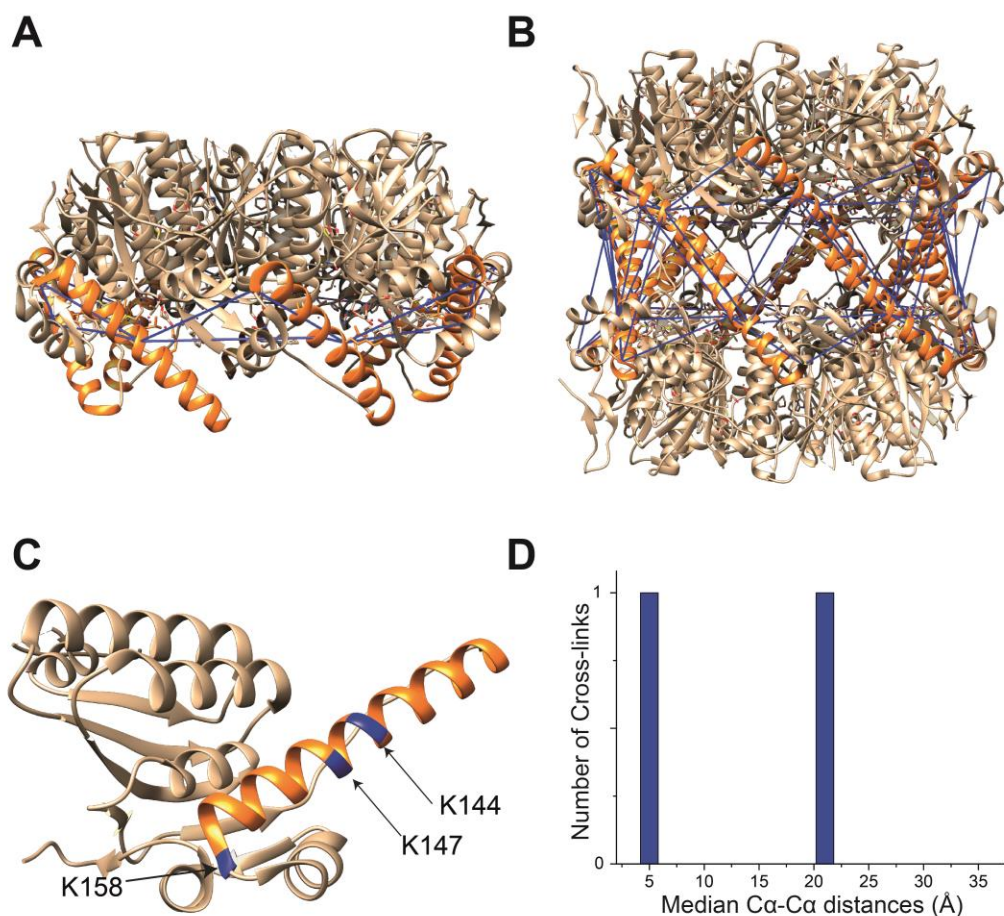
Figure S2 Cross-linking hClpP with the MS-cleavable linker DSSO. (**A**) Identified cross-links (FDR<0.01) were mapped onto the crystal structure of hClpP (PDB: 1TG6) which resembles a heptameric state. (**B**) Cross-links were further mapped on a tetradecameric model structure (**C**) Residues involved in cross-linking contacts are highlighted in blue in one hClpP monomer. E-helices are colored orange. (**D**) The histogram of median Cα-Cα distances of heptameric hClpP reveals that all cross-links are within the maximum distance span of DSSO (35 Å).[110]
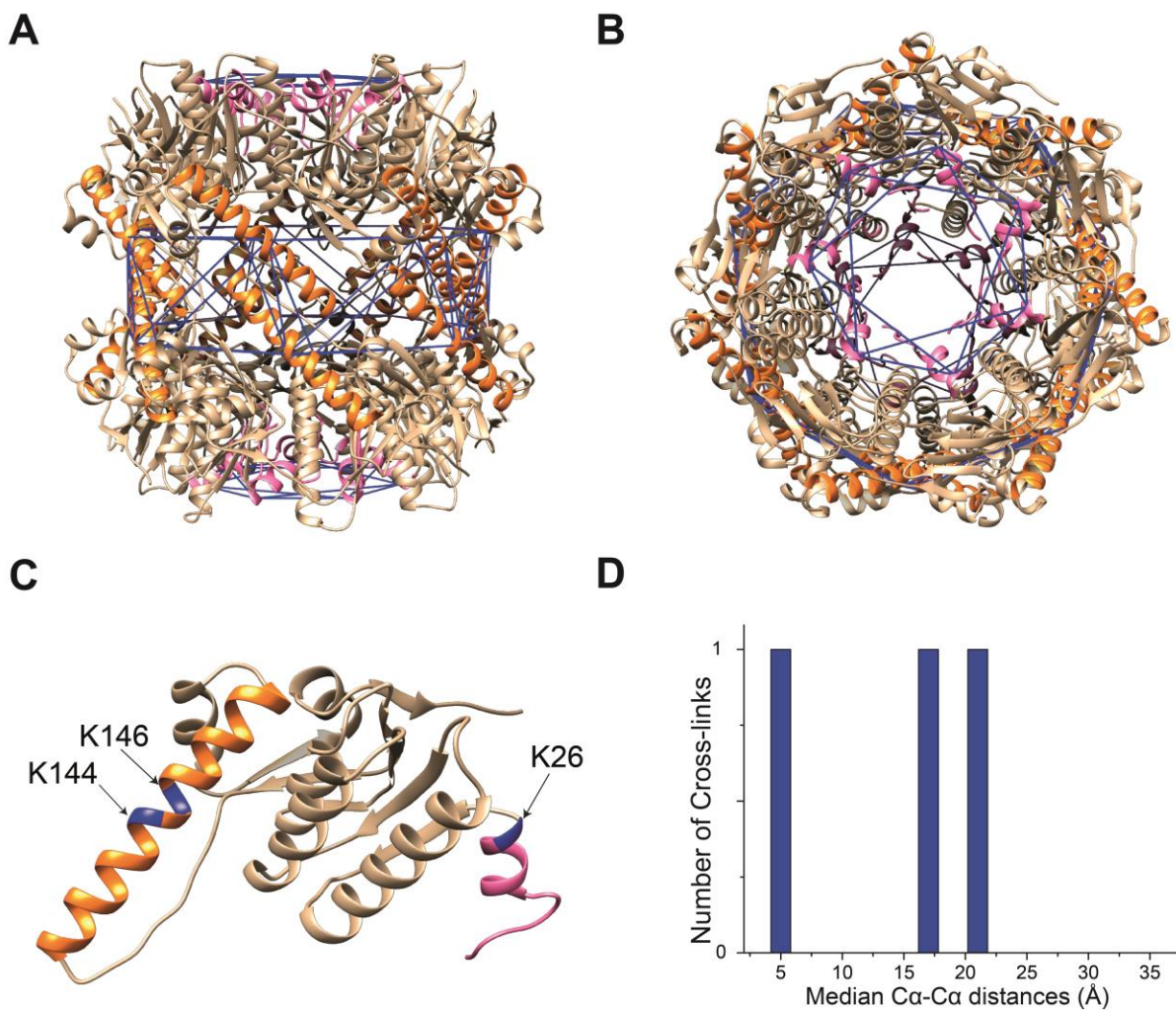
Figure S3 Cross-linking of EcClpP with the BS3 cross-linker. Identified cross-links were mapped onto the crystal structure of EcClpP (PDB: 1YG6) in the side-view (**A**) and the top-view (**B**). (**C**) Residues involved in cross-linking contacts are highlighted in blue in one EcClpP monomer. E-helices are colored orange and N-terminal residues pink. (**D**) The histogram of median Cα-Cα distances of EcClpP reveals that all cross-links are within the maximum distance span of BS3 (36 Å).[110]
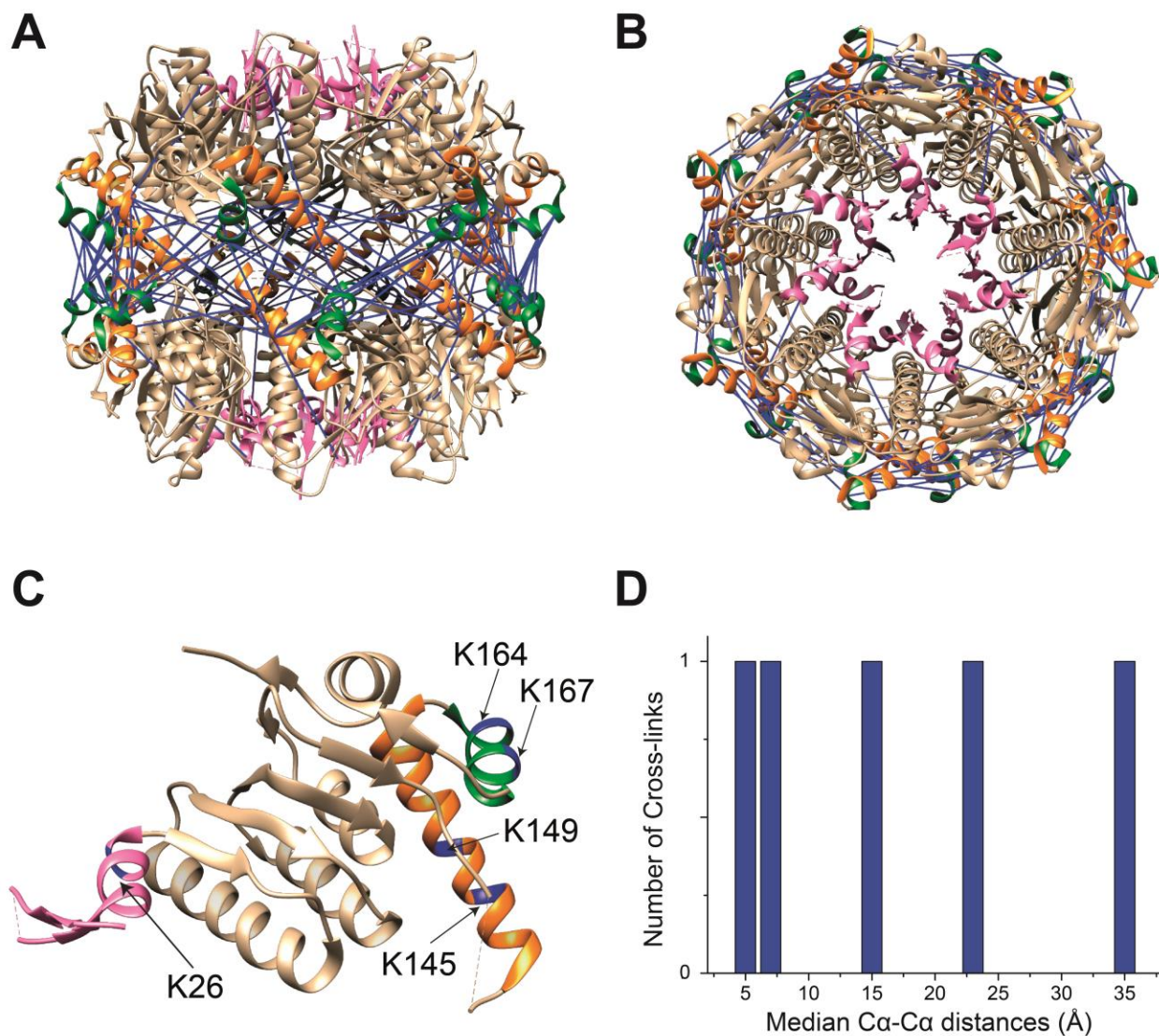
Figure S4 Cross-linking of SaClpP with the MS-cleavable linker DSSO. Identified cross-links (FDR<0.01) of the protease were mapped onto the crystal structure of SaClpP in an inactive, compressed conformation (PDB: 4EMM) in the side-view (**A**) and the top-view (**B**). (**C**) Residues involved in cross-linking contacts are