universität wien

DISSERTATION | DOCTORAL THESIS

Titel | Title

Exploring the impact of posttranslational modifications on function and structure of Heat shock protein 90 via protein semisynthesis

verfasst von | submitted by

Oliver Gajsek BSc MSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 419
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Chemie
Betreut von | Supervisor:	Univ.-Prof. Dr. Christian Friedrich Wilhelm Becker

Acknowledgments

I thank the Austrian Academy of Sciences for financial support.

I gratefully acknowledge the scientific support received by Prof. Christian Becker and Asst. Prof. Anne Conibear.

Furthermore, I thank the members of the Institute of Biological Chemistry for providing an exciting working environment and interesting conversations inside and outside of the laboratory.

I also want to thank my family, friends, and significant other for their moral support and always having an open ear.

Abstract

As molecular chaperones, heat shock proteins are essential guardians of the cell by ensuring proteostasis under basal and stress conditions. Heat shock protein 90 (Hsp90) belongs to one of the most abundant representatives of the Hsp family, with several hundred known client proteins. Many of these client proteins are part of signal transduction cascades, making the chaperoning activity of Hsp90 play an essential role in many cellular processes.

Hsp90 is highly regulated by posttranslational modifications (PTMs), which influence all significant steps of the chaperoning process. These modifications include phosphorylation, acetylation, and methylation, which occur in all major domains: The amino-terminal domain NTD, middle domain (MD), and C-terminal domain (CTD). Also, rarer PTMs such as SUMOylation, *S*-nitrosylation, and ubiquitination have been identified. However, the impact of specific modifications and their crosstalk are still primarily unknown or hinted at, giving birth to the term chaperone code.

This thesis aims to apply protein semi-synthesis technology to gain access to site-specifically modified homogenous Hsp90 variants, which can provide novel insight into the chaperone code. Three semi-synthesis routes have been explored to achieve this goal, which aim to generate human Hsp90α amino-terminal domain (NTD), carboxy-terminal domain (CTD), and full-length protein variants. A native chemical ligation strategy generated NTD variants bearing two different phosphorylations. The expressed protein selenoester ligation (EPSL) methodology successfully provided access to four phosphorylated and unmodified Hsp90 CTD variants. The CTD variants were refolded, and the impact of the PTMs on the secondary structure was characterized by CD spectroscopy as well as their thermal denaturation and chaperone activity assessed, which indicates a minor effect of double phosphorylation on the thermal stability of Hsp90 CTD. Furthermore, the EPSL strategy for generating semisynthetic CTD was extended by applying two additional C-to-N sequential ligation steps, allowing the generation of (unmodified) full-length hHsp90α.

This strategy will allow further study of the chaperone code by potentially generating a variety of full-length Hsp90 variants bearing site-selective modifications in the CTD and/or the MD.

Zusammenfassung

Molekulare Chaperone, wie die Hitzeschockproteine (Hsp), spielen eine entscheidende Rolle als Wächter der Zelle, indem sie sowohl unter normalen als auch unter Stressbedingungen die Proteostase aufrechterhalten. Innerhalb dieser Proteinfamilie gehört das Hitzeschockprotein 90 (Hsp90) zu den am häufigsten vorkommenden Vertretern und ist essenziell für die Faltung, Stabilisierung und Aktivierung hunderter Substrate – Proteine, die auf die Chaperonaktivität von Hsp90 angewiesen sind.

Ein zentraler Regulationsmechanismus von Hsp90 sind posttranslationale Modifikationen (PTMs), die nachweislich alle wichtigen Schritte des Chaperonprozesses beeinflussen. Zu den häufigsten PTMs zählen Phosphorylierung, Acetylierung und Methylierung, die in allen drei Hauptdomänen von Hsp90 – der Amino-terminalen Domäne (NTD), der mittleren Domäne (MD) und der Carboxy-terminalen Domäne (CTD) – auftreten. Auch seltenere PTMs wie SUMOylierung, S-Nitrosylierung oder Ubiquitinierung wurden bei Hsp90 identifiziert. Trotz ihrer Bedeutung sind die spezifischen Auswirkungen dieser Modifikationen sowie ihre Wechselwirkungen bislang nur unzureichend verstanden, was zur Prägung des Begriffs „Chaperon-Code“ führte.

Im Rahmen dieser Arbeit wurde mithilfe der Proteinsemisynthese ein Zugang zu ortsspezifisch modifizierten, homogenen Hsp90-Varianten geschaffen, um neue Einblicke in diesen Chaperon-Code zu gewinnen. Hierzu wurden drei verschiedene Semisynthesestrategien untersucht. Durch native chemische Ligation wurde versucht, phosphorylierte Varianten der N-terminalen Domäne (NTD) von Hsp90 herzustellen. Gleichzeitig gelang es mithilfe der EPSL-Methode (expressed protein selenoester ligation), vier verschiedene Varianten der C-terminalen Domäne (CTD) zu generieren, sowohl phosphorylierte als auch unmodifizierte. Die Auswirkungen dieser PTMs auf die Sekundärstruktur der CTD wurden durch CD-Spektroskopie untersucht. Außerdem wurde deren Einfluss auf die thermische Denaturierung und Chaperonaktivität der CTD analysiert. Die EPSL-Strategie wurde durch zwei zusätzliche sequenzielle Ligationen erweitert, wodurch erstmals die Synthese von unmodifiziertem humanen Hsp90α in voller Länge gelang.

Diese Methode eröffnet neue Möglichkeiten zur Erforschung des Chaperon-Codes, da sie die gezielte Erzeugung einer Vielzahl von Hsp90-Varianten mit ortsspezifischen Modifikationen in der CTD und/oder MD ermöglicht.

Abbreviations

aa	Amino acid
ACN	Acetonitrile
ATP	Adenosintriphosphate
Boc	*Tert*-Butyloxycarbonyl
CD	Cirular dichroism
CTD	Carboxy-Terminal Domain
DCM	Dichloromethane
DIC	*N,N-Diisopropylcarbodiimide*
DIPEA	*N,N-Diisopropylethylamine*
DMF	*N,N*-Dimethylformamide
DSL	Diselenide-selenoester ligation
DTT	1,4-dithiothreitol
E. coli	*Escherichia coli*
EDT	1,2-ethanedithiol
EPL	Expressed protein ligation
EPSL	Expressed protein selenoester ligation
ESI	Electrospray Ionization
Fmoc	9-Fluorenylmethoxycarbonyl
Gdn.HCl	Guanidine hydrochloride
GR	Glucocorticoid receptor
HSE	Heat shock element
HSF	Heat shock transcriptional factor
HSP	Heat shock protein
HSR	Heat shock response

IB	Inclusion body
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
NCL	Native chemical ligation
NTD	Amino-Terminal Domain
MALDI	Matrix-assisted laser desorption/ionization
mesna	2-mercaptoethanesulfonic acid
MD	Middle Domain
MPAA	4-mercaptophenylacetic acid
MS	Mass spectrometry
MTG	Methylthioglycolat
RP-HPLC	Reversed-phase high-performance liquid chromatography
rt	Room temperature
S. cerevisiae	*Saccharomyces cerevisiae*
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SrtA	Sortase A

Table of Contents

1 Introduction

1.1 The Heat shock response and Heat shock transcription factors

All living organisms feature an optimum growth temperature, and evolutionary adaptations have enabled life to exist, from freezing water temperatures to above 100°C (at elevated pressure). De*via*tion of temperature only moderately above the optimal conditions of a specific organism induces stress reactions in cells that can be potentially lethal without any countermeasures.[1] The overlying cellular mechanism that deals with sudden temperature increases and detrimental effects on cellular structures and functions is termed heat shock response (HSR).[2] Later, the term widened as it became clear that HSR is triggered by unfolded proteins resulting from various biotic and abiotic stressors, including chemicals, toxins, UV radiation, and oxidizing agents. It became clear that this is an ancient, universal, and highly evolutionary conserved protection mechanism in all three domains of life: Archaea, Bacteria, and Eukarya.[3] The term heat shock originates from the first experiments that examined the effect of suddenly increased temperatures on the fruit fly *Drosophila melanogaster* on a chromosomal level. In the early 1960s, F. Rosetta observed the induction of a specific set of puffs in the larval salivary gland polytene chromosomes after heat shock exposure, indicating the emergence of locally enhanced transcription.[4]

The HSR requires specific transcription factors, termed heat shock transcription factors (HSFs). HSFs are a conserved family of proteins that feature DNA binding activity in their active form. They were initially described as proteins specifically recognizing the heat shock element (HSE) DNA binding site and activating the genes encoding proteins termed Heat shock proteins (Hsps). A significant class of Hsps are molecular chaperones, which prevent protein aggregation and facilitate folding.[5] The latest research revealed the complexity of the HSF's structural and functional characteristics, revealing their modulation via post-translational modifications, protein-protein interactions, and DNA-binding selectivity. Although there are six known HSF isoforms in the human genome: HSF1, HSF2, HSF4, HSF5, HSFX, and HSFY with HSF1, research has focused on HSF1 and HSF2 due to their direct role in the expression of stress-responsive genes and their involvement in a variety of diseases.[6] The detailed activation mechanism of HSFs represents an ongoing challenge but has been directly linked to its oligomerization state. In eukaryotes, the HSF family member Hsf1 predominantly exists in an inactive state under standard cellular growth conditions. In this state, the Hsf1 monomer forms a complex with other proteins from the Hsp90 chaperone machinery and the Hsp70/40 complex in the cytoplasm. Regulation of Hsf1 also involves a variety of post-translational modifications and other protein interactions; the classical chaperone titration model delivers a simplified explanation of the activation of Hsf1 through its release from the Hsp complex, which leads to the trimerization and transport to the nucleus (**Figure 1**).[7] If internal or

external stress factors lead to the increase of unfolded proteins that are recognized by molecular chaperones such as Hsp90 or Hsp70, the former unengaged chaperones release Hsf1 upon binging to the unfolded proteins, allowing Hsf1 to trimerize and travel to the nucleus in its active state. If basal cellular conditions are recovered, excess free chaperones lead to the binding of Hsf1 and downregulation of the heat shock response.

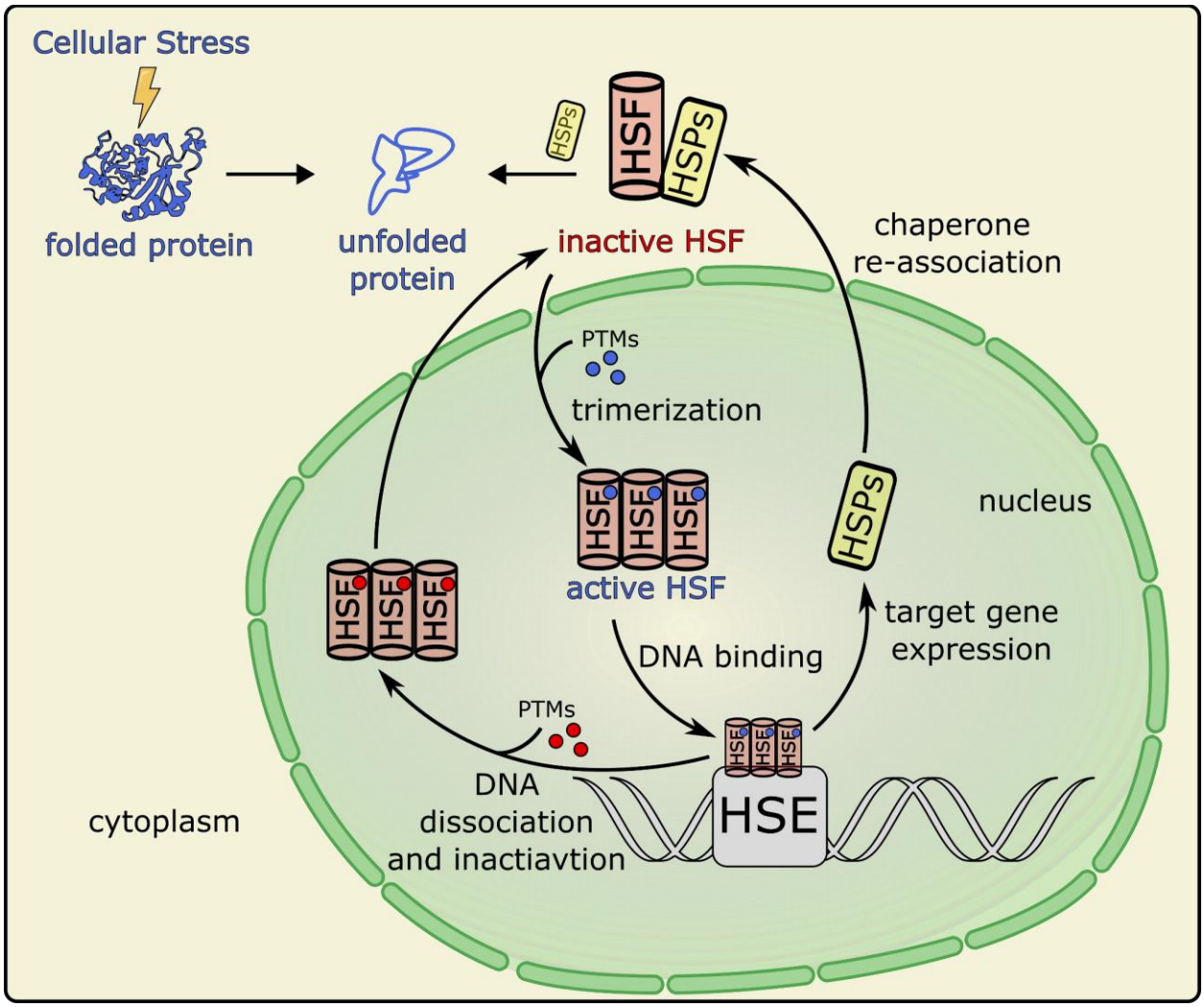

Figure 1 Chaperone titration cycle showcasing the activation of the heat shock transcription factor. Cellular stress conditions cause the increase of unfolded proteins and the binding of chaperones to these, which in response release HSF1 monomer retained in the cytoplasm. Upon release, HSF1 trimerizes and relocates to the nucleus, where it undergoes various posttranslational modifications (PTMs) upon binding to heat shock element (HSE) DNA regions, therefore causing transcription activation of HSP target genes. After release from the DNA region, a process regulated by PTMs, the increased Hsps causes inhibition of HSF activation *via* a negative feedback mechanism.

Genome-wide studies of the expression of Heat shock proteins revealed that about 50-200 genes are significantly induced in various prokaryotic and eukaryotic model organisms.[2] These Hsps have previously been grouped into seven different classes based on their primary function, of which the class with the highest expression level belongs to the molecular chaperones.[8] Other classes include members of the proteolytic system, including ubiquitin, which clears no longer needed, misfolded and aggregated proteins from the cell, and DNA and RNA repairing enzymes that reverse covalent nuclei acid modifications like stress-induced methylation.[9] Another class of Hsps includes metabolic enzymes that reorganize and stabilize cellular energy supply during stress and a class of heat shock-induced transcriptional factors. The final two classes of Hsps comprise cytoskeletal proteins providing cellular structure and transport, detoxifying, and membrane modulating enzymes.[10] The following section will provide an overview of the molecular chaperone family of Hsps and their general mode of action before focusing on one of the members of this protein superfamily, which is the main target of this thesis: Hsp90.

1.2 Heat shock Proteins and the chaperone mechanism

1.2.1 The Hsp family of proteins and the underlying chaperoning function

The term protein homeostasis (sometimes also referred to as proteostasis) is used to describe the maintenance of the integrity of the cellular protein network. To fulfill the plethora of operations required to sustain cellular function, proteins are intrinsically dynamic by nature. However, this comes with the cost of being on the edge of unfolding, even under physiological conditions.[11] Cellular compartments represent a crowded space with proteins constantly coming in proximity, which favors unspecific interactions and aggregation and needs to handle ever-changing environmental conditions with internal and external stress factors. Therefore, it is unsurprising that molecular chaperones make up to 1–2% of total protein in unstressed cells, while their concentration can be increased up to 4–6% under stress to ensure protein homeostasis.[12] The term chaperone originates from the French word “chaperon” and is used for a matron who accompanies young people to ensure proper behaviors and prevent "undesired" social interactions. It was first used by Ron Laskey in 1978 to describe the ability of the protein nucleoplasm to avoid the aggregation of folded histone proteins during the assembly of nucleosomes.[13]

With the major class of Hsps being molecular chaperones, the term Heat shock proteins is frequently directly used for the five ubiquitous conserved protein families that are classified based on their molecular weight into large Hsps (e.g., Hsp100), Hsp90s, Hsp70s, Hsp60s and small heat shock proteins (sHsps). The functioning of these molecular chaperones is in many ways different from classical enzymes as all the chaperones act promiscuously with a wide range of unfolded proteins and generally have to be present at stoichiometric ratios to effectively decrease the concentration of non-native proteins.[14] Another classification is usually based on how the chaperoning is achieved since so-called “holdases”, including the sHsps, bind to their substrates to prevent aggregation but are not directly involved in the refolding. Furthermore, holdases work in an ATP-independent manner and are usually only stress-induced, while “foldases” such as Hsp90 and Hsp70 refold their clients as ATP-dependent and exist in constitutively active and stressed-induced isoforms.[15,16] Chaperones generally do not provide structural information for refolding their substrates, often referred to as “client” proteins, but act more as preventers of unwanted intermolecular interactions. They achieve this through controlled binding and release of their substrates, which is driven by a change of affinity coupled to the binding and hydrolysis of ATP. The exposure of hydrophobic amino acids due to partially or entirely incorrect folding of globular proteins is a recognizable feature for chaperones, and binding usually occurs at hydrophobic patches, specific peptide sequences, or other structural elements of the non-native substrate.

1.2.2 Small Hsps

The group of sHsps represents the most heterogeneous family of heat shock proteins defined by their molecular weights between 12 and 43 kDa. The ATP-independent proteins are found in the nucleus and cytosol, constitutively expressed, and rapidly upregulated upon stress exposure to counter unfolding pressure.[17] Most sHsps function as holdases forming large hollow sphere-shaped oligomers that bind to a wide variety of partially folded target proteins to prevent their aggregation and prepare them for folding through the Hsp70/Hsp90 machinery. Structurally, they share a conserved α-crystallin domain comprising a β-sandwich structure, flanked by N- and C-terminal regions of a different length and sequence.[18]

1.2.3 Hsp60

A unique feature of the Hsp60 chaperone family is its ring-shaped subunit, which leads to the assembly of barrel-like protein complexes by stacking onto each other. This feature gave this class of Hsps its nomenclature, and they are generally named chaperonins (Cpn). Chaperonins encapsulate their substrate during the refolding and belong to the ATP-dependent Hsps, with their most prominent members being GroEL in bacteria and Hsp60 in eukaryotes, respectively. Their specific role as mitochondrial chaperonins includes transporting and refolding proteins from the cytoplasm into the mitochondrial matrix, making it essential to ensure mitochondrial homeostasis.[19]

1.2.4 Hsp70

Hsp70 and its prokaryotic equivalent, DnaK, represent one of the most highly conserved chaperones in the Hsp family. These ubiquitous proteins are expressed in different cellular locations, including the cytosol, nucleus, mitochondria, and the endoplasmatic reticulum (ER). They are involved in the *de novo* folding of proteins under physiological conditions and prevent the unfolding of proteins under stress. There are 17 genes in the human cell encoding for Hsp70, and they can be distinguished between constitutively expressed isoforms termed HSC70 (heat shock cognate) and a stress-induced form (HSP70s). The Hsp70 isoforms share a conserved structure consisting of two major domains, an N-terminal nucleotide-binding domain (NTD) and a C-terminal substrate binding domain (SBD) connected *via* a flexible linker. The SBD domain contains a C-terminal tail with an EEVD tetratricopeptide binding (TPR) domain motif recognized by various co-chaperones. J-domain containing co-chaperones that bind incorrectly folded proteins and deliver them to Hsp70 whilst interacting with its ATPase domain are important for regulating this ATP-dependent

protein. They stimulate ATP hydrolysis, leading to the ADP-bound state of Hsp70, which exhibits the highest affinity for its substrates. Nucleotide-exchange factors accelerate the exchange of ADP to ATP, enabling the release of the substrate before another cycle of J-domain protein and client binding.[20]

1.2.5 Large Hsps

Large Hsps include the eukaryotic Hsp110 and Grp170, considered conserved homologs of the Hsp70 family. While Hsp110 resides universally distributed in the cytosol and nucleus, Grp170 is specifically located in the ER. The bacterial Hsp100 subfamily belongs to the AAA+ superfamily of ATPases. It is further subdivided into two classes: Class one exhibits two AAA+ nucleotide binding modules and comprises the bacterial proteins ClpA, ClB, ClC, and ClpE, as well as the plant homolog ClpD, while class two contains only one nucleotide-binding domain such as in ClpX, ClpY and heat-shock-like U (HslU). They function as ATP-dependent foldases that form ring-structured hexamers and are believed to drag misfolded target proteins into their pore, where folding is assisted. Furthermore, large Hsps can form disaggregation machinery that can dissolve and reactivate aggregated proteins, as observed for the Hsp110/Hsp70/Hsp40 machinery in humans that can decompose α-synuclein amyloid fibrils efficiently.[21,22]

1.2.6 Hsp90

Hsp90 represents another highly abundant member of the Hsp protein superfamily of foldases. This chaperone supports the folding and activation of a plethora of client proteins besides adding in protein degradation and assembling multiprotein complexes. As the stress-inducible isoform, Hsp90α is the target protein of this thesis, and the following chapter will give a more detailed introduction to the knowledge of Hsp90 structure, function, and regulation.[23]

1.3 Heat shock protein 90 (Hsp90) and the chaperone code

1.3.1 Structure and function of Hsp90

Heat shock protein 90 (Hsp90) represents one of the most abundant members of the Hsp family in eukaryotic cells under basal and stress conditions.[24] The protein has several hundred known client proteins that are directly or indirectly involved in all major cellular processes, which include apoptosis, cell cycle control, and cell *via*bility.[25] Many of these clients are part of the signal transduction cascade and therefore include many kinases (e.g., Src tyrosine kinases, EGF receptor or Cdc2), steroid hormone receptors (e.g., glucocorticoid receptor) but also transcriptional factors (e.g., p53, c-Myc) and ubiquitin ligases (e.g., Hakai).[26] Besides chaperoning its important clients, Hsp90 tasks also include aiding in protein degradation *via* directing substrates towards the ubiquitin-proteasome system and mediating protein complex formation by stabilizing otherwise labile subunits like in RNA polymerase II assembly. Another factor that makes Hsp90 a vital research target is the observation that many oncogenic protein variants depend on Hsp90.[27] Chaperoning through Hsp90 enables the stabilization of mutations, a crucial factor in malignant progression and adaptivity of cancers. Due to this role, Hsp90 has been directly linked to leukemia-, prostate- and breast cancer.[28,29] Furthermore, it has been shown that it is involved in the initiation and progression of a variety of other disorders, including infectious diseases (caused by Herpesviridae, Paramyxoviridae) and neurodegenerative diseases (Alzheimer's), which sparked the search for Hsp90 inhibitors in the last two decades.[30] Since then, more than a dozen Hsp90 inhibitors have been developed and clinically evaluated, with only recently receiving approval for the first treatment against gastrointestinal stromal cancer.[31,32]

Two major cytosolic isoforms of Hsp90 have been identified in mammals: the stress-inducible form Hsp90α and the constitutively expressed Hsp90β. Both isoforms share ~85% sequence identity and are considered highly homologous, although recently, it has been shown that they have distinctive functions.[33,34,35] Both variants form a homodimer in their active state, with each monomer comprising three highly conserved domains with distinct features (**Figure 2**). The N-terminal domain (NTD), with the 'lid' region, contains the critical nucleotide pocket where binding and hydrolysis of ATP occurs during the chaperoning cycle. The NTD is connected *via* a flexible and highly charged linker region to the middle domain (MD), the protein's largest domain and a primary binding site for clients and co-chaperones. Finally, the C-terminal domain (CTD) enables dimerization of Hsp90 and carries a distinct MEEVD tetratricopeptide repeat (TPR) domain binding motif at its C-terminus.[36] This motif is recognized by a variety of TPR domain containing co-chaperones, and each Hsp90 monomer can only interact with one TPR cochaperone at a time.

Competition of this specific set of chaperoning helpers is another complex level of regulation of chaperone activity.[37]

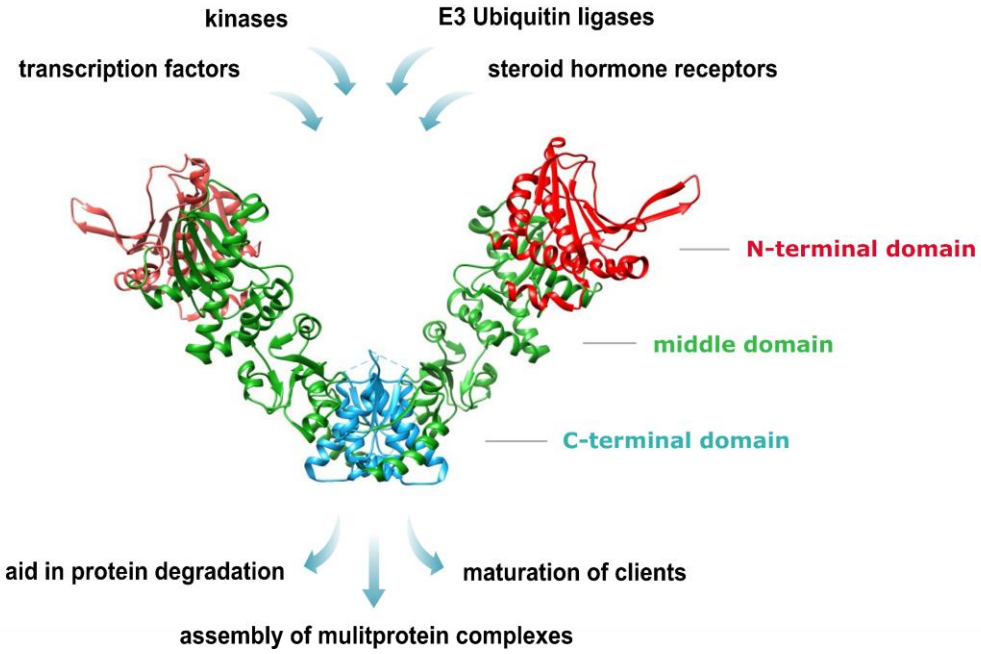

Figure 2 Crystal structure of *E. coli* Hsp90 analog HTPG (IPD:2IOQ) with colored N-terminal domain (red), middle domain (green), and C-terminal domain (blue).[38] Main client classes include signaling pathway proteins like kinases, steroid hormone receptors, transcriptional factors, and ubiquitin ligases. Besides chaperoning, Hsp90 is directly involved in protein degradation by directing substrates towards the ubiquitin-proteasome system and the assembly of larger multiprotein complexes.