highlighted in blue in one SaClpP monomer. E-helices are colored orange, F-helices green, and N-terminal residues pink. (**D**) The histogram of median Cα-Cα distances of SaClpP reveals that all cross-links are within the maximum distance span of DSSO (35 Å).[110]
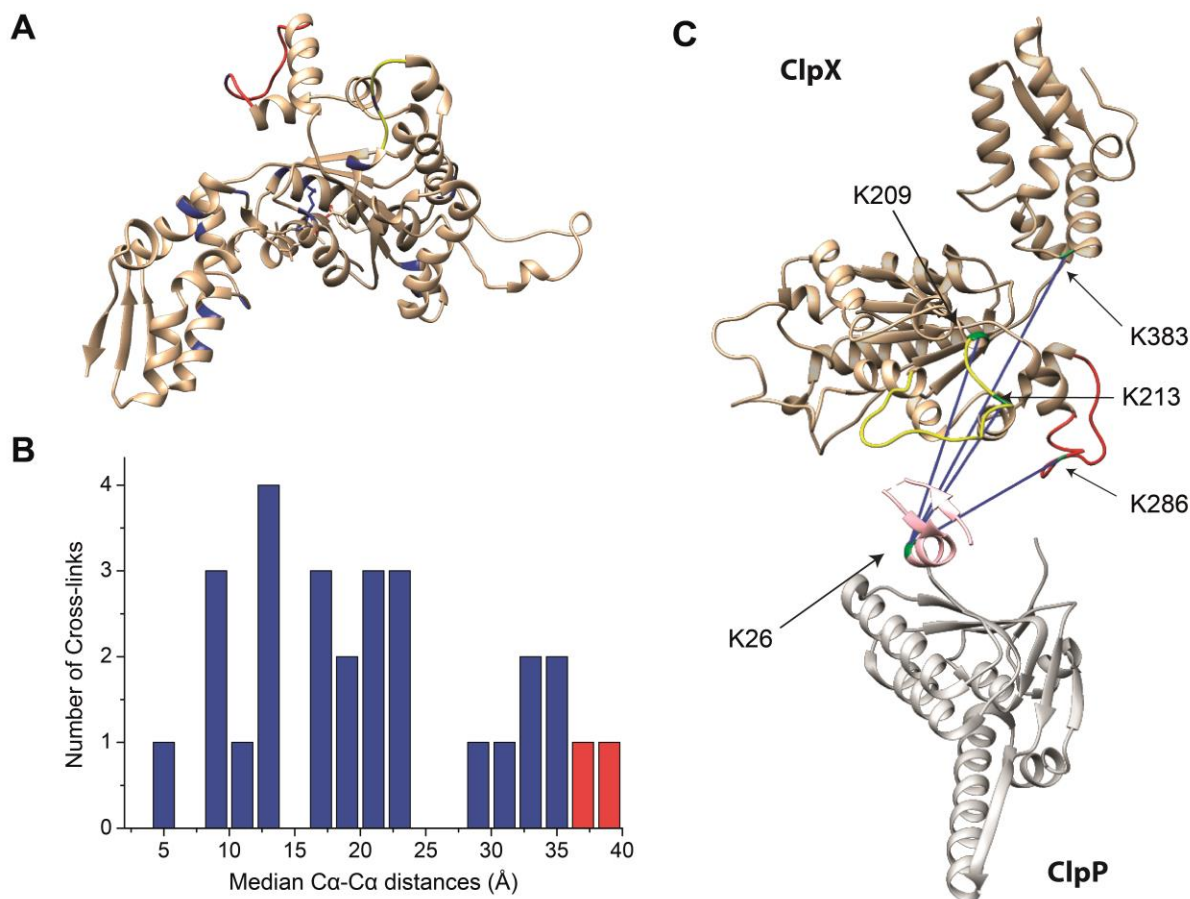
Figure S5 Analysis of the SaClpXP interface. (**A**) The largest proportion of cross-links (FDR<0.01) detected in the SaClpXP sample are intra-ClpX links distributed all over the sequence (involved lysines are highlighted in blue in one monomer). The IGF-loop is colored red, the pore-2-loop yellow. (**B**) Cross-links were displayed onto the *S. aureus* ClpX (SaClpX) hexameric model derived from the *E. coli* crystal structure (PDB: 4I81). The histogram of median Cα-Cα distances displays most of the cross-links within the maximum distance threshold of 35 Å, indicating a good quality of our model. Cross-links exceeding the maximum distance threshold are highlighted in red. (**C**) SaClpX-ClpP cross-linking contacts: One subunit of SaClpP (PDB: 3V5E, colored grey) and one subunit of SaClpX (colored beige) are positioned next to each other and inter-facial cross-links (involved lysine residues highlighted in green) are displayed from SaClpP to SaClpX. The ClpX IGF-loop is colored red and the pore-2-loop yellow. The ClpP N-terminus is colored pink. The SaClpX hexamer model was derived from the *E. coli* pseudohexameric crystal structure (PDB: 4I81) using the SWISS-MODEL platform.[110]

A

B
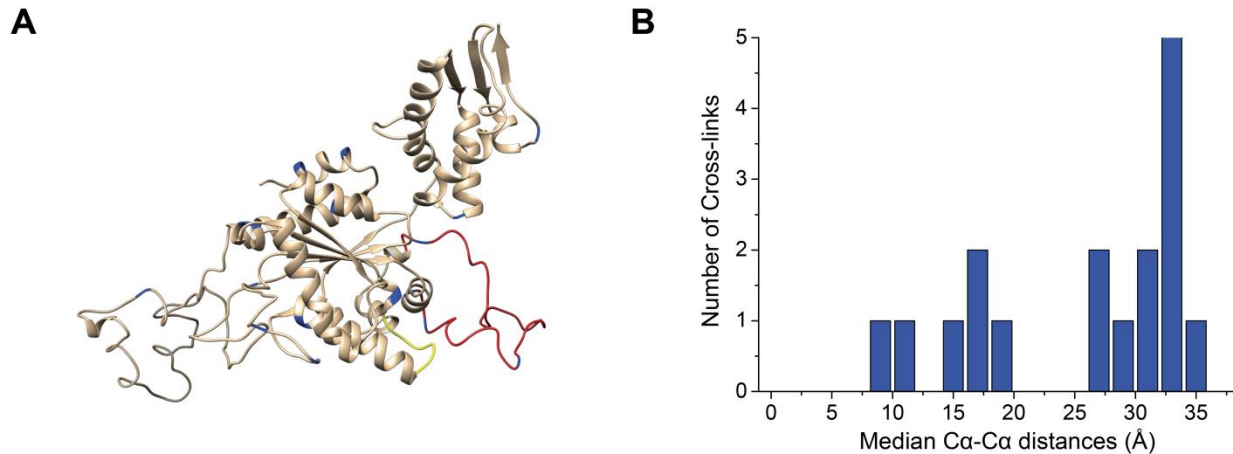
Figure S6 Analysis of the hClpXP interface. (**A**) The largest proportion of cross-links (FDR<0.01) detected in the hClpXP sample are intra-ClpX links distributed all over the sequence (involved lysines are highlighted in blue in one monomer). The LGF-loop is colored red, the pore-2-loop yellow. (**B**) Cross-links were displayed onto the human ClpX (hClpX) hexameric model derived from the *E. coli* crystal structure (PDB: 4I81). The histogram of median Cα-Cα distances displays all of the cross-links within the maximum distance threshold of 35 Å, indicating a good quality of our model. The hClpX hexamer model was derived from the *E. coli* pseudohexameric crystal structure (PDB: 4I81) using the SWISS-MODEL platform.[110]
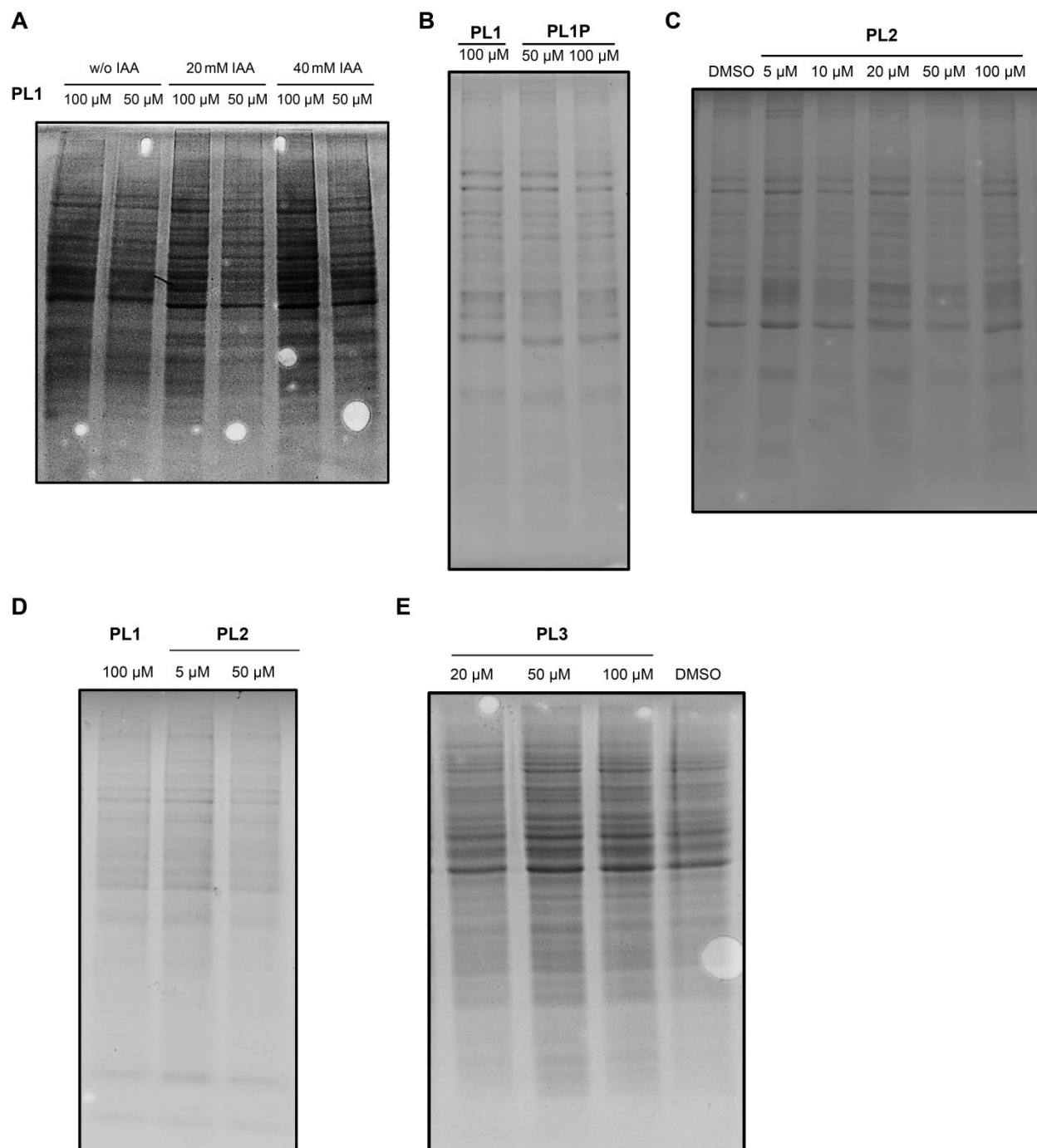
Figure S7 Coomassie-stained gels of PL-probe labeling. Gels representing protein loading of **PL1** labeling and IAA test (**A**), **PL1** and **PL1P** comparison (**B**), **PL2** concentration-dependent labeling (**C**), **PL1** and **PL2** comparison (**D**), and **PL3** labeling (**E**).
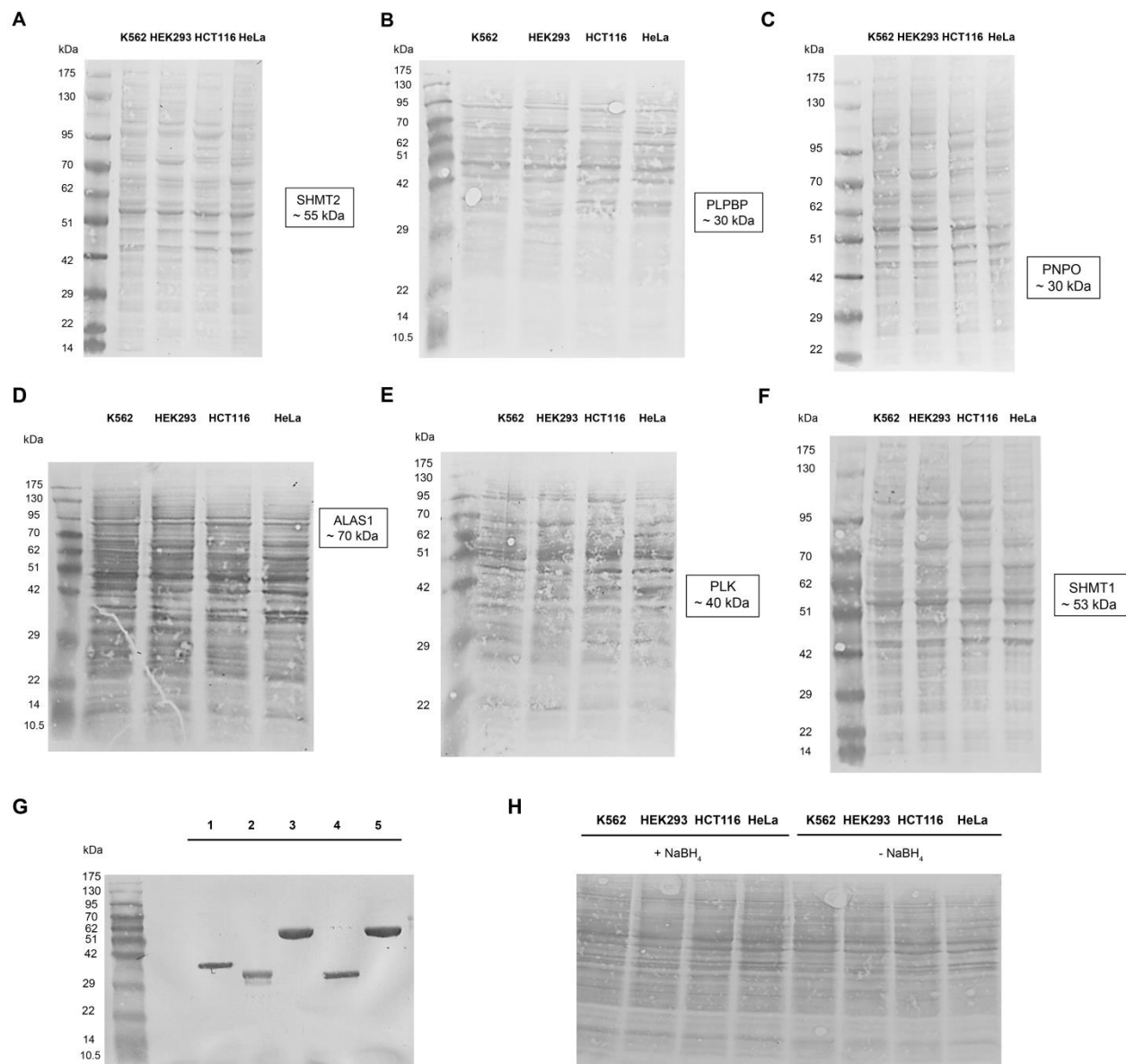
Figure S8 Ponceau S-stained western blot membranes. Membranes of western blots against SHMT2 (**A**), PLPBP (**B**), PNPO (**C**), ALAS1 (**D**), PLK (**E**), and SHMT1 (**F**) reveal equal protein loading amounts for all cell lines. (**G**) Membrane of western blot using the anti-**PL** antibody with hPLK (1), NaBH₄-reduced forms of PLPBP (2) and SHMT1 (3), as well as the non-reduced PLPBP (4) and SHMT1 (5) counterparts. (**H**) Membrane of western blot analysis of cofactor loading states of **PLP**-DEs in NaBH₄-treated as well as non-reduced proteomes reveals equal protein loading.[239]
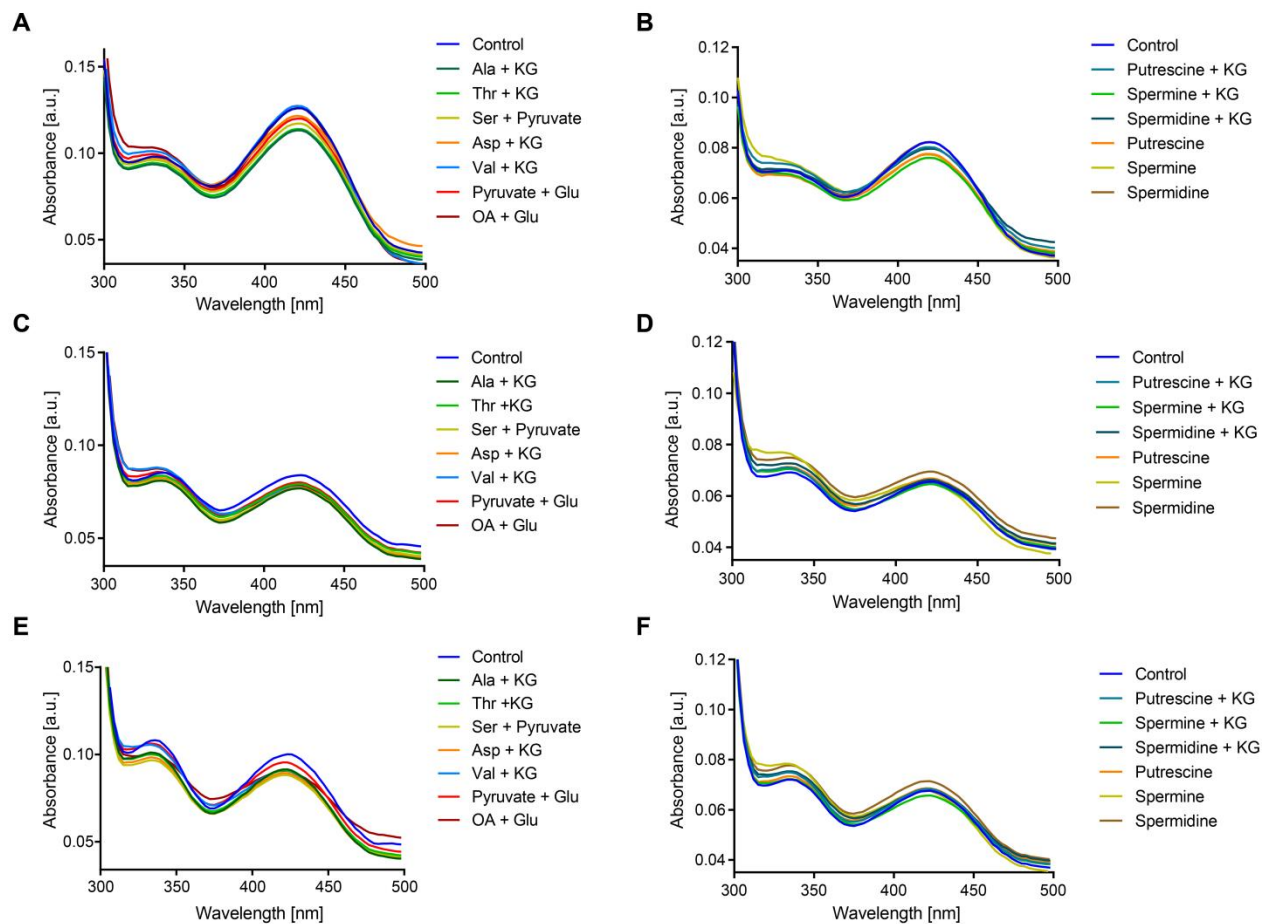
Figure S9 Incubation of PLPBP with transaminase educts. UV/Vis spectra of monomeric PLPBP after 25 min incubation at 37 °C with 20-fold molar excess of amino acids (**B**) and polyamines (**C**). UV/Vis spectra of dimeric PLPBP after incubation at 37 °C with 20-fold molar excess of amino acids for 5 min (**C**) and for 25 min (**E**), as well as with polyamines for 5 min (**D**) and for 25 min (**F**; KG = α-ketoglutaric acid, OA = oxaloacetic acid).