During chaperoning, Hsp90 undergoes a complex conformational change that is referred to as the chaperone cycle (**Figure 3**). In its apo state, the Hsp90 homodimer exhibits an open V-shaped conformation where, after the binding of ATP and a client, the nucleotide lid region closes, which causes large structural rearrangements leading to the dimerization of the NTD (closed state 1**, Figure 3**). Subsequently, a side swap of the NTD segments occurs, followed by compression of the MD domains and the subsequent ATP hydrolysis (closed state 2, **Figure 3**). The release of ADP, phosphate, and the correctly folded client protein, the simultaneous dissociation of the NTDs, and the regain of the initial open conformation completes the chaperone cycle. During the different conformational stages of this cycle, various co-chaperones have been identified that dynamically associate and dissociate with/from the Hsp90-client complex to support and regulate the chaperoning activity.[39] For example, the co-chaperones Cdc37 and HOP are known to be essential for client protein recruiting, which they enable by blocking the ATPase activity of Hsp90, thereby facilitating binding to clients in the open conformation state.[40] In contrast, the co-chaperone Aha1

accelerates the ATPase activity of Hsp90 by promoting the rate-limiting conformational changes during the chaperone cycle.[41] This shows the complex interplay of co-chaperones with Hsp90 as a regulatory mechanism of the chaperone activity.

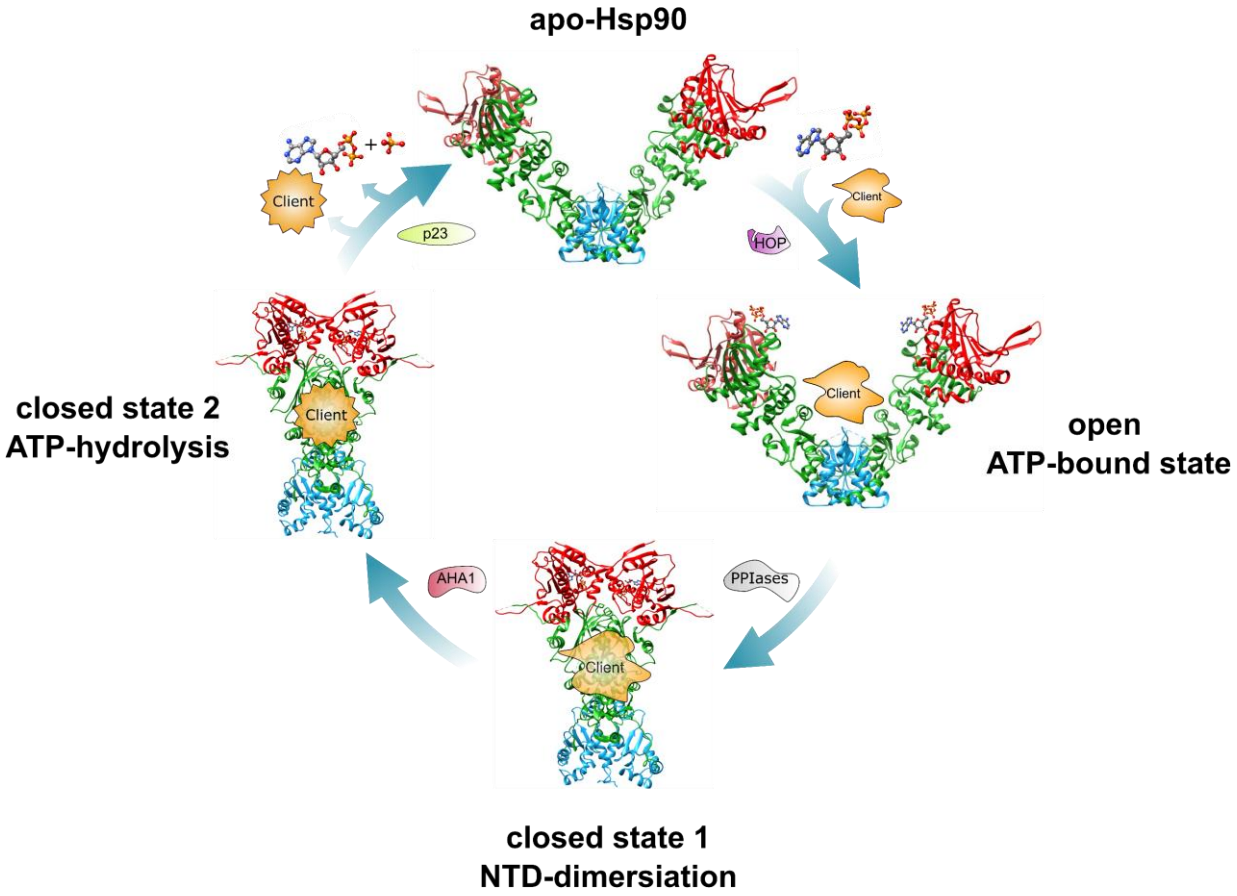

Figure 3 Schematic presentation of the chaperone cycle. Crystal structure of *E. coli* Hsp90 analog htpG (IPD:2IOQ) representing the apo- conformation while closed conformations are crystal structures of *S. cerevisiae* Hsp90 analog Hsp82 (IDP: 2CG9).[38,42] During its chaperoning activity, the protein cycles through different open and closed conformational states, driven by the hydrolysis of ATP and binding of co-chaperones and client proteins. The proteins HOP, PPIase, AHA1, and p23 are important co-chaperone representatives that dynamically bind and dissociate during the different cycle steps as a regulatory aid for the chaperone cycle.

1.3.2 Posttranslational modification of Hsp90: The chaperone code

Another level of Hsp90 regulation, besides transcriptional regulators and co-chaperones, are distinct site-selective chemical modifications that occur after the translation process of Hsp90, hence termed posttranslational modifications (PTMs). Most PTMs found on Hsp90 include phosphorylation and acetylation that occur on all domains of the protein, but also methylation, SUMOylation, S-nitrosylation, and ubiquitination sites have been identified (**Figure 4**).[43]

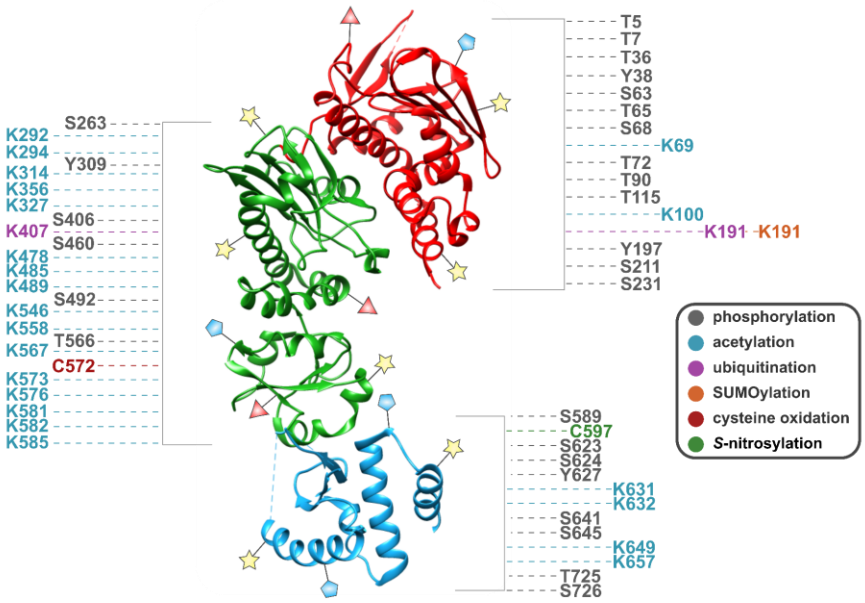

Figure 4 Known sites of posttranslational modifications of the human Hsp90α isoform. The major PTM sites of Hsp90 are the phosphorylation of Thr, Ser, and Tyr residues and the acetylation of Lys residues, occurring in all three major domains. Other rarer modifications, including cysteine oxidation, nitrosylation of serine, or Lysin SUMOylation/ubiquitination, have also been identified.[43]

Recent investigations have shown that PTMs are essential modulators of Hsp90 that modulate fundamental aspects of Hsp90's chaperoning function, including the kinetics of the chaperone cycle, the maturation of client proteins, and the binding of co-chaperones and nucleotides. Besides the general effects of hyper-phosphorylation and acetylation of Hsp90, different site-specific PTMs have been investigated over the past three decades, indicating the crucial impact of single modification sites on the overall functioning of the protein.[44,45]

For example, Muller *et al.* showed that phosphomimetic mutations close to the -EEVD TPR recognition motif at the C-terminus of Hsp70 and Hsp90 influence the binding to the co-chaperones CHIP and HOP and increase cell proliferation.[46] In contrast, another study from Backe *et al.* showed that a phosphomimetic mutation of a conserved serine residue in the NTD (yeast Hsp82 S25/ human Hsp90α S39) inhibits ATPase activity of Hsp90, thereby causing the dissociation of

Hsp90 from the protein kinase complex Atg1/Ulk1, which leads to initiation of autophagy.[47] Another example by Mollapour *et al.* indicates the effect of phosphorylation of a conserved threonine residue (T22) in the NTD of yeast Hsp90 N-domain.[48] The position of this modified residue is located in an α-helical region, known to participate in hydrophobic interactions with the middle domain during ATP binding. The authors showed that phosphomimic mutants displayed altered ATPase activity and chaperone function *in vitro* and *in vivo* and that overexpression of the co-chaperone Aha1 compensated for these functional defects. Additionally, they examined the effect of the mutation on the drug sensitivity of Hsp90, indicating that the phosphorylation of this site increased the sensitivity of Hsp90 towards the NTD binding inhibitors geldanamycin, radicicol, and two synthetic inhibitors in a cell-based assay, an important factor when considering Hsp90 inhibitor development for oncological indications.[49] Besides phosphorylation, other PTMs have been studied, including acetylation of single sites. For example, Scroggins *et al.* studied the impact of K294 acetylation in Hsp90α at the junction of the charged linker region and the middle domain using point mutations that were supposed to mimic acetylated (Gln and Ala) or unacetylated (Arg) lysine. Their discovery suggested that a single acetylation weakened the interaction of Hsp90 with the client kinase ErbB2 and androgen receptor (AR) and directly abrogated the protein's ability to chaperone Chk1 kinase in an *in vitro* assay.[50] In a final example, Lees-Miller *et al.* identified that the dsDNA-activated protein kinase (DNA-PK) specifically phosphorylated two threonine residues (Thr5 and Thr7) only present at the amino terminus of the α isoform of human Hsp90 in vitro using 32P labeling and tryptic digestion followed by autoradiography of the fragments.[51] Following that, Quanz *et al.* used stabilized double-stranded DNA molecules or irradiation to induce a damage response *in vivo* to investigate novel DNA-PK targets. They thereby confirmed the phosphorylation of Thr7 Hsp90α in MRC-5 cells. Using immunoblotting, they revealed that this phosphorylation event is directly caused by DNA damage, which resulted in the accumulation of Hsp90 at sites of DNA double-strand breaks where it formed repair foci, thereby linking this modification to Hsp90 chaperoning in response to DNA damage.[52,53]

As shown above, most previous studies on the impact of single PTMs on Hsp90's function were based on either utilizing point mutations, which mimic PTMs in terms of their charge and steric properties, or by performing *in vivo* or *in vitro* enzymatic modifications of Hsp90. However, PTM mimicking mutations only allow for an estimation of the actual alteration in reactivity since differences between the phosphomimetics and the phosphorylated residue caused by de*via*tions in the charge states, steric properties, and Ramachandran distributions can alter the protein's structure and function significantly.[54]

Enzymatic modifications often result in inseparable heterogeneous mixtures due to the limited activity and specificity of the modifying enzyme.[55] The complexity of this level of modulation and the specific obstacles of studying the particular PTM impact gave rise to the term “chaperone

code", which makes the accessibility of site-selective homogenously modified Hsp90 variants highly desirable but also makes this an enormous challenge.[56]

In summary, the pivotal task of Hsp90 as a chaperone and guardian of the proteome with a plethora of clients and being involved in many major cellular processes and various diseases makes this protein a desirable target. However, the complex mechanisms of Hsp90s function, including the underlying regulation through PTMs, make it an incredibly challenging target to comprehend fully. The following chapters of this thesis introduce the strategies applied to generate selectively modified Hsp90 variants to understand Hsp90 function and regulation through PTMs further.