Table S1 Detected cross-links for SaClpP. Cross-links of DSSO between corresponding peptides A and B are shown together with XlinkX scores, position in the crystal structure (PDB: 3V5E, extended and 4EMM, compressed) and Cα-Cα distances. For each cross-link we indicated in which sample it was detected.

1:500_MS3:	No. 1	1:500_EThcD:	No. 2	1:300_MS3:	No. 3
1:300_EThcD:	No. 4	1:1,000_MS3:	No. 5	1:1,000_EThcD:	No. 6

Peptide A	Peptide B	Detected in No.	XlinkX Score (in No.)	Position in Crystal	Median Distance in Crystal (Å; Extended)	Median Distance in Crystal (Å; Compressed)
IQ[K]DTDRDNFLTAEEAK	IQ[K]DTDR	1, 2, 3, 4, 5, 6	371.96, 108.47, 313.82, 105.21, 288.16, 123.79	167_167	14.481	22.105
IQ[K]DTDRDNFLTAEEAK	TGQSIE[K]IQK	2, 3, 4	91.58, 285.85, 154.38	167_164	5.080	5.336
EYGLIDEVMPET[K]WSHPQFEK	TRE[K]LNR	2, 3, 4, 6	69.72, 154.72, 66.57, 70.92	Not Covered	Not Covered	Not Covered
E[K]LNR	LL[K]DR	3	32.12	149_26	34.220	34.241
EYGLIDEVMPET[K]WSHPQFEK	LL[K]DR	3	27.03	Not Covered	Not Covered	Not Covered
IQ[K]DTDRDNFLTAEEAK	E[K]LNR	3, 4	24.64, 28.73	167_149	15.218	14.445
FALPNAEVMIHQPLGGAQATEIEIAANHIL[K]TR	E[K]LNR	4	59.80	145_149	6.221	6.506

Table S2 Detected cross-links for hClpP. Cross-links of DSSO between corresponding peptides A and B are shown together with XlinkX scores, position in the crystal structure (PDB: 1TG6, heptameric structure) and Cα-Cα distances. For each cross-link we indicated in which sample it was detected.

1:300_MS3:	No. 1	1:500_MS3:	No. 2	1:1,000_MS3:	No. 3
1:300_EThcD:	No. 4	1:500_EThcD:	No. 5	1:1,000_EThcD:	No. 6

Peptide A	Peptide B	Detected in No.	XlinkX Score (in No.)	Position in Crystal	Median Distance in Crystal (Å)
GQATDIAIQAEEIM[K]LKK	[K]QLYNIYAK	1, 2, 3, 4, 6	227.94, 273.49, 344.32, 143.42, 121.28, 164.57	144_147	5.194

GQATDIAIQAEEIM[K]LKK	HT[K]QSLQVIESAMER	6	22.10	144_158	20.508

Table S3 Detected cross-links for LmClpP1/2. Cross-links of DSSO between corresponding peptides A and B are shown together with XlinkX scores, position in the crystal structure (PDB: 4RYF) and Cα-Cα distances. For each cross-link we indicated in which sample it was detected.

1:1,000_MS3: No. 1 1:500_MS3: No. 2 1:300_MS3: No. 3

1:1,000_EThcD: No. 4 1:500_EThcD: No. 5 1:300_EThcD: No. 6

Peptide A (ClpP1)	Peptide B (ClpP2)	Detected in No.	XlinkX Score (in No.)	Position in Crystal	Median Distance in Crystal (Å)
YMIHQPAGGVQGQSTEIEIEA[K]EIIR	HIL[K]IK	1, 2, 3, 4, 5, 6	153.93, 251.78, 172.68, 51.93, 66.66, 76.29	141_145	15.780
AENT[K]NENITNILTQK	LL[K]DR	1, 2, 3, 4, 5, 6	110.77, 287.54, 271.49, 47.30, 64.24, 82.92	Not Covered	Not Covered
AENT[K]NENITNILTQK	MNTIMAE[K]TGQPYEVIAR	2, 3, 6	278.55, 253.45, 78.84	Not Covered	Not Covered
AENT[K]NENITNILTQK	DYGLIDDIIIN[K]SGLK	2, 3, 6	153.92, 262.39, 77.73	Not Covered	Not Covered
Peptide A (ClpP1)	**Peptide B (ClpP1)**	**Detected in No.**	**XlinkX Score (in No.)**	**Position in Crystal**	**Median Distance in Crystal (Å)**
AENT[K]NENITNILTQK	YMIHQPAGGVQGQSTEIEIEA[K]EIIR	6	71.70	Not Covered	Not Covered
AENT[K]NENITNILTQK	DYGIVNEIIENRDGL[K]WSHPQFEK	6	121.25	Not Covered	Not Covered
AENT[K]NENITNILTQK	FI[K]PTVK	2, 3, 5, 6	289.35, 265.36, 51.29, 70.12,	Not Covered	Not Covered
AENTKNENITNILTQ[K]	ENTKNENITNILTQ[K]LIDTR	2,3, 5, 6	81.64, 301.74, 204.47, 205.75	24_24	16.628
FI[K]PTVK	DGL[K]WSHPQFEK	3, 6	170.88, 29.80	Not Covered	Not Covered
Peptide A (ClpP2)	**Peptide B (ClpP2)**	**Detected in No.**	**XlinkX Score (in No.)**	**Position in Crystal**	**Median Distance in Crystal (Å)**
MNTIMAE[K]TGQPYEVIAR	HIL[K]IK	3, 5, 6	188.41, 39.51, 57.99	157_145	18.613
MNTIMAE[K]TGQPYEVIAR	LL[K]DR	2, 3, 5	225.80, 317.67, 38.09	157_26	29.053
DYGLIDDIIIN[K]SGLK	ERMNTIMAE[K]TGQPYEVIAR	1, 2, 3, 4, 5, 6	118.52, 258.09, 154.69, 64.30,	193_157	25.996

			168.27, 88.39,		
DYGLIDDIIIN[K]SGLK	LL[K]DR	2	131.69	193_26	18.670
DTDRDNFMTAQEA[K]DYGLIDDIIINK	HIL[K]IK	6	71.35	181_145	17.881
AYDIYSRLL[K]DR	LL[K]DR	6	28.17	26_26	18.086

Table S4 Detected cross-links for EcClpP. Cross-links of BS3 between corresponding peptides A and B are shown together with position in the crystal structure (PDB: 1YG6) and Cα-Cα distances. For each cross-link we indicated in which sample it was detected.

1:300_BS3:   No. 1   1:500_BS3:   No. 2

Peptide A	Peptide B	Detected in No.	Position in Crystal	Median Distance in Crystal (Å)
LL[K]ER	LL[K]ER	1, 2	25_25	17.821
EILKV[K]GR	EIL[K]VK	1	146_144	5.423
EIL[K]VK	EIL[K]VK	1, 2	144_144	21.587

Table S5 Detected cross-links for SaClpXP complex. Cross-links of DSSO between corresponding peptides A and B are shown together with XlinkX scores, position in the crystal structure (ClpX derived from PDB: 4I81 and ClpP is PDB: 3V5E) and Cα-Cα distances. For each cross-link we indicated in which sample it was detected.