1.4 Protein semi-synthesis as a tool to unravel the chaperone code

1.4.1 Background

As the chaperone code exemplifies, posttranslational modifications serve as crucial molecular regulatory mechanisms. With more than 400 known types of reversible or irreversible PTMs, they have an enormous impact on the structure and functionality of proteins. Therefore, they play a key role in a plethora of biological processes, such as signal transduction, gene activation, DNA repair, and cell cycle control.[57,58] In recent years, different methodologies, including computational methods and classical immunoprecipitation in combination with mass spectrometry, have expanded the possibilities for detecting PTMs.[59] However, systematic studies of individual PTMs and their cross-talk still represent a challenge and requires access to proteins with defined and homogenous PTMs.[60] The following section provides a concise overview of the synthetic and recombinant technology methods that enable access to site-selective modified peptides and proteins. It also introduces the strategies applied in this thesis for starting to unravel the chaperone code.

1.4.2 Chemoenzymatic ligation strategies for protein synthesis

The introduction of stepwise solid phase peptide synthesis (SPPS) by Merrifield in 1963 marked a revolution in synthetic organic chemistry, especially in peptide synthesis.[61] This method is based on immobilizing amino acids via its C-terminal carboxyl group onto a solid support, followed by iterative deprotection and coupling cycles for elongation of the peptide chain and, therefore, building up the desired peptide. In the following decades, the development of a variety of solid supports with different linkers and coupling reagents, as well as the introduction of orthogonal protection group strategies, including the base labile 9-fluorenylmethyloxycarbonyl (Fmoc) group as α-amine protection, made SPPS an extremely popular and widespread method for generating native peptide sequences.[62] Furthermore, it enabled the introduction of non-native modifications, such as installing probes or site-specific post-translational modifications. However, its linear nature usually causes insufficient quality and yield when peptides exceed the length of 40~50 amino acids due to aggregation and steric crowding. Recently, exceptional cases of smaller proteins with up to ~300 amino acids were successfully synthesized using a flow chemistry setup.[63,64]

The introduction of native chemical ligation (NCL) as the currently most used chemoselective ligation method that allows the connection of synthetic peptide segments bearing natural or

unnatural modifications to build up larger polypeptides and proteins marked another big leap in the protein chemistry field (**Figure 5A**).[65] The reaction is based on an N-terminal peptide bearing a C-terminal α-thioester that chemoselectively reacts with another peptide featuring an N-terminal cysteine residue in an initial transthioesterification step. This step is facilitated through the nucleophilic attack of the cysteine thiol moiety on the thioester, followed by an S → N acyl shift that generates the stable native amide bond, making the reaction traceless and irreversible. The fact that this reaction can be performed using unprotected peptide segments and under mild, aqueous conditions made it highly versatile for generating a variety of small proteins. It sparked various methodological approaches for generating the necessary C-terminal thioester moiety on peptide segments.[66] However, one drawback of NCL was always the requirement for a cysteine residue at the ligation site, which can be challenging as Cys represents one of the least common proteinogenic amino acids (~2% abundance).[67] This limitation was overcome by the introduction of post-ligation desulfurization chemistry of the sulfhydryl side chain to the corresponding alanine, significantly increasing the number of possible ligation junctions due to the relatively high abundance of alanine (~9% in the human proteome).[68] This procedure and the introduction of β-mercapto amino acid derivatives made this an extremely widespread and versatile methodology. However, if native cysteines are present in the target sequence, these might require temporary protecting groups to be retained during desulfurization reactions, which represent a sometimes tedious or not feasible obstacle to overcome.[69]

Over the past 15 years, other ligation methods emerged based on different proteinogenic functional groups or the introduction of new reactive ligation handles, such as DSL, KAHA, and Serine Threonine ligation (STL) (**Figure 5B**).[70–72] STL is based on a C-terminal salicylaldehyde (SAL) ester and a peptide with an N-terminal serine or threonine residue. The mechanism involves an imine capture followed by the generation of an *N,O*-benzylidene acetal at the conjugation site through a 5-endo-trig ring–chain tautomerization and an *O*-to-*N* [1,5] acyl transfer. Acidolysis of the *N,O*-benzylidene acetal releases the natural peptide bond at the conjugation of the peptide segments, which can be applied without side chain protection groups, making this a viable alternative for NCL if the desired ligation site contains a Ser/Thr residue and the generation of the C-terminal SAL ester is compatible with the desired peptide sequence.

Another thiol-independent ligation technique is the α-Ketoacid-hydroxylamine amide-forming (KAHA) ligation.[72] This reaction involves the reaction of N-terminal hydroxylamines on one peptide segment with another peptide carrying a C-terminal α-ketoacid and leads to the formation of an amide bond. This reaction can also be applied to unprotected peptide segments without catalysts (**Figure 5C**). Two different types of KAHA ligation have been classified based on the use of O-unsubstituted hydroxylamine in type I and *O*-benzoylhydroxylamine in the type II reaction.[73] Although differences in the specific mechanisms have been proposed, the general

mechanism is envisioned to start with the nucleophilic attack by the hydroxylamine upon the carbonyl group, followed by an elimination of H_2O or benzoic acid, depending on the type of hydroxylamine used. Subsequent decarboxylation yields a carbanion, which, after a hydroxyl transfer, generates a hydroxyimine that tautomerizes to the final peptide bond product.[74] With the most widely used hydroxylamine being the 5-membered cyclic hydroxylamine (*S*)-5-oxaproline, the reaction leaves homoserine at the ligation junction, making it not fully traceless.[75] However, recently developed hydroxyl amines, e.g., 4,4-difluoro-5-oxaproline, generate an aspartic acid at the ligation junction, making this a traceless ligation technique, although synthetic access to the α-ketoacid and hydroxylamine peptide derivatives have to be achieved.[76]

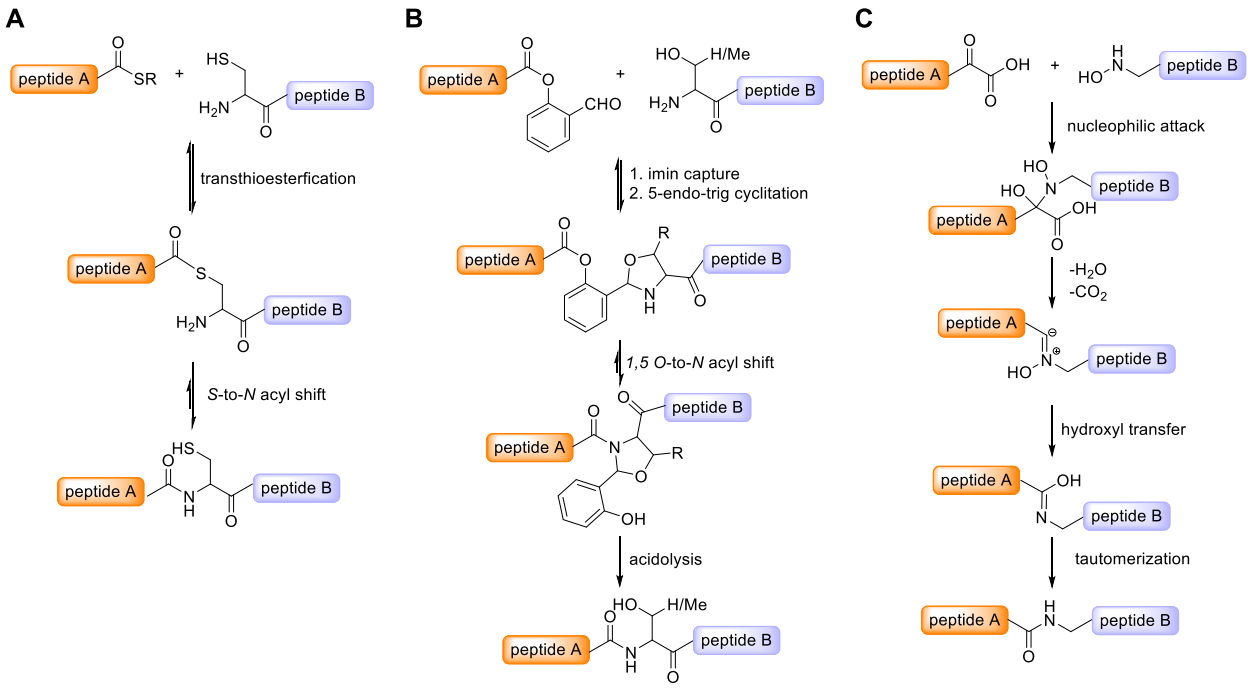

Figure 5 Different chemoselective ligation techniques for the synthesis of peptides and proteins. **A:** The Native chemical ligation (NCL) is based on the reaction of a C-terminal thioester as acyl donor segment with the thiol moiety of an N-terminal Cys side chain. **B:** During the Serine/Threonine Ligation (STL), a C-terminal salicylaldehyde ester reacts with peptides containing N-terminal serine or threonine residues to form a peptide bond after the hydrolysis of an *N,O*-benzylidene acetal intermediated. **C:** The α-Ketoacid-hydroxylamine amide (KAHA) ligation utilizes the reactivity of N-terminal hydroxylamine-modified peptides with C-terminal α-ketoacids peptide segments.

A more recent methodology and closely related to NCL is the Diselenide-Selenoester Ligation (DSL), in which the thioester and cysteine ligation handles are replaced with their respective α-selenoester and selenocysteine counterparts (**Figure 6**).[71]

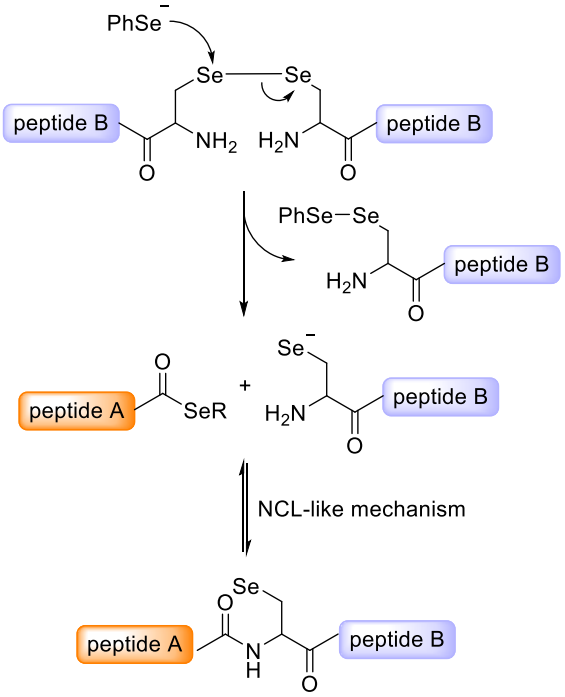

Figure 6 Mechanism of Diselenide Selenoester Ligation (DSL).[71]

Due to the low reduction potential of selenocysteine (-381 mV compared to -180 mV for Cys), N-terminal Sec peptides are usually only observed as the corresponding diselenide dimers in solution and require an initial reduction to release free selenol to perform the ligation.[77] DSL exhibits an almost equal reaction mechanism to that of the NCL. However, the beneficial nucleophilic properties and increased polarizability of Se are also indicated by the lower pK_a of Sec (5.2-5.6) compared to Cys (8.2), as well as the increased electrophilic nature of the corresponding selenoester improve reaction rates significantly.[78] This feature enables successful ligations of peptide segments at lower µM concentrations and difficult ligation junctions, overcoming limitations of NCL, which proceeds best at mM concentrations and where C-terminal, β-branched amino acids bearing the thioester moiety are ideally avoided.[79] Furthermore, the possibility of performing this ligation at highly diluted concentrations is essential when poorly water-soluble hydrophobic peptide segments or lipidated proteins are the target. Finally, another significant benefit of this methodology that is also essential for the objective of this thesis is the possibility of selective deselenization of the selenocysteine residue, which is left at the ligation junction after the reaction.[80] As described above, desulfurization chemistry allowed the expansion of the NCL methodology by using cysteines as surrogates for native alanine residues at the ligation site; however, if native Cys is present, these require protection group chemistry to be retained when performing desulfurization.[81] Deselenization can be performed under mild conditions using only TCEP and a hydrogen atom donor (DTT) without the need for a radical initiator, as used for the desulfurization of Cys, which makes the transformation chemoselective and does not touch unprotected Cys residues. This unique

advantage is especially important when Cys protection group chemistry cannot be applied, e.g., in a recombinant segment as a ligation partner.[82]

Finally, another alternative approach to expand the scope of NCL beyond reactions at cysteine is the utilization of ligation auxiliaries in the form of universal, generically applicable aminothiol groups attached to the peptide N-terminus. An example of this would be the use of 2-mercapto-2-(pyridin-2-yl)ethyl derivatives, which can be introduced by reductive alkylation and can be removed after the auxiliary mediated ligation in a traceless manner using a radical-based reaction or under mildly basic conditions. This methodology provides an acceptable reaction rate even for junctions containing proline or a β-branched amino acid and was applied to synthesize eight MUC5AC tandem repeats containing 120 aa long peptide *via* the ligation of two 60mers.[83,84]