1:80_MS3:   No. 1   1.80_EThcD: No. 2   1:137.5_MS3:   No. 3   1:137.5_EThcD: No. 4

Peptide A (ClpP)	Peptide B (ClpP)	Detected in No.	XlinkX Score (in No.)	Position in Crystal	Median Distance in Crystal (Å)
EYGLIDEVMVPET[K]WSHPQFEK	E[K]LNR	1, 2, 4	179.84, 73.25, 83.26	Not covered	Not Covered
LL[K]DR	E[K]LNR	1, 2, 3	110.54, 24.64, 49.48	26_149	34.252
IQ[K]DTDRDNFLTAEEAK	TGQSIE[K]IQK	3, 4	136.58, 116.00	164_167	5.048
Peptide A (ClpX)	**Peptide B (ClpP)**	**Detected in No.**	**XlinkX Score (in No.)**	**Position in Crystal**	**Median Distance in Crystal (Å)**
[K]SENTSITRDVSGEGVQQALLK	LL[K]DR	1, 2	349.41, 60.44	213(X)_26(P)	N/A
[K]TGAR	EYGLIDEVMVPET[K]WSHPQFEK	1	130.50	Not covered	N/A
[K]TGAR	LL[K]DR	1, 2, 3	109.94, 59.80, 109.94	383(X)_26(P)	N/A
GIIYVDQID[K]IAR	LL[K]DR	1, 3, 4	77.08, 59.39, 35.62	209(X)_26(P)	N/A
LGE[K]VIGFSSNEADK	LL[K]DR	3, 4	155.09, 82.92	286(X)_26(P)	N/A

Peptide A (ClpX)	Peptide B (ClpX)	Detected in No.	XlinkX Score (in No.)	Position in Crystal	Median Distance in Crystal (Å)
EIMDHLNEYVIGQE[K]AK	[K]SLAVAVYNHYK	1, 2, 3, 4	399.40, 179.25, 406.71, 133.48	103_106	5.170
LIQAADFDIDKAE[K]GIIYVDQIDK	RIQQLGP[K]EDDVELQK	1, 2, 3, 4	332.36, 163.67, 329.08, 198.87	199_125	16.955
RIQQLGP[K]EDDVELQK	NILTQP[K]NALVK	1, 2, 3, 4	325.77, 159.19, 261.32, 79.60	125_347	14.362
RIQQLGP[K]EDDVELQK	NILTQPKNALV[K]QYTK	1, 2, 3, 4	309.78, 126.83, 296.11, 122.46	125_352	13.747
[K]SENTSITRDVSGEGVQQALLK	[K]SENTSITR	1, 2	307.67, 96.73	213_213	19.433
IQQLGP[K]EDDVELQK	[K]AGFENLYFQGFK	1, 2, 3, 4	284.26, 122.74, 251.01, 117.93	Not Covered	Not Covered
IQQLGP[K]EDDVELQK	LIQAADFDID[K]AEK	1, 2, 3, 4	281.94, 253.91, 253.50, 218.79	125_196	19.604
IQQLGP[K]EDDVELQK	SLAVAVYNHY[K]R	1, 2, 3, 4	279.20, 165.02, 263.50, 153.36	125_117	11.994
IQQLGP[K]EDDVELQK	DQDQV[K]K	1, 2, 3, 4	266.44, 108.47, 342.11, 97.81	Not Covered	Not Covered
EIMDHLNEYVIGQE[K]AK	[K]TGAR	1, 2, 3, 4	250.81, 71.50, 284.86, 60.44	103_383	21.254
DVSGEGVQQALL[K]ILEGTTASVPPQGGR	IQQLGP[K]EDDVELQK	1, 2, 3, 4	236.22, 161.68, 251.02, 121.62	234_125	22.021
GIIYVDQID[K]IAR	[K]SENTSITRDVSGEGVQQALLK	1, 2, 3, 4	195.98, 181.68, 222.28, 109.00	209_213	9.940
SNIALIGPTGSG[K]TLLAQTLAK	[K]SENTSITR	1, 2, 3, 4	168.86, 58.07, 131.09, 67.12	146_213	23.615
RIQQLGP[K]EDDVELQK	[K]SENTSITR	1, 2	158.91, 113.82	125_213	37.713
LIQAADFDIDKAE[K]GIIYVDQIDK	[K]SENTSITR	1	154.14	199_213	34.262
GIIYVDQID[K]IAR	[K]TGAR	1, 2, 3, 4	147.24, 82.60, 142.76, 72.47	209_383	16.944
QYT[K]MLELDDVDLEFTEEALSAISEK	IQQLGP[K]EDDVELQK	1, 2, 3, 4	146.54, 151.27, 258.62, 162.50	356_125	9.649
VPIVANLETLDVTAL[K]NILTQPK	NALV[K]QYTK	1, 2, 3, 4	110.24, 98.03, 119.23, 109.68	340_352	12.917
QYT[K]MLELDDVDLEFTEEALSAISEK	NILTQP[K]NALVK	1, 2, 3, 4	93.26, 84.24, 139.29, 107.66	356_347	11.885
VVITAQTINEETEPELYDAEGNLINNS[K]TSA	[K]SLAVAVYNHYK	1, 2	97.47, 85.36	Not Covered	Not Covered
DVSGEGVQQALL[K]ILEGTTASVPPQGGR	[K]SENTSITR	1, 2, 3, 4	51.69, 70.89, 51.69,98.94	234_213	21.767

DVSGEGVQQALL[K]ILEGTTASVPPQGGR	GIIYVDQID[K]IAR	1, 2, 3, 4	50.70, 71.26, 43.81, 63.86	234_209	19.122
LIQAADFDID[K]AEK	[K]SENTSITR	2, 3, 4	97.55, 234.41, 89.23	196_213	34.191
LIQAADFDID[K]AEKGIIYVDQIDK	DQDQV[K]K	3, 4	208.36, 119.53	Not Covered	Not Covered
LIQAADFDID[K]AEK	[K]TGAR	3	110.55	196_383	37.508
MLELDDVDLEFTEEALSAISE[K]AIER	[K]TGAR	4	59.25	378_383	8.808
SNIALIGPTGSG[K]TLLAQTLAK	IQQLGP[K]EDDVELQK	2	125.81	146_125	31.926
DVSGEGVQQALL[K]ILEGTTASVPPQGGR	LIQAADFDID[K]AEK	2	113.89	234_196	20.710
IQQLGP[K]EDDVELQK	GIIYVDQID[K]IAR	2, 4	82.45, 53.77	125_209	30.161
VIGFSSNEAD[K]YDEQALLAQIRPEDLQAYGLIPEFIGR	[K]TGAR	2, 4	50.91, 50.91	297_383	20.326
[K]AGFENLYFQGFK	[K]SENTSITR	4	79.03	Not Covered	Not Covered
NALV[K]QYTK	[K]SENTSITR	4	55.21	352_213	34.947
DVSGEGVQQALL[K]ILEGTTASVPPQGGR	[K]TGAR	4	37.72	234_383	33.759

Table S6 Detected cross-links for hClpXP complex. Cross-links of DSSO between corresponding peptides A and B are shown together with XlinkX scores, position in the crystal structure (ClpX derived from PDB: 4I81 and ClpP is PDB: 1TG6) and Cα-Cα distances. For each cross-link we indicated in which sample it was detected.

1:80_MS3:	No. 1		1.80_EThcD:	No. 2		1:100_MS3:	No. 3

1:100_EThcD:	No.4		1:137.5_MS3:	No. 5		1:137.5_EThcD:	No. 6

Peptide A (ClpP)	Peptide B (ClpP)	Detected in No.	XlinkX Score (in No.)	Position in Crystal	Median Distance in Crystal (Å)
GQATDIAIQAEEIM[K]LKK	[K]QLYNIYAK	1, 2, 3, 4, 5, 6	208.66, 141.38, 228.90, 153.38, 214.54, 147.33	144_147	5.194
Peptide A (ClpX)	**Peptide B (ClpX)**	**Detected in No.**	**XlinkX Score (in No.)**	**Position in Crystal**	**Median Distance in Crystal (Å)**
LLEGTIVNVPE[K]NSR	DGIS[K]DGSGDGNKK	1, 2, 3, 4	291.92, 92.29, 193.94, 127.97	Not Covered	Not Covered
[K]SIIKEPESAAEAVK	DGSGDGN[K]K	1, 2, 3, 4	224.13, 86.29, 221.30, 86.29	Not Covered	Not Covered
[K]SIIKEPESAAEAVK	SII[K]EPESAAEAVK	1, 2, 3, 4	224.13, 156.39, 226.76, 157.52	Not Covered	Not Covered

IYNNIPANLRQQAEVE[K]QTSLTPR	SII[K]EPESAAEAVK	1, 2, 4, 6	216.06, 107.45, 132.74, 78.41	Not Covered	Not Covered
SII[K]EPESAAEAVK	SASEGSS[K]K	1, 2, 3, 5, 6	212.68, 118.37, 153.81, 207.12, 85.74	Not Covered	Not Covered
LAFQQ[K]PPPPPK	[K]IYNYLDK	1, 2, 3, 4, 5, 6	203.59, 113.83, 182.55, 91.32, 156.01, 106.49	Not Covered	Not Covered
LLQDANYNVE[K]AQQGIVFLDEVDK	SII[K]EPESAAEAVK	1, 2, 5	199.27, 101.28, 103.76	Not Covered	Not Covered
SII[K]EPESAAEAVK	DGIS[K]DGSGDGNK	1, 2, 3, 4, 5	198.89, 43.18, 157.56, 95.67, 74.85	Not Covered	Not Covered
SII[K]EPESAAEAVK	DGSGDGN[K]K	1, 2, 3, 4, 5, 6	179.28, 107.53, 189.72, 91.64, 186.41, 57.29	Not Covered	Not Covered
LLEGTIVNVPE[K]NSRK	LLEGTIVNVPE[K]NSR	1, 2, 3, 4, 5, 6	177.58, 178.77, 171.38, 186.56, 160.75, 196.31	354_354	16.990
LLEGTIVNVPE[K]NSRK	[K]SGSGNSGK	1, 2, 3, 4, 5, 6	175.88, 75.59, 118.21, 109.00, 108.43, 49.48	Not Covered	Not Covered
EVVEG[K]KEPGYIR	R[K]NEK	1, 2, 3, 5	170.18, 34.66, 189.52, 200.13	553_390	19.610
LLEGTIVNVPE[K]NSR	SASEGSS[K]K	1, 2, 3, 4, 5, 6	168.67, 86.47, 128.71, 132.14, 136.02, 86.47	Not Covered	Not Covered
YLGFGTPSNLG[K]GR	[K]TGAR	1, 3, 5	160.11, 178.85, 129.73	405_514	33.316
KSII[K]EPESAAEAVK	LAFQQ[K]PPPPPKK	1, 2, 3, 4, 5, 6	156.01, 127.12, 109.11, 130.04, 151.98, 123.99	Not Covered	Not Covered
LLEGTIVNVPE[K]NSR	DGSGDGN[K]K	1, 2, 3, 4, 5	155.64, 69.31, 182.37, 75.59, 153.05	Not Covered	Not Covered
AAAAADLANRSGESNTHQDIEE[K]DRLLR	[K]TGAR	1, 2, 3, 4, 5, 6	152.59, 65.09, 213.91 72.47, 198.31, 70.43	431_514	13.455
IYNNIPANLRQQAEVE[K]QTSLTPR	[K]IYNYLDK	1, 2, 3, 4, 5, 6	145.65, 106.01, 151.50, 129.84, 149.65, 79.49	173_127	30.121
LLEGTIVNVPE[K]NSR	SII[K]EPESAAEAVKLAFQQKPPPPPK	1, 2, 3, 4, 5, 6	142.10, 156.67, 149.68, 138.03,	Not Covered	Not Covered

			188.68, 138.03		
SII[K]EPESAAEAVK	[K]IYNYLDK	1, 2, 3, 4, 5, 6	139.77, 115.33, 142.08, 101.71, 89.26, 82.36	Not Covered	Not Covered
DVGGEGVQQGLL[K]LLEGTIVNVPEK	LLEGTIVNVPE[K]NSRK	1, 2, 3, 4, 5, 6	139.65, 117.10, 113.52, 65.42, 128.06, 89.40	342_354	25.046
LLQDANYNVE[K]AQQGIVFLPEVDK	LAFQQ[K]PPPPPKK	1, 2, 3, 4, 5, 6	138.63, 107.65, 171.74, 143.37, 137.74, 107.65	Not Covered	Not Covered
YLGFGTPSNLG[K]GR	SASEGSS[K]K	1, 2	136.93, 33.80	Not Covered	Not Covered
LAFQQ[K]PPPPPK	[K]SASEGSSK	1	135.43	Not Covered	Not Covered
IYNNIPANLRQQAEVE[K]QTSLTPR	LLEGTIVNVPE[K]NSRK	1, 2, 3, 4, 5, 6	132.57, 128.41, 122.74, 147.34, 140.78, 129.02	173_354	20.900
GGEVLDSSHDDI[K]LEK	[K]IYNYLDK	1, 3, 4, 5	120.75, 136.27, 50.56, 96.95	239_127	26.932
AQQGIVFLDEVD[K]IGSVPGIHQLRDVGGEGVQQGLLK	[K]TGAR	1, 2, 3, 4, 5, 6	110.89, 46.13, 87.81, 36.69, 86.10, 46.13	318_514	16.966
QQAEVE[K]QTSLTPR	LAFQQ[K]PPPPPK	1, 3, 4, 5, 6	110.61, 94.41, 94.41, 110.61, 87.62	Not Covered	Not Covered
SII[K]EPESAAEAVK	[K]SGSGNSGK	1, 2, 3, 5	101.42, 67.12, 89.86, 97.27	Not Covered	Not Covered
GFTETPAYFAS[K]DGISK	SII[K]EPESAAEAVK	1, 2, 3, 4, 5, 6	97.43, 139.17, 55.06, 124.00, 63.82, 66.45	Not Covered	Not Covered
LLQDANYNVE[K]AQQGIVFLDEVDK	QQAEVE[K]QTSLTPR	1, 2, 3, 4, 5	96.19, 60.97, 113.89, 120.62, 86.79	305_173	33.437
IYNNIPANLRQQAEVE[K]QTSLTPR	SASEGSS[K]K	1, 2	76.24, 45.86	Not Covered	Not Covered
GFTETPAYFAS[K]DGISK	LLEGTIVNVPE[K]NSR	1, 2, 4, 5	60.56, 100.44, 62.58, 55.06	Not Covered	Not Covered
SII[K]EPESAAEAVK	YLGFGTPSNLG[K]GR	1, 2	51.35, 89.48	Not Covered	Not Covered
LLEGTIVNVPE[K]NSR	[K]SIIKEPESAAEAVK	1, 3, 4, 5, 6	34.93, 69.63, 120.23, 201.86, 111.66	Not Covered	Not Covered
GFTETPAYFAS[K]DGISK	DGSGDGN[K]K	1, 3, 4, 6	34.89, 50.12, 125.89, 122.46	Not Covered	Not Covered
[K]SIIKEPESAAEAVK	[K]IYNYLDK	2, 4	95.40, 84.57	Not Covered	Not Covered
GFTETPAYFAS[K]DGISK	[K]SGSGNSGK	2	93.35	Not Covered	Not Covered
IYNNIPANLRQQAEVE[K]QTSLTPR	LAFQQ[K]PPPPPK	2	88.03	Not Covered	Not Covered

[K]SIIKEPESAAEAVK	[K]SGSGNSGK	2, 3	64.87, 39.54	Not Covered	Not Covered
SGESNTHQDIEE[K]DRLLR	YLGFGTPSNLG[K]GR	2, 3, 4, 5, 6	60.76, 121.93, 49.49, 151.51, 34.01	431_405	22.793
GFTETPAYFAS[K]DGISK	SASEGSS[K]K	2, 4, 6	51.64, 121.28, 70.89	Not Covered	Not Covered
[K]SIIKEPESAAEVK	LAFQQ[K]PPPPPK	2, 3, 4	38.88, 256.05, 94.30	Not Covered	Not Covered
LAFQQ[K]PPPPPK	DGSGDGN[K]K	2, 3, 4, 5	36.68, 169.73, 51.35, 149.49	Not Covered	Not Covered
GFTETPAYFAS[K]DGISK	[K]SASEGSSKK	2, 4, 5	33.70, 44.16, 50.12	Not Covered	Not Covered
LAFQQ[K]PPPPPK	SASEGSS[K]K	2, 3	32.13, 160.67	Not Covered	Not Covered
NE[K]YLGFGTPSNLGK	[K]TGAR	2, 3, 4	32.13, 40.95, 47.30	393_514	30.543
[K]SIIKEPESAAEAVK	SGSGNSG[K]GGNQLR	3, 5	242.48, 139.30	Not Covered	Not Covered
SII[K]EPESAAEAVK	IYNNIPANLRQQAEVE[K]QTSLTPR	3, 5	191.06, 139.73	Not Covered	Not Covered
LLEGTIVNVPE[K]NSRK	[K]SASEGSSK	3	120.64	Not Covered	Not Covered
YLGFGTPSNLG[K]GR	R[K]NEK	3, 5	79.87, 114.02	405_390	16.977
SII[K]EPESAAEAVK	SGSGNSG[K]GGNQLR	3	59.86	Not Covered	Not Covered
SIIKEPESAAEAV[K]LAFQQKPPPPPK	LLEGTIVNVPE[K]NSR	4	103.31	Not Covered	Not Covered
RGGEVLDSSHDDI[K]LEK	IYNYLD[K]YVVGQSFAK	4	81.48	239_134	32.758
[K]SIIKEPESAAEAVK	SASEGSS[K]K	4	76.59	Not Covered	Not Covered
[K]SIIKEPESAAEAVK	[K]SASEGSSK	4	75.46	Not Covered	Not Covered
SII[K]EPESAAEAVK	[K]SASEGSSK	4	72.68	Not Covered	Not Covered
[K]SIIKEPESAAEAVK	QQAEVE[K]QTSLTPR	4	60.97	Not Covered	Not Covered
SIIKEPESAAEAV[K]LAFQQKPPPPPK	[K]IYNYLDK	4	54.54	Not Covered	Not Covered
NAVIPQYQALFSMD[K]CELNVTEDALK	LAFQQ[K]PPPPPK	4	51.21	Not Covered	Not Covered
RGGEVLDSSHDDI[K]LEK	QQAEVE[K]QTSLTPR	4	38.32	Not Covered	Not Covered
LLQDANYNVE[K]AQQGIVFLDEVDK	LLEGTIVNVPE[K]NSR	4	34.68	305_354	25.179
LAFQQKPPPPP[K]K	[K]VLSVAVYNHYK	5	177.57	126_144	8.048
[K]LRGETVQVDTTNILFVASGAFNGLDR	LLEGTIVNVPE[K]NSR	5	106.04	358_354	11.538
[K]SGSGNSGK	[K]IYNYLDK	5	87.31	Not Covered	Not Covered

Table S7 Detected cross-links for LmClpXP1P2 complex. Cross-links of DSSO between corresponding peptides A and B are shown together with XlinkX scores, position in the crystal structure (ClpX derived from PDB: 4I81 and ClpP is PDB: 4RYF) and Cα-Cα distances. For each cross-link we indicated in which sample it was detected.