1.4.4 Protein semi-synthesis through Expressed Protein Ligation and Expressed Protein Selenoester Ligation

Taken together, a variety of chemoselective ligation strategies enable synthetic access to tailored, site-selectively modified proteins, but although sequential use of chemoselective ligation reactions allows for the assembly of multiple synthetic peptide segments into larger proteins, the efficient generation of proteins beyond 200 aa still requires the use of recombinant technology through biological expression in prokaryotic and eukaryotic systems. However, introducing site-selective modifications through the cellular machinery through either enzymatic processing or genetic code expansion remains highly challenging and frequently results in inseparable heterogeneous mixtures of the target protein or very low amounts of target protein, respectively.[85,86] Protein semi-synthesis can overcome these challenges by utilizing chemical ligation methods to connect small synthetic peptide segments bearing natural or unnatural modifications to larger recombinantly generated polypeptides.[87] Expression of protein segments with a free N-terminal thiol moiety through the protease-based release of Cys, e.g., by cleavage of a tag, such as a TEV recognition sequence or SUMO, represents a well-established methodology that allows the use of NCL in combination with a synthetically generated peptide carrying a C-terminal α-thioesters. This enabled the incorporation of N-terminal modifications into larger proteins and naturally sparked the desire to expand this method to introduce C-terminal modifications into larger proteins. Further research in this direction fulfilled this desire by extending the versatility of protein semi-synthesis through Expressed Protein Ligation (EPL). This methodology involves the reaction of a protein segment with a C-terminal α-thioester (produced by expression) with a synthetic peptide or protein equipped with an N-terminal Cys. When using SPPS to generate the C-terminal peptide, this strategy also allows the introduction of PTMs close to the C-terminus of the target protein.[88]

This recombinant C-terminal thioester is generated by exploiting a naturally occurring autocatalytic process termed protein splicing. Analogous to RNA splicing, protein splicing is achieved through inteins, which are defined as protein domains that autocatalytically excise themselves from a protein sequence and fuse the two lateral protein domains termed the N-extein and C-extein together *via* a peptide bond.[89,90] The mechanism of protein splicing starts with an N-to-S acyl shift at a Cys residue at the N-terminus of the intein sequence, generating a thioester intermediate. This is followed by a transesterification step through an acyl transfer initiated by the side chain of either a Cys or Ser residue at the N-terminal position of the C-extein. This branched intermediate is released through the cleavage by an Asn residue at the intein domain's C-terminus, resulting in the intein's excision. Finally, the released (thio-)ester bond between the N-extein and C-extein protein domains re-equilibrates to the native backbone, forming the stable amide bond (**Figure 7**).[91]

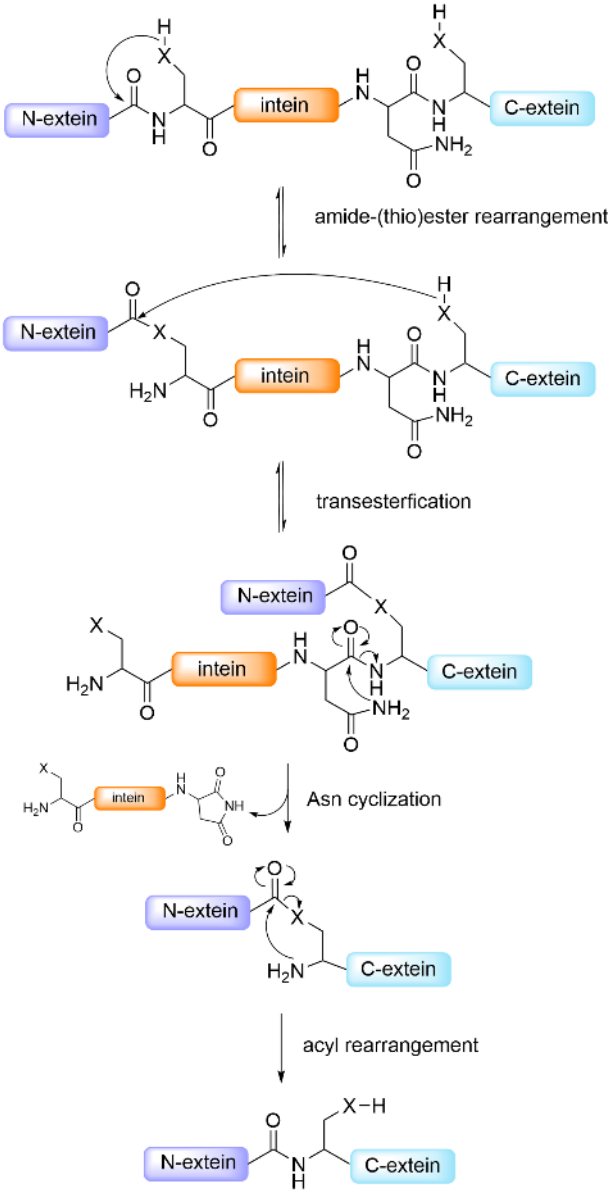

Figure 7 Mechanism of protein splicing: An internal protein segment (intein) autocatalytically excises itself from two flanking protein regions termed N-extein and C-extein. X represents either a Cys or Ser/Thr residue. The first step involves a N to S/O shift through a (thio)ester-arrangement at the N-terminus of the intein followed by a transesterification by the attack of either a Cys/Ser or Thr residue as the first residue of the C-extein. A subsequent Asn cyclization at the intein segment's C-terminus leads to the peptide bond cleavage between intein and C-extein. Finally, an acyl rearrangement forms an N- and C-extein peptide bond.

In EPL, specifically engineered inteins are fused to the C-terminus of the target protein, allowing the cleaving of the thioester intermediate by the addition of an excess of (thiol) nucleophiles and thereby generating the desired C-terminal α-thioester for the ligation to the Cys peptide segment (**Figure 8A**).[92] This strategy usually also involves an affinity tag, e.g., a chitin-binding domain (CBD) or a His6-tag, attached to the C-terminus of the intein that enables immobilization on an affinity matrix and, therefore, affinity purification and the separation of intein-tag from the generated thioester segment.[93] Similarly to the development of the selenium-based DSL strategy as a continuation of NCL, the desire to be able to apply DSL for ligating recombinant selenoester segments to selenocysteine-containing peptides recently gave rise to the Expressed Protein Selenoester Ligation (EPSL) method. EPSL combines the advantages of DSL with EPL concerning ligations at low concentrations and selective deselenization. In EPSL, analogously to EPL, an intein fusion construct with the target protein segment is generated. However, the thiolysis step is replaced by a hydrazine-based intein cleavage to generate the respective C-terminal acyl hydrazide (**Figure 8B**). The hydrazide can then be converted into the respective α-selenoester at acidic pH *via* an acyl pyrazole intermediate through Knorr Pyrazole chemistry.[94]

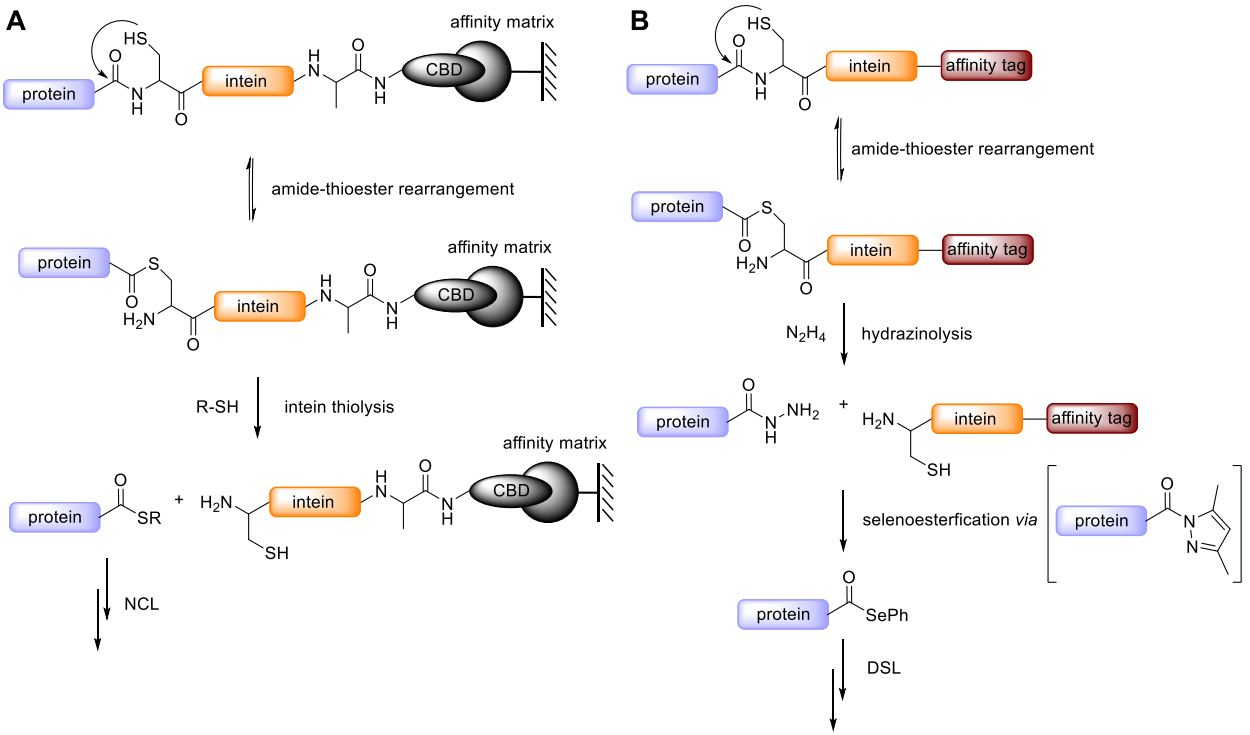

Figure 8 Principal mechanisms of Expressed Protein Ligation (EPL) and Expressed Protein selenoester Ligation (EPSL). **A:** In EPL, the recombinant protein segment is expressed as an intein fusion construct, cleaved through thiolysis, giving the corresponding thioester derivative. Adding an affinity tag, e.g., a chitin-binding domain, to the C-terminus of the intein furthermore allows for direct purification of the desired thioester through immobilized affinity chromatography. The thioester segment can then be used to ligate the synthetic peptide segment of choice. **B:** Analogously to EPL, Expressed Protein Selenoester Ligation (EPSL) uses intein fusion technology to generate a C-terminal peptide hydrazide through hydrazinolyses. The peptide hydrazide can be converted to the selenoester through the generation of an acyl pyrazole intermediate, e.g., *via* the use of acetylacetone, which readily exchanges with an aryl selenol to give the respective aryl selenoester.

The importance of protein synthesis and semi-synthesis is highlighted by hundreds of reported modified or unmodified proteins accessed through this methodology. In many cases, these strategies allowed access to proteins carrying site-selective PTMs where the effects of the introduced modification were studied, which allowed new insights into the often highly complex impact of PTMs. For example, phosphorylation and acetylation of tau and N- and O-glycosylation of α-synuclein, both proteins that play a central role in neurodegeneration, were extensively studied using protein semi-synthesis.[93]

Furthermore, a variety of PTMs affecting the nucleosome core proteins histone H4, H3, and histone 2A, as well as the nucleosome binding protein HMGN1, have been examined by protein semi-synthesis.[95] However, larger proteins with more conserved structures are scarcer as protein semi-synthesis targets, presumably due to challenges in generating the protein segments and the

underlying ligation chemistry and downstream characterization, which usually requires the successful refolding of the protein.

In this instance, if soluble protein segments can be produced, another alternative for the assembly of protein segments without the need to apply denaturing chemistry would be the use of enzyme-based ligations, including the use of specifically engineered transpeptidases, e.g., Sortase A (SrtA) or the asparaginyl endopeptidase Butilase 1.[99,100] The enzyme SrtA catalyzes the formation of a new peptide bond between a C-terminal LPXTG (X can be any amino acid) motif and an N-terminal oligo-glycine (GGG motif). This approach recently gave novel insights into Hsp90's conformational state dynamics during the ATPase cycle by generating site-selectively fluorescently labeled Hsp90 variants.[101] Another successful example of this methodology is displayed by the segmental isotope labeling of Hsp90 NTD before SrtA-mediated ligation towards Hsp90 MD, which allowed the elucidation of cochaperone-GR-Hsp90 interactions using NMR spectroscopy.[102] However, the required recognition motif is not fully cleaved during the enzymatic ligation through SrtA. Therefore, this cannot be considered a traceless methodology if the required sequence is not naturally present at the target protein. In comparison, butelase 1 is structurally a cysteine protease that belongs to the asparaginyl endoprotease (AEP) protein family but does not display any protease activity but instead Asn/Asp-specific ligase activity. The protein recognizes the C-terminal Asn/Asp -His-Val tripeptide motif and enables the formation of an intermolecular or intramolecular Asn/Asp - X1X2 aa peptide bond (X1 = any amino acid except proline, X_2 = I/L/V/C). Another benefit of Butelase 1-based peptide/protein ligation is its high catalytic efficiency since it is the fastest known ligase and typically requires 100- to 1,000-fold less enzyme than a SrtA ligation-based ligation.[103] Taken together, the short recognition sequence (Asx–His–Val) and broad tolerance for the first N-terminal residue of the nucleophilic donor peptide make this a versatile tool for protein ligation showcased by a variety of examples where it was applied for N-terminal labeling, site-specific peptide and protein ligation, macrocyclization and bioconjugation of dendrimers.[104–106] However, one drawback besides hydrolytic activity and reversibility was the lack of sufficient access to Butelase 1 as it required extraction from plants, and only recently, access *via* recombinant expression was successfully established.[107]

2 Objectives

This thesis aims to establish semi-synthetic routes for accessing site-selective post-translationally modified human Hsp90α variants. Semi-synthetic routes were designed to gain access to N-terminally phosphorylated variants of Hsp90 NTD (**Figure 9 A**) and C-terminally phosphorylated Hsp90 CTD (**Figure 8 B**). Furthermore, a ligation strategy for accessing full-length Hsp90 for which PTMs in the CTD and/or MD could be introduced was explored (
Figure 9 **C**). Access to Hsp90 NTD (232 aa) and CTD domains (161 aa) should be provided *via* a single ligation step based on an NCL for the NTD and an EPSL strategy for CTD variants, respectively. Full-length Hsp90 (732 aa) should be accessed *via* two sequential C- to N-terminal ligation reactions utilizing a combination of NCL and EPSL, which enables the introduction of PTMs into the MD, specifically acetylation of lysin residue K548 and phosphorylation of the CTD while maintaining all native cysteine residues.

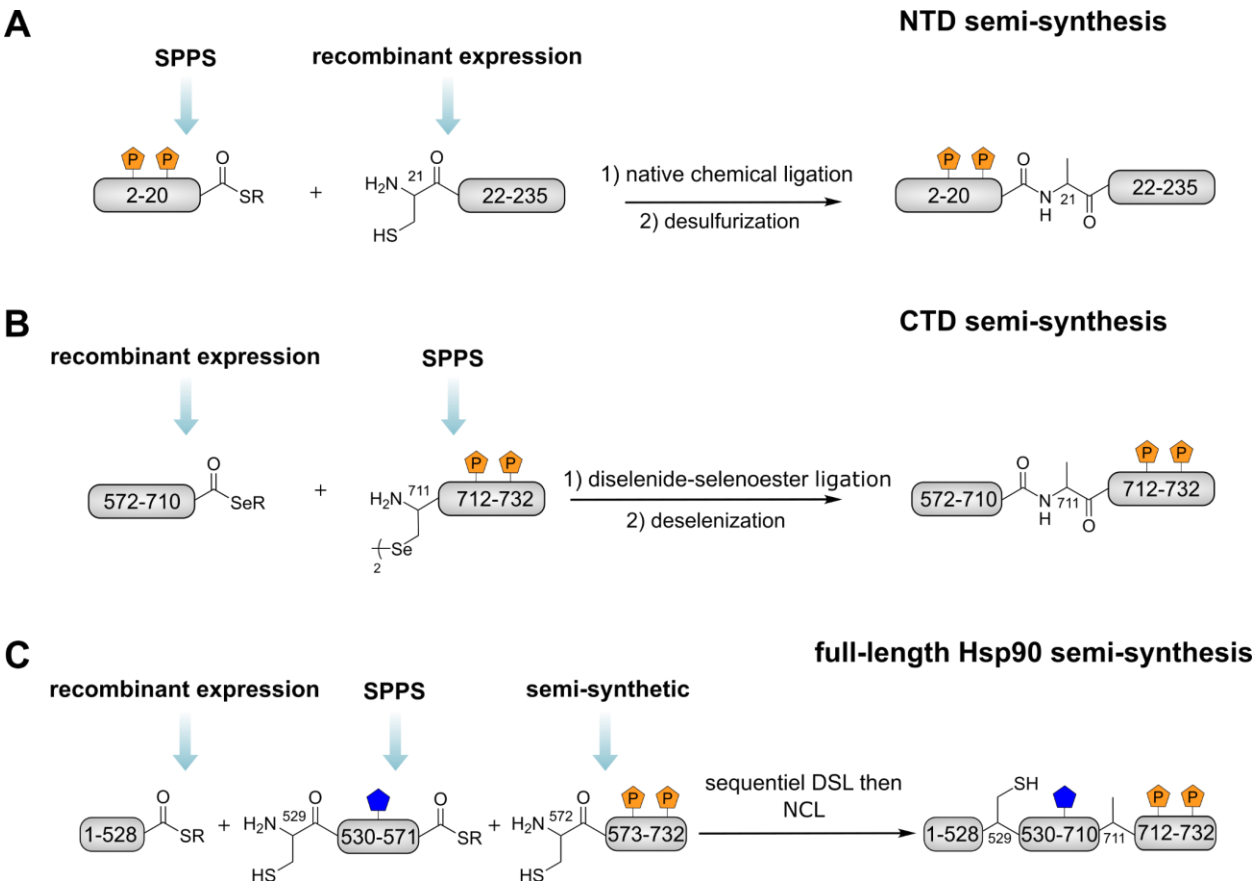

Figure 9 Schematic overview of envisioned semi-synthetic strategies for the generation of modified human Hsp90alpha NTD (A), CTD (B), and full-length (C) using EPL, EPSL, and a combination of both methods. Modifications in the form of phosphorylation are represented as orange pentagons labeled with a P inside, while acetylation is depicted as a blue pentagon.

Furthermore, upon successfully generating the desired variants, the impact of the introduced PTMs on specific structural and functional properties will be examined through various biophysical characterization techniques and functional assays. In particular, the effect on the protein's folding and secondary structure elements should be assessed using circular dichroism spectroscopy, and the overall structural stability should be assessed using thermal aggregation analysis. The modulation on functional aspects, including chaperoning capabilities as well as dimerization properties, should be evaluated using established model substrate aggregation assays comparing the modified variants with unmodified semi-synthetic NTD, CTD, or full-length Hsp90 as well as with the expressed unmodified counterparts as controls. Hereby, we should be able to gain unique and novel insights into the specific effect(s) of single PTMs. We should establish a basis for further exploration of the chaperone code by protein semi-synthesis on a quite complex protein.

CHAPTER ONE

Semi-synthesis of Hsp90α N-terminal Domain

1. Introduction

This chapter describes the semis-synthesis of N-terminally modified variants of human Hsp90α NTD. It builds on my master thesis project („Modulating the cell's guardian: The impact of post-translational modification on function and structure of Heat shock protein 90", University of Vienna). The interest modifications are two phosphorylations at N-terminal Thr residues (Thr5 and Thr7), which have been reported to be phosphorylated by DNA-dependent protein kinase (DNA-PK) in response to DNA damage. However, their specific impact on the structural and functional properties of the NTD remains to be elucidated.[42,51] Since there are no native cysteine residues in the NTD, Ala21 was chosen as a ligation site since a non-native cysteine residue can be converted into alanine after ligation. This ligation site allows the synthesis of a peptide of reasonable length (aa 2-20) comprising the desired phosphorylation sites (Thr 5 & 7). Since the initiating methionine (Met1) of the NTD has been reported to be cleaved *in vivo* by endogenous methionine aminopeptidases, this residue was omitted in our synthetic peptide.[108] The unmodified and phosphorylated peptides can be prepared as peptide hydrazides on hydrazine-modified 2-chlorotrityl resin and used as thioester surrogates.[109] The recombinant segment Hsp90 21-235 was designed with a C-terminal His-tag for affinity purification and an N-terminal Met-Cys sequence so that direct cleavage by endogenous *E. coli* Met-aminopeptidases (MAP) can release the N-terminal cysteine for NCL. Post-ligation desulfurization will restore the native Ala residue at position 21, allowing access to the desired protein variants (

Figure 10). The generated Hsp90 NTD variants are refolded, and the impact on secondary structural elements is determined *via* circular dichroism spectroscopy. Furthermore, functional modulations, including the ATP binding capabilities and affinities towards known Hsp90 NTD binding proteins, e.g., the co-chaperones p23 and Cdc37, through ATP binding and pulldown assays.

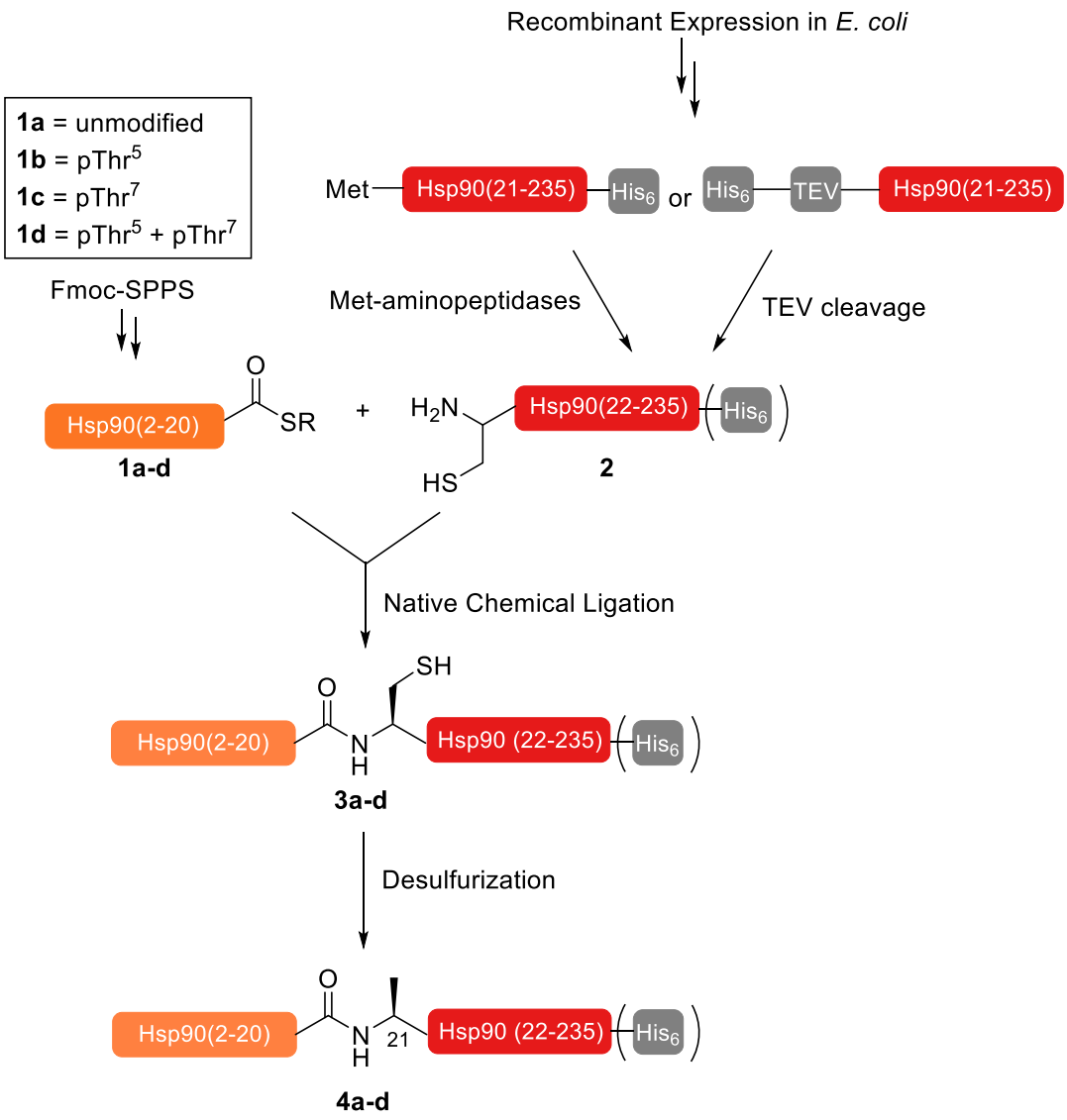

Figure 10 Schematic overview of the generation of site selectively modified Hsp90 NTD using a Native Chemical ligation-based synthesis strategy.

2. Materials and Methods

2.1 Solid phase peptide synthesis of Hsp90 2-20 peptide variants 1a-d

Synthetic peptide variants Hsp90 2-20 1a-d were previously synthesized as described in my master thesis, “Modulating the cell’s guardian: The impact of the posttranslational modification on function and structure of Heat shock protein 90” (**Figure A 4**). Briefly, modification of 2-Chlorotritly resin *via* hydrazine before coupling of the first amino acid gave access to the desired peptide hydrazides **1a-d** *via* standard Fmoc-based SPPS synthesis in a yield of ~15-25% according to the procedure of Liu and coworkers.[109] These peptides were either directly used for *in situ* thioester formation and ligation to recombinant segment Hsp90 21-235 (**2**) or separately converted to the corresponding mesna thioester, as described in the following section.