1:80_MS3:	No. 1		1.80_EThcD:	No. 2	
1:100_EThcD:	No. 4		1:137.5_MS3:	No. 5	

1:100_MS3:	No. 3
1:137.5_EThcD:	No. 6

Peptide A (ClpP1)	Peptide B (ClpP1)	Detected in No.	XlinkX Score (in No.)	Position in Crystal	Median Distance in Crystal (Å)
NENITNILTQ[K]LIDTR	AENTKNENITNILTQ[K]NENITNILTQK	1, 3, 4, 6	304.21, 50.80, 179.54, 192.60	Not Covered	Not Covered
FI[K]PTVK	AENT[K]NENITNILTQK	2, 5, 6	72.65, 169.20, 62.41	Not Covered	Not Covered
NENITNILTQ[K]LIDTR	AENT[K]NENITNILTQK	2, 5	186.13, 261.38	Not Covered	Not Covered
Peptide A (ClpP2)	**Peptide B (ClpP2)**	**Detected in No.**	**XlinkX Score (in No.)**	**Position in Crystal**	**Median Distance in Crystal (Å)**
MNTIMAE[K]TGQPYEVIAR	DYGLIDDIIIN[K]SGLK	1, 2, 3, 4, 5, 6	308.59, 195.06, 328.56, 93.18, 235.40, 179.68	157_193	25.996
DTDRDNFMTAQEA[K]DYGLIDDIIINK	LL[K]DR	1	153.29	181_26	29.737
Peptide A (ClpP2)	**Peptide B (ClpP1)**	**Detected in No.**	**XlinkX Score (in No.)**	**Position in Crystal**	**Median Distance in Crystal (Å)**
HIL[K]IK	YMIHQPAGGVQGQSTEIEIEA[K]EIIR	1, 2, 3, 4, 5, 6	222.20, 66.66, 200.59, 82.50, 254.20, 90.13	145_141	15.780
LL[K]DR	AENT[K]NENITNILTQK	1, 3, 6	151.04, 169.20, 62.58	Not Covered	Not Covered
Peptide A (ClpP2)	**Peptide B (ClpX)**	**Detected in No.**	**XlinkX Score (in No.)**	**Position in Crystal**	**Median Distance in Crystal (Å)**
LL[K]DR	MGE[K]VIGFGTDNAK	1, 2, 4, 5, 6	272.04, 103.52, 91.27, 290.61, 91.27	26_262	N/A
LL[K]DR	[K]SENPSITR	1, 2, 3, 4, 5, 6	273.62, 70.43, 202.39, 65.09, 187.10, 64.24	26_190	N/A
DYGLIDDIIIN[K]SGLK	L[K]DDETYLSR	1, 2, 3, 4, 5, 6	170.85, 127.26, 202.20, 101.36, 209.39, 113.63	193_274	N/A
DYGLIDDIIIN[K]SGLK	MGE[K]VIGFGTDNAK	1, 2, 3, 5, 6	220.06, 49.42, 117.48, 202.14, 71.09	193_262	N/A

Peptide A	Peptide B	Detected in No.	XlinkX Score (in No.)	Position in Crystal	Median Distance in Crystal (Å)
LL[K]DR	[K]TGAR	1, 2, 3, 4, 5, 6	113.82, 59.80, 109.80, 72.47, 107.65, 73.25	26_359	N/A
LL[K]DR	VIGFGTDNA[K]LKDDETYLSR	1, 2, 4, 5, 6	91.27, 72.47, 72.47, 35.62, 46.13	26_272	N/A
LL[K]DR	GIIYIDQID[K]VAR	1, 2, 3, 4, 6	204.00, 75.90, 265.65, 75.90, 75.90, 59.39	26_186	N/A
DTDRDNFMTAQEA[K]DYGLIDDIIINK	L[K]DDETYLSR	1, 5	33.12, 33.12	181_274	N/A
HIL[K]IK	[K]TGAR	1	125.86	145_359	N/A
DYGLIDDIIIN[K]SGLK	VIGFGTDNA[K]LKDDETYLSR	5	228.09	193_272	N/A
MNTIMAE[K]TGQPYEVIAR	[K]SENPSITR	1, 3, 4, 5, 6	188.41, 175.80, 79.03, 210.98, 47.91	157_190	N/A
Peptide A (ClpX)	**Peptide B (ClpX)**	**Detected in No.**	**XlinkX Score (in No.)**	**Position in Crystal**	**Median Distance in Crystal (Å)**
RINSNET[K]EDEVELSK	[K]SENPSITR	1, 2, 3, 4, 5, 6	354.23, 128.61, 389.75, 110.24, 417.21, 119.39	102_190	33.113
RINSNET[K]EDEVELSK	FNDE[K]GQLK	1, 2, 3, 5, 6	394.82, 127.57, 431.45, 402.05, 136.19	Not Covered	Not Covered
RINSNET[K]EDEVELSK	NALV[K]QYKR	1, 2, 3, 4, 5	429.70, 123.07, 420.30, 119.22, 435.18	102_328	13.119
RINSNET[K]EDEVELSK	[K]TGAR	1, 2, 3, 4	389.75, 47.30, 314.98, 456.51	102_359	34.061
INSNET[K]EDEVELSK	KALAVAVYNHY[K]R	1, 2, 3, 4, 5, 6	391.82, 109.11, 394.82, 123.99, 382.06, 120.29	102_94	11.420
RINSNET[K]EDEVELSK	LIQSADYDVE[K]AEKGIIYIDQIDK	1, 2, 3, 4, 5, 6	349.81, 221.28, 331.82, 223.75, 223.75, 236.03	102_173	19.526
LPVIATLEQLDEAALVSILTEP[K]NALVK	INSNET[K]EDEVELSK	1, 2, 3, 4, 5, 6	270.07, 140.63, 37.19, 124.34, 131.39, 129.02	323_102	13.944
INSNET[K]EDEVELSK	MGE[K]VIGFGTDNAK	1, 2, 4, 5, 6	314.50, 98.30, 102.07, 309.06, 132.74	102_262	41.321
RINSNET[K]EDEVELSK	LIQSADYDVEKAE[K]GIIYIDQIDK	2, 3, 4, 5, 6	183.26, 326.36, 215.10, 349.52, 181.81	102_176	16.620
GIIYIDQID[K]VARK	[K]SENPSITRDVSGEGVQQALLK	1, 2, 3, 4,	217.65, 181.68,	186_190	9.810

		5, 6	312.82, 126.07, 202.05, 117.54		
LIQSADYDVE[K]AEK	FNDE[K]GQLK	1, 2, 4, 5, 6	236.89, 122.46, 67.12, 88.63, 89.46	Not Covered	Not Covered
VVPEDLL[K]FGLIPEFIGR	MGE[K]VIGFGTDNAK	1, 2, 3, 4, 5, 6	268.74, 84.11, 241.78, 110.60, 216.68, 87.77	290_262	20.652
MGE[K]VIGFGTDNAK	[K]TGAR	1, 2, 3, 4, 5, 6	218.40, 59.17, 224.25, 59.17, 299.11, 80.63	262_359	29.034
[K]SENPSITRDVSGEGVQQALLK	[K]SENPSITR	1, 2, 3, 4, 5, 6	278.40, 148.62, 319.00, 130.65, 351.70, 127.57	190_190	16.029
AE[K]GIIYIDQIDK	FNDE[K]GQLK	5, 6	282.29, 118.37	Not Covered	Not Covered
L[K]DDETYLSR	[K]SENPSITR	1, 2, 3, 4, 5, 6	183.00, 127.57, 210.67, 136.58, 193.04, 138.38	274_190	24.737
MGE[K]VIGFGTDNAK	[K]SENPSITR	1, 2, 3, 4, 5	259.06, 98.03, 262.65, 119.23, 260.80	262_190	25.000
[K]ALAVAVYNHYKR	L[K]DDETYLSR	1, 2, 3, 4, 5, 6	228.66, 154.11, 223.79, 143.88, 229.99, 144.73	83_274	23.954
RINSNET[K]EDEVELSK	GIIYIDQID[K]VAR	5	223.14	102_186	29.699
LIQSADYDVE[K]AEK	[K]SENPSITR	1, 2, 3, 4, 5, 6	204.59, 103.41, 225.68, 119.23, 210.67, 113.82	173_190	35.910
VVPEDLL[K]FGLIPEFIGR	GIIYIDQID[K]VAR	1, 2	65.25, 65.25	290_186	13.810
SNICLIGPTGSG[K]TLLAQTLAR	INSNET[K]EDEVELSK	3, 6	226.09, 105.25	123_102	31.196
AE[K]GIIYIDQIDK	ALAVAVYNHY[K]R	2, 5, 6	120.29, 179.93, 137.61	176_94	11.203
LIQSADYDVE[K]AEK	[K]TGAR	1, 2, 3, 5, 6	142.67, 40.44, 134.37, 164.56, 40.44	173_359	37.478
DVSGEGVQQALL[K]ILEGTVASVPPQGGRK	LIQSADYDVE[K]AEK	1, 2	32.32, 104.37	211_173	20.684
VVPEDLL[K]FGLIPEFIGR	[K]SENPSITR	1, 2, 3, 4, 5, 6	169.93, 96.73, 187.97, 109.68, 175.41, 76.80	290_190	7.798
MLELDDVELEFEPTALIEIA[K]EAIER	[K]SENPSITR	1, 2	166.66, 32.13	353_190	32.318
AARGEEEPQLQLEDGSIIPI[K]TSA	[K]ALAVAVYNHYKR	1, 2, 3, 4, 5, 6	209.80, 157.11, 169.93, 153.95,	Not Covered	Not Covered