2.2 Thioester conversion of synthetic peptide hydrazide 1d

Double phosphorylated Hsp90 2-20 peptide hydrazide **1d** was converted to the respective C-terminal mesna thioester by dissolving the lyophilized peptide in 200 mM NaPi, 6M Gdn HCl pH 3 buffer at a concentration of 5 mM. The solution was precooled at -15°C before adding a 15 eq. $NaNO_2$ using a 0.5 M stock solution in H_2O and was stirred for another 15 min at -15°C before adding mesna (100 eq.) using a 1M mesna stock solution in 200 mM NaPi, 6M Gdn HCl, pH 6.8. The solution was warmed to room temperature, the pH adjusted to 6.8, and stirred for 30 min before diluting, filtering, and loading on a C18 semi-preparative column for purification. A linear gradient of 5-65% of solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) at a flow rate of 5 ml/min yielded the desired doubly phosphorylated mesna-thioester peptide **1d**.

2.3 Recombinant expression and purification of Hsp90 21-235

The gene for the N-terminal domain (NTD) construct of human Hsp90α 21-235 (UniProt entry: P07900; isoform 2) bearing a C-terminal His_6-tag was ordered in a pET-21a+ bacterial expression vector from BioCat GmbH (69120 Heidelberg, Germany) and transformed into *E. coli* competent cell strains BL21 (DE3) and Rosetta II *via* heat shock transformation. Overnight cultures with a final concentration of 100 µg mL^{-1} ampicillin (BL21) and 100 µg mL^{-1} ampicillin + 30 µg mL^{-1} chloramphenicol (Rosetta II) were prepared using LB media and incubated at 37°C at 170 rpm. Large-scale expressions involved inoculating 2 L of LB medium with 50-80 mL of overnight culture (starting OD_{600} ~ 0.2) before incubation at 37°C at 170 rpm until an OD_{600} ~ 0.7-0.8. Expression was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and cells incubated at 24°C for 4 hours before harvesting by centrifugation (5500 rpm, 14°C, 20 min). Cell pellets were resuspended in 30-50 ml TBS (150 mM NaCl, 50 mM Tris, pH 7.5) homogenized using an UltraTurrax (IKA GmbH, 79219 Staufen, Germany) and lysed *via* sonication at 4°C for 40 min (pulse: 15 ON 45 seconds OFF; ampl. 60%) or by disruption (Constant Systems Limited, Northants United Kingdom) at 1.93 kbar (three cycles). The lysate was centrifuged (22000 rpm, 45 min, 7°C), the supernatant removed, and the inclusion body pellets solubilized in 6M Gdn·HCl, TBS overnight. The protein was purified using immobilized Ni-NTA affinity chromatography before adding 0.1M methoxyamine hydrochloride at pH 4.2 to release the N-terminal Cys from formed thiazolidine derivatives. Final purification was performed *via* RP-HPLC on a preparative C4 column at 60°C using a linear gradient of 45-75% solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 40min at flow rate of 10 ml/min. Peak fractions were submitted for LC-MS analysis and pooled accordingly before lyophilization, yielding protein **2** at an average yield of 8 mg/L culture.

2.4 Ligation of synthetic peptide Hsp90 2-20 to recombinant Hsp90 21-235

2.4.1 *In-situ* thioester conversion and ligation of Hsp90 2-20 hydrazide 1a to recombinant Hsp90 21-235 (2)

For the *in-situ* thioester conversion and ligation of synthetic peptide hydrazide **1a** to recombinant protein **2**, the peptide hydrazide (1.2 eq.) was dissolved in 6M Gdn·HCl, 150 mM NaPi pH 3 buffer at a concentration of 5 mM and precooled at -15°C min before the addition of 20 eq. $NaNO_2$ using a 0.5M stock in ddH_2O. The solution was stirred for another 15 minutes at -15°C before combining with recombinant segment **2** dissolved in ligation buffer and incubated at rt for 1-2 h whilst stirring**.** Final ligation concentrations contained protein segment **2** in the range from 0.4-08 mM in a ligation buffer of 6M Gdn.HCl, 150 mM NaPi, and 100 mM MPAA at a pH of 2. The reaction was monitored *via* LC-MS and SDS-PAGE and, upon completion, quenched *via* 1:1 dilution with a solution of 6M Gdn.HCl, 20% B-Me (v/v) in ddH_2O for 15 min. Product purification was performed on a semi-preparative C4 column using a linear gradient of 5-75% solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 40 min at 60°C and a flow rate of 5 ml/min. Peak fractions were pooled and lyophilized, generating semi-synthetic Hsp90 2-235 (A21C) (**3a**) at 10-20% yield.

2.4.2 Post-ligation desulfurization of semi-synthetic Hsp90 2-235 (A21C)

For the desulfurization of RP-HPLC purified ligation product **3a**, the protein was dissolved in a freshly prepared and degassed desulfurization buffer containing 6M Gdn.HCl, 200 mM NaPi, 200 mM TCEP, 50 mM mesna, and 50 mM VA-044 at a final concentration of 0.4 mM. The pH was adjusted to 6.8, and the solution was incubated at 40°C (water bath) while stirring for 3-4 h. Reaction monitoring was performed *via* LC-MS, and purification was performed of an analytical C4 column using a linear gradient of 5-75% of solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 30 min at 60°C and a flow rate of 1 ml/min. Peak fractions were manually collected, lyophilized, and characterized, yielding sub-milligram amounts of unmodified semi-synthetic Hsp90 NTD.

2.4.3 Ligation and one pot desulfurization of synthetic double phosphorylated Hsp90 2-20-SR thioester (1d) to recombinant Hsp90 21-235 (2)

Small-scale ligations of **2** to peptide thioester **1d** (1.5 equiv.) were performed by dissolving the recombinant protein in the ligation buffer consisting of 6M Gdn.HCl, 200 mM NaPi, 50 mM TCEP, and 100 mM thiol catalyst (MPAA, MTG, mesna) to a concentration of 0.8 mM. Subsequently, the pH was adjusted to 7.2, and the solution was added to the peptide thioester while stirring at room temperature. The pH was measured again and adjusted to 7.2 if needed. The reaction was monitored every 30 minutes by LC-MS and SDS-PAGE. After the desulfurization, the pH was adjusted to 3.0, and ethyl acetate extraction was performed three times using MPAA. After adding an equivalent volume of sodium citrate buffer (desulfurization buffer) with a concentration of 500 mM TCEP to the reaction mixture, the pH was adjusted to 4.5 – 5.0. Then NaBEt4 was added to a final concentration of 100mM, according to the procedure of Li and coworkers, by preparing a 1 g/ml NaBEt4 stock solution in ddH_2O.[110] The reaction was monitored by LC-MS and, after completion, diluted with equal volumes of 6M Gdn.HCl in ddH_2O and β-mercaptoethanol (20%, v/v), frozen and stored at -80°C until further purification. Semipreparative purification of the ligation product was performed on a C4 Waters column using a gradient of 5-40% solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 5 min and 40-75% in 30 min at a flow rate of 5 mL/min at 60°C. Analysis was done *via* HPLC on a C4 Waters column at 60°C, using a linear gradient of 5-75% solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 30 min., detecting wavelengths 214 and 280 nm on a Dionex ultimate 3000 detector.

3. Results and Discussion

3.1 Synthesis of Hsp90 2-20 peptide hydrazides and conversion to mesna thioester

As mentioned in section 2.1, Hsp90 2-20 (**1a-d**) peptide hydrazides were successfully prepared using standard Fmoc-SPPS, yielding the segments in satisfactory purity and yield. Analytical data corresponding to the peptide hydrazides (**Figure A 1**) and the sequence 2-20 can be found in the appendix (section 5.1.3). Minor Glutamic deletion and addition side products were observed, presumably due to the significant Glu residues in the short peptide (seven Glu residues in a sequence of 19 aa total). Our initial ligation strategy involved the *in situ* conversion of the peptide hydrazides to the respective mesna thioester directly before ligating them to the recombinant part using the hydrazide activation and ligation protocol described by Liu *et al.*.[111] However, since we wanted to exclude the possibility of side product formation and provide ideal reaction conditions (especially for the post-ligation desulfurization step), we decided to convert and isolate the doubly phosphorylated peptide hydrazide **1d** the mesna thioester before the ligation step. Therefore, we could also use the isolated peptide thioester **1d** in test ligations to compare the *in situ* conversion and ligation of peptide hydrazides **1a-d**. Activation and thioester conversion of peptide hydrazide **1d** yielded the corresponding C-terminal mesna thioester in excellent purity and yield (~80% based on the peptide hydrazide), as depicted in **Figure 11**.

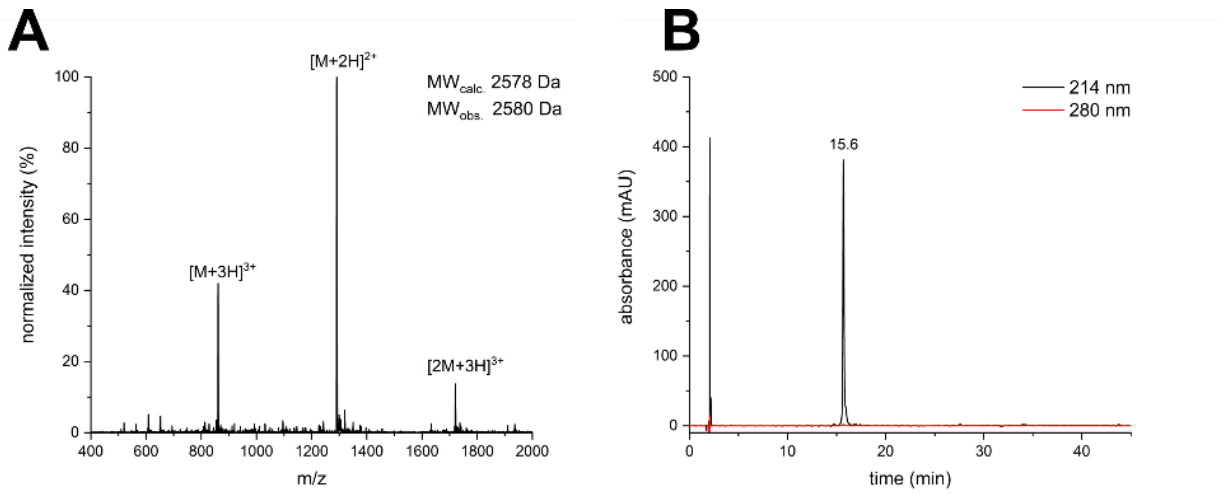

Figure 11 A: ESI-MS characterization of doubly phosphorylated Hsp90 2-20-SR peptide. **B:** RP-HPLC analysis of double phosphorylated Hsp90 2-20-SR peptide.

3.2 Expression and purification of Hsp90 21-235

Careful design of the recombinant protein segments plays a crucial role in the success of a semi-synthesis strategy, as it can influence the expression yield and final purity of the segment used for ligation. Expression in *E. coli* should ideally deliver pure starting material in at least lower mg/liter culture amounts. Therefore, semi-synthesis strategies usually take advantage of proteases engineered for stability and efficiency, e.g., the Tobacco Etch Virus (TEV) cysteine protease or SUMO protease, to release the Cys residue necessary for NCL in a traceless manner.[111,112] Another benefit would be the possibility of including an affinity purification tag, e.g., a His_6-tag, enabling initial separation of the cellular protein background.[114] If the affinity tag is not needed after this initial purification, proteases allow for the simultaneous cleavage of the tag and release the N-terminal Cys residue for the ligation step. Accordingly, our initial design of the recombinant segment Hsp90 21-235) (A21C) included an N-terminal His_6-tag followed by a TEV recognition site adjacent to the non-native Cys21 introduced as a ligation handle. Although decent expression yields of this construct were achieved, the protein was expressed exclusively in insoluble inclusion bodies (**Figure A 5**). This required high amounts of chaotropic salts (6M Gdn.HCl or 8M Urea) for initial solubilization and affinity purification. Unfortunately, all attempts to lower the amount of chaotropic salts < 4M Urea/ Gdn.HCl using dialysis or pulsatile dilution led to direct and/or delayed significant precipitation of the protein, which is described in detail in my master's thesis ("Modulating the cell's guardian: The impact of posttranslational modification on function and structure of Heat shock protein 90" Chapter 3.1.2.2). This led to significant problems when the crucial TEV cleavage had to be performed. Non-denaturing conditions need to be applied for proper protease activity; however, this limits the concentration of chaotropic salts (usually < 3M Urea) applicable, and no matching conditions were found that would keep the highly hydrophobic and aggregation-prone recombinant segment in solution while observing proteolytic cleavage. Interestingly, literature describing the expression and purification of Hsp90α N-terminal domain (NTD) typically includes the purification of the protein from the soluble fraction of the cell lysate, hinting towards the importance of the first 20 residues for proper folding and solubility properties of the highly conserved NTD.