			193.83, 155.03		
[K]SENPSITR	[K]TGAR	1, 2, 4, 5, 6	164.04, 86.83, 52.48, 135.44, 52.48	190_359	22.803
VIGFGTDNA[K]LKDDETYLSR	CIITE[K]AAR	2, 4, 5, 6	68.03, 98.03, 132.76, 49.48	272_394	28.188
NALV[K]QYKR	[K]TGAR	1, 3, 5, 6	90.14, 82.26, 91.55, 40.95	328_359	23.692
L[K]DDETYLSR	[K]TGAR	1, 2, 3, 4, 5, 6	64.24, 55.37, 32.12, 50.04, 238.36, 50.04	274_359	19.658
DVSGEGVQQALL[K]ILEGTVASVPPQGGR	MGE[K]VIGFGTDNAK	1	98.87	211_262	36.417
VVPEDLL[K]FGLIPEFIGR	INSNET[K]EDEVELSK	1, 2	222.75, 41.13	290_102	34.307
DVSGEGVQQALL[K]ILEGTVASVPPQGGR	L[K]DDETYLSR	1, 2, 5	45.31, 45.31, 47.91	211_274	32.990
DVSGEGVQQALL[K]ILEGTVASVPPQGGR	[K]SENPSITR	1, 2, 4, 5, 6	143.40, 67.23, 47.91, 47.91, 47.91	211_190	21.686
VIGFGTDNA[K]LKDDETYLSR	[K]TGAR	1, 2, 3, 4, 5, 6	32.12, 32.12, 49,86	272_359	24.354
DVSGEGVQQALL[K]ILEGTVASVPPQGGR	[K]TGAR	1, 2, 4	40.95, 40.35, 38.09	211_359	33.769
AARGEEEPQLQLEDGSIIPI[K]TSA	RINSNET[K]EDEVELSK	2, 3	133.46, 236.17	Not Covered	Not Covered
LIQSADYDVE[K]AEK	GIIYIDQID[K]VAR	1, 2, 4	170.06, 87.32, 158.36	173_186	30.572
MGE[K]VIGFGTDNAK	L[K]DDETYLSR	2, 3, 4, 6	138.53, 259.59, 138.75, 166.39	262_274	11.347
VVPEDLL[K]FGLIPEFIGR	L[K]DDETYLSR	3, 4	216.06, 98.94	290_274	20.190
GEEEPQLQLEDGSIIPI[K]TSA	CIITE[K]AAR	2, 3, 4	130.10, 39.54, 141.38	Not Covered	Not Covered
GIIYIDQID[K]VARK	[K]TGAR	1, 2, 3, 4, 5, 6	185.25, 75.90, 190.36, 78.53, 169.89, 78.53	186_359	16.946
DVSGEGVQQALL[K]ILEGTVASVPPQGGR	INSNET[K]EDEVELSK	2	57.96	211_102	21.625
MLELDDVELEFEPTALIEIA[K]EAIER	[K]TGAR	2, 4	70.92, 47.82	353_359	9.764
VIGFGTDNA[K]LKDDETYLSR	[K]SENPSITR	2, 4, 6	143.03, 146.44, 116.49	272_190	28.887
VIGFGTDNA[K]LKDDETYLSR	MGE[K]VIGFGTDNAK	1, 2, 5	98.30, 121.40, 90.11	272_262	9.044
SNICLIGPTGSG[K]TLLAQTLAR	[K]SENPSITR	4	88.63	123_190	23.515
VIGFGTDNAKL[K]DDETYLSR	CIITE[K]AAR	2, 3, 4, 6	98.94, 208.66, 132.31, 119.22	274_394	25.499
GIIYIDQID[K]VAR	L[K]DDETYLSR	3, 4	57.46, 92.99	186_274	25.595

VIGFGTDNA[K]LKDDETYLSR	LIQSADYDVE[K]AEK	2, 6	31.75, 78.66	272_173	40.232
VIGFGTDNA[K]LKDDETYLSR	L[K]DDETYLSR	2, 3, 5, 6	43.63, 84.17, 60.96, 127.26	272_274	6.189
ALAVAVYNHY[K]R	[K]SENPSITR	5, 6	50.13, 50.13	94_190	36.338
MLELDDVELEFEPTALIEIA[K]EAIER	L[K]DDETYLSR	1, 2, 3, 4, 5	205.63, 45.31, 119.53, 48.50, 171.70	353_274	22.751
AARGEEEPQLQLEDGSIIPI[K]TSA	[K]TGAR	2	71.50	Not Covered	Not Covered
VVPEDLL[K]FGLIPEFIGR	[K]TGAR	5	32.12	290_359	23.893
RINSNET[K]EDEVELSK	NALVKQY[K]R	6	123.07	102_331	13.099
AE[K]GIIYIDQIDK	[K]TGAR	1, 4, 6	133.98, 40.95, 47.30	Not Covered	Not Covered
GEEEPQLQLDGSIIPI[K]TSA	L[K]DDETYLSR	1, 6	41.50, 148.91	Not Covered	Not Covered

Table S8 Shortest average Cα distances between cross-linking lysines based on molecular dynamics simulations of two ClpP2/ClpX trimers. The subunit position is given according to Figure I-17: **c**enter, **l**eft, **r**ight. σ is the standard deviation.

ClpP2 subunit[a]	ClpX subunit[a]	Residue 1	Residue 2	Mean [Å]	σ[b] [Å]
c	c	157	190	34.95	6.16
c	r	145	359	45.67	7.97
c	r	26	272	29.81	6.65
c	c	26	186	25.11	1.78
c	c	26	190	20.83	2.54
c	r	26	262	27.86	5.42
c	c	26	359	28.26	1.66
c	c	181	274	27.76	11.44
c	l	193	272	18.44	4.65
c	l	193	274	15.95	4.37
c	l	193	262	19.9	3.97

Table S9 DAVID analysis of significantly enriched proteins upon hClpP co-IP from *in situ* cross-linked HepG2. Functional annotation, *p*-value, and number of genes concerning cross-link/co-IP against hClpP from HepG2 (4-fold enrichment upon anti-hClpP immunoprecipitation compared to isotype control ($n=4$) and with a -log$_{10}$ (*p*-value) of more than 1.3).

HepG2 co-IP		
Functional Annotation	***p*-value**	**Number of Genes**
Nucleotide binding	3.5 E-10	36
Carbon metabolism	7.3 E-11	15
Protein biosynthesis	1.6 E-10	13
Oxidation-reduction process	8.2 E-6	17
tRNA aminoacylation for protein translation	1.6 E-8	8
Pyruvate metabolism	8.5 E-6	7
Arginine and proline metabolism	3.2 E-5	7
One carbon pool by folate	9.3 E-5	5
Electron carrier activity	5.5 E-5	7
Protein folding	4.0 E-4	8

Table S10 Cross-linking of PLPBP with MS-cleavable DSSO linker. Detected cross-links for PLPBP monomer and dimer (FDR<0.01). Cross-links of DSSO between corresponding peptides A and B are shown together with XlinkX scores and position in the sequence. For each cross-link we indicate, in which sample it was detected (excess of DSSO, fragmentation strategy, oligomeric state of PLPBP). The lysine in parentheses is the site of the detected cross-link.

1 1:50_EThcD (Monomer)   **2** 1:50_MS3 (Monomer)   **3** 1:100_EThcD (Monomer)

4 1:100_MS3 (Monomer)   **5** 1:50_EThcD (Dimer)   **6** 1:50_MS3 (Dimer)

7 1:100_EThcD (Dimer)   **8** 1:100_MS3 (Dimer)

Peptide A	Peptide B	Position	Sample No.	XlinkX Score
LAD[K]VNSSWQR	[K]GSPER	125_133	1, 2, 4, 5, 6, 7, 8	55.06, 156.12, 75.67, 90.13, 198.45, 75.67, 187.45
ASNP[K]ILSLCPEIK	[K]GSPER	81_133	1, 2, 5, 6, 7, 8	81.88, 244.15, 100.09, 229.49, 65.29, 185.67
[K]PTPDK	CAADV[K]APLEVAQEH	255_266	2, 5, 6, 7, 8	100.11, 44.12, 148.22, 50.16, 151.74
T[K]PADMVIEAYGHGQR	[K]GSPER	49_133	7	35.78
T[K]PADMVIEAYGHGQR	ASNP[K]ILSLCPEIK	49_81	5, 6, 8	118.74, 216.22, 69.81
LVAVS[K]TKPADMVIEAYGHGQR	[K]GSPER	47_133	5, 6, 8	58.74, 124.37, 58.74
[K]PTPDK	ASNP[K]ILSLCPEIK	255_81	5, 6	65.29, 200.97
LVAVS[K]TKPADMVIEAYGHGQR	ASNP[K]ILSLCPEIK	47_81	5	41.48
T[K]PADMVIEAYGHGQR	LAD[K]VNSSWQR	49_125	5	39.15
[K]PTPDK	T[K]PADMVIEAYGHGQR	255_49	5, 6	36.47, 124.37
QNVN[K]LMAVPNLFMLETVDSVK	[K]GSPER	104_133	5	33.41

VI – CURRICULUM VITAE

Anja Stefanie Fux, born 28th January 1993 in Schrobenhausen, Germany

Doctoral studies:

Since 11/2016

Supervisor Prof. Dr. Stephan A. Sieber, Chair of Organic Chemistry II, Technische Universität München

Projects:

(I) *ClpXP protease networks studied by chemical cross-linking combined with mass spectrometry*

(II) *A proteomic survey on human pyridoxal 5'-phosphate-binding proteins*

Academic education:

10/2014-09/2016

Master of Science in Biochemistry (Grade: 1.0)

Technische Universität München

<u>Master's Thesis</u>

Supervisor Prof. Dr. Stephan A. Sieber, Chair of Organic Chemistry II, Technische Universität München

Projects

(I) *Crosslinking combined with mass spectrometry for structural investigations of the ATP-dependent protease ClpXP*

(II) *Small molecules targeting ClpXP for whole proteome analysis by mass spectrometry in human cancer cell lines and Staphylococcus aureus*

10/2011-09/2014

Bachelor of Science in Biochemistry (Grade: 1.7)

Technische Universität München

Bachelor's Thesis

Supervisor Prof. Dr. Matthias Rief, Chair of Molecular Biophysics, Technische Universität München

Project

Detection of open-close transitions of adenylate kinase with fluorescence dyes

Scholar education:

10/2003-09/2011

Gymnasium Schrobenhausen, Schrobenhausen

Higher education entrance qualification (Abitur, Grade: 1.4)

10/1999-09/2003

Grundschule Mühlried, Schrobenhausen

Basic education

Conferences:

06/2019   Gordon Research Conference (Bioorganic Chemistry)

Poster presentation, Andover, NH, US

05/2018   SFB1035-Conference

Poster presentation, Venice, Italy

Seminars/ Workshops:

06/2019	Gordon Research Seminar (Chemical tools for understanding biological systems)
	Andover, NH, US
05/2018	TUM Graduate School (Develop your entrepreneurial thinking)
	Reitenhaslach, Germany
05/2018	Origin Basic and Intermediate Training
	Venice, Italy
07/2017	MaxQuant Summer School
	Berlin, Germany
02/2017	TUM Graduate School (New developments in the research of the pharma industry)
	Munich, Germany
01/2017	European Bioinformatics Winter School
	Semmering, Austria
11/2016	Mass-Spec Proteomics on Campus 2016
	Martinsried, Germany

Industrial Internship:

11/2015-01/2016

Research and Development (GMP-compliant)

Labor LS (Bad Bocklet, Germany)