Therefore, we envisioned redesigning the segment, leading us back to a straightforward but more unpredictable approach. We directly located Cys21 adjacent to the transcription initializing Met residues. This residue is usually not part of the functional protein sequence and is frequently cleaved by endogenous methionine aminopeptidase.[115] We added an His_6-tag to the C-terminus of the protein, allowing us to purify it using affinity methods while still giving the possibility to purify or immobilize the protein downstream when studying the impact of the introduced modifications *via* pulldown assays. Indeed, the first expression attempts of the Met-21-235-His6 did

show good overexpression of the protein, although still expressed as inclusion bodies (**Figure A 6**). Direct analysis of the pellet revealed initiator methionine cleavage to a large extent (**Figure 12** and **Figure A 7**). The initial mass de*via*tion from the calculated mass (+24 Da) is explained by N-terminal aldehyde-driven thiazolidine formation, which is frequently observed in the expression of proteins with N-terminal cysteine residues in *E. coli* and was quickly reversed by the addition of an excess of methoxyamine under acidic conditions. Further optimization *via* expression strain screening and expression conditions screening yielded optimal expression conditions where, based on MS signal intensity approximation, ~90% of the desired initiator meth-cleaved protein is produced (**Figure A 8**).

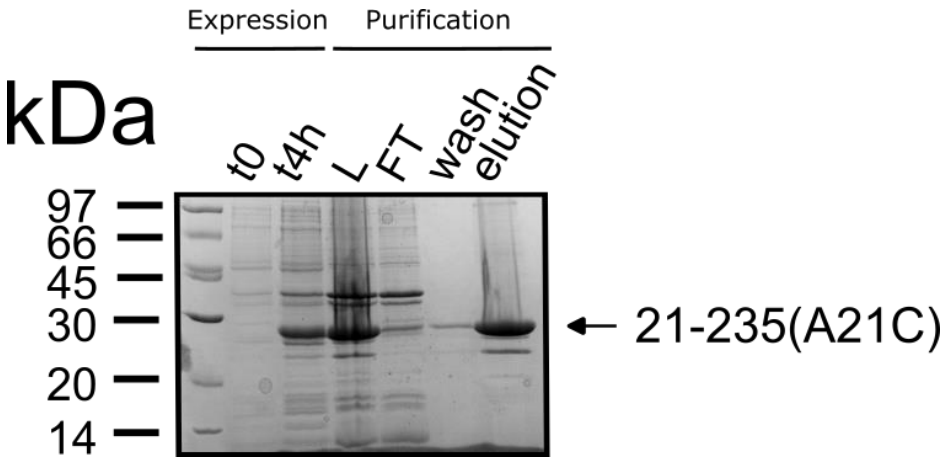

Figure 12 Monitoring of Expression (t0 and t4 h) and Purification of Hsp90 21-235(A21C). Expression was performed at 24°C for 4 hours and induced *via* adding 1mM IPTG. **L:** Solubilized inclusion body pellets of Hsp90 21-235. **FT:** Solubilized Hsp90 21-235 after loading onto Ni-NTA resin. **Wash:** Wash of Ni-NTA resin with 50 mM Imidazole, 6M Gdn.HCl, TBS buffer after loading of Hsp90 21-235. **Elution:** Eluted Hsp90 21-235 from Ni-NTA resin *via* incubation in 500 mM Imidazole 6M Gdn.HCl, TBS buffer.

Nevertheless, protein purification *via* RP-HPLC remained challenging, presumably due to strong interactions with the stationary phase. This resulted in late elution profiles, strong peak tailing, and poor recovery rates of around ~15%. Nevertheless, the overall yield of 8 mg/L culture and purity (**Figure 13**) of the desired segment 21-235 were satisfactory enough to continue with ligations to the synthetic segment 2-20.

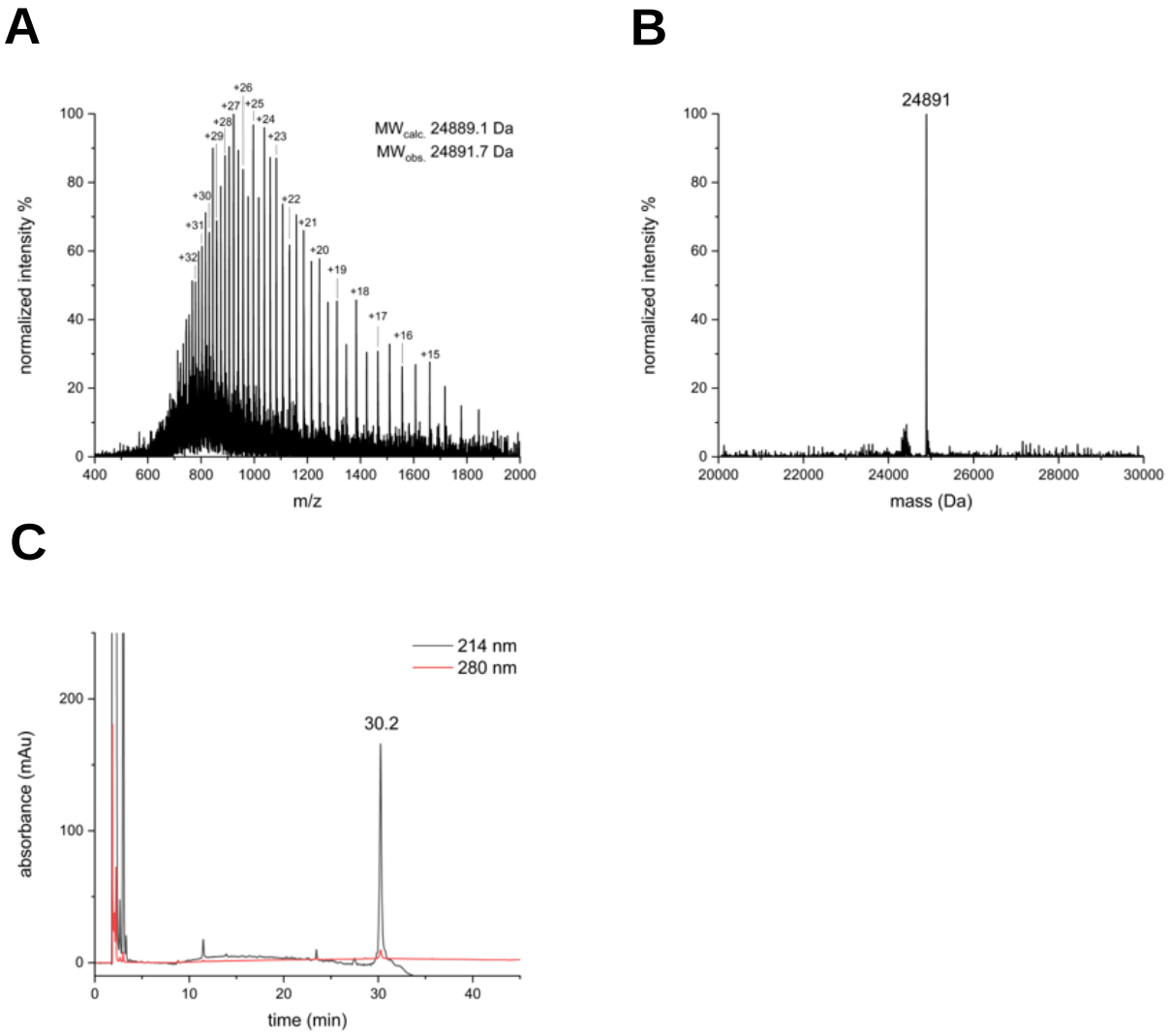

Figure 13 A: Characterization of Hsp90 21-235 *via* ESI-MS. **B:** Deconvoluted mass spectra of Hsp90 21-235, the observed mass (24891.7 Da) is in good agreement with the calculated mass (24889 Da). **C:** Characterization of recombinant Hsp90 21-235 *via* RP-HPLC.

3.3 Ligation of synthetic Hsp90 2-20 peptides to recombinant Hsp90 21-235

After Having sufficient amounts of the recombinant segment **2** and the synthetic peptides hydrazides **1a-d** at hand, we aimed to test and optimize the ligation of these two segments (**Figure 14**).

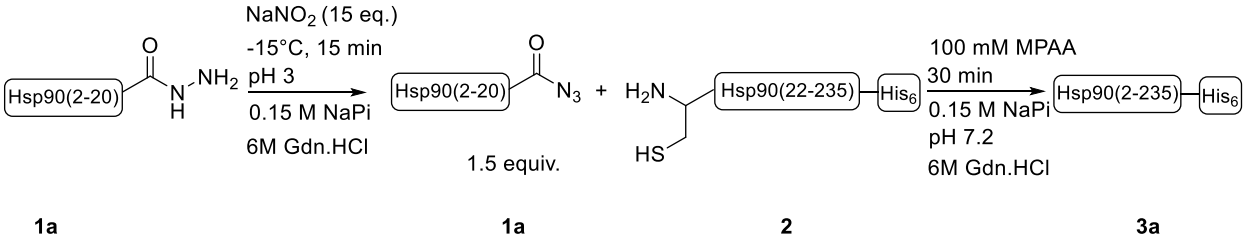

Figure 14 Schematic of the *in situ* thioester conversion and ligation of synthetic peptide hydrazide **1a** to recombinant segment **2**.

We were excited to see a good conversion (~80-90%) and fast product formation (~30 min) when performing *in situ* thioester conversion of the unmodified synthetic peptide hydrazide **1a** using MPAA as an aryl thiol catalyst, as depicted in **Figure 15**.

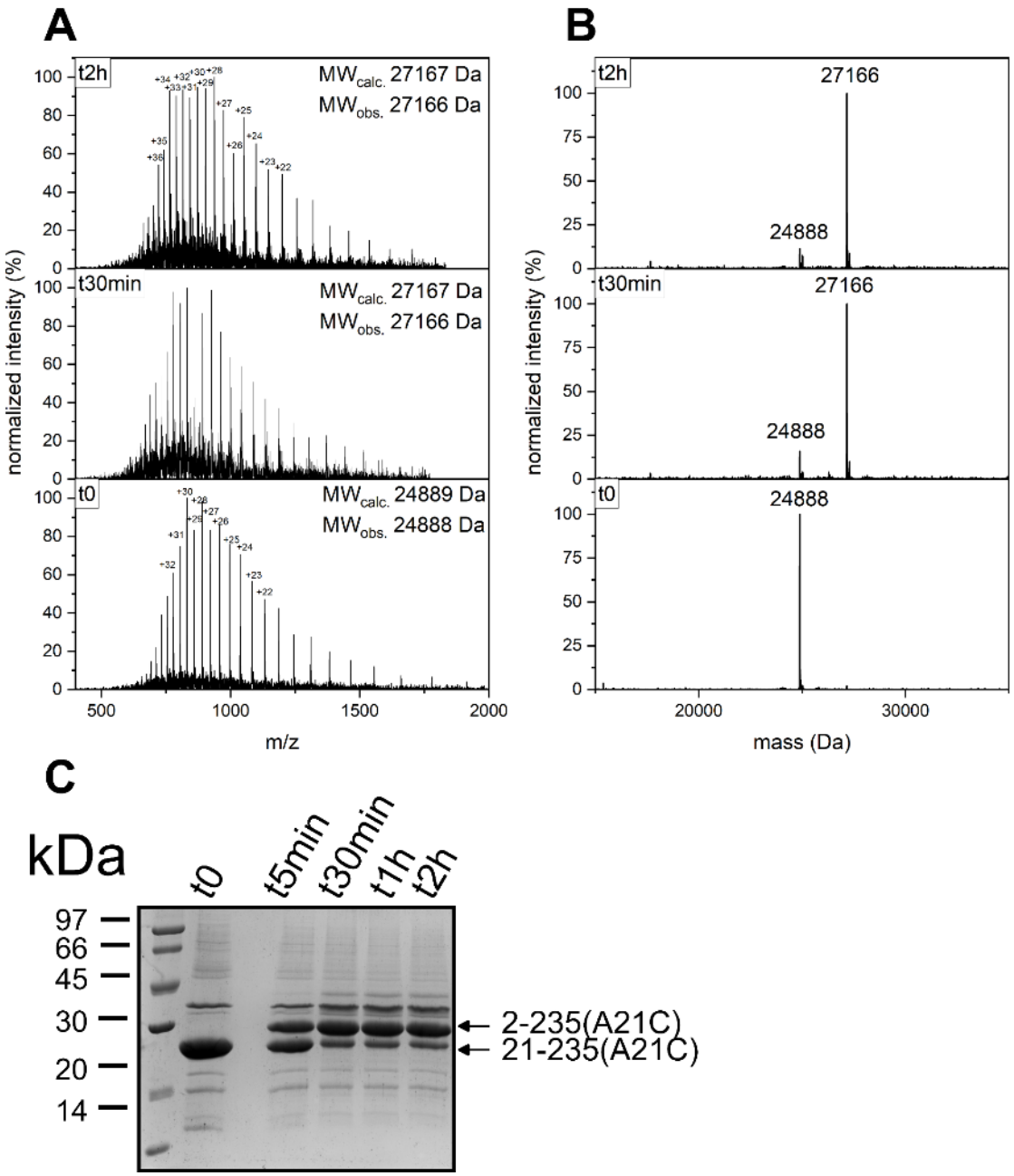

Figure 15 Ligation Monitoring of in-situ generated synthetic Hsp90 2-20 thioester to recombinant 21-235. **A:** Ligation monitoring via ESI-MS **B:** Deconvoluted ESI-MS spectra of ligation monitoring. **C:** Ligation monitoring via SDS-PAGE.

Analytical characterization *via* ESI-MS and RP-HPLC indicated good purity of the isolated product after purification *via* RP-HPLC (**Figure 16**).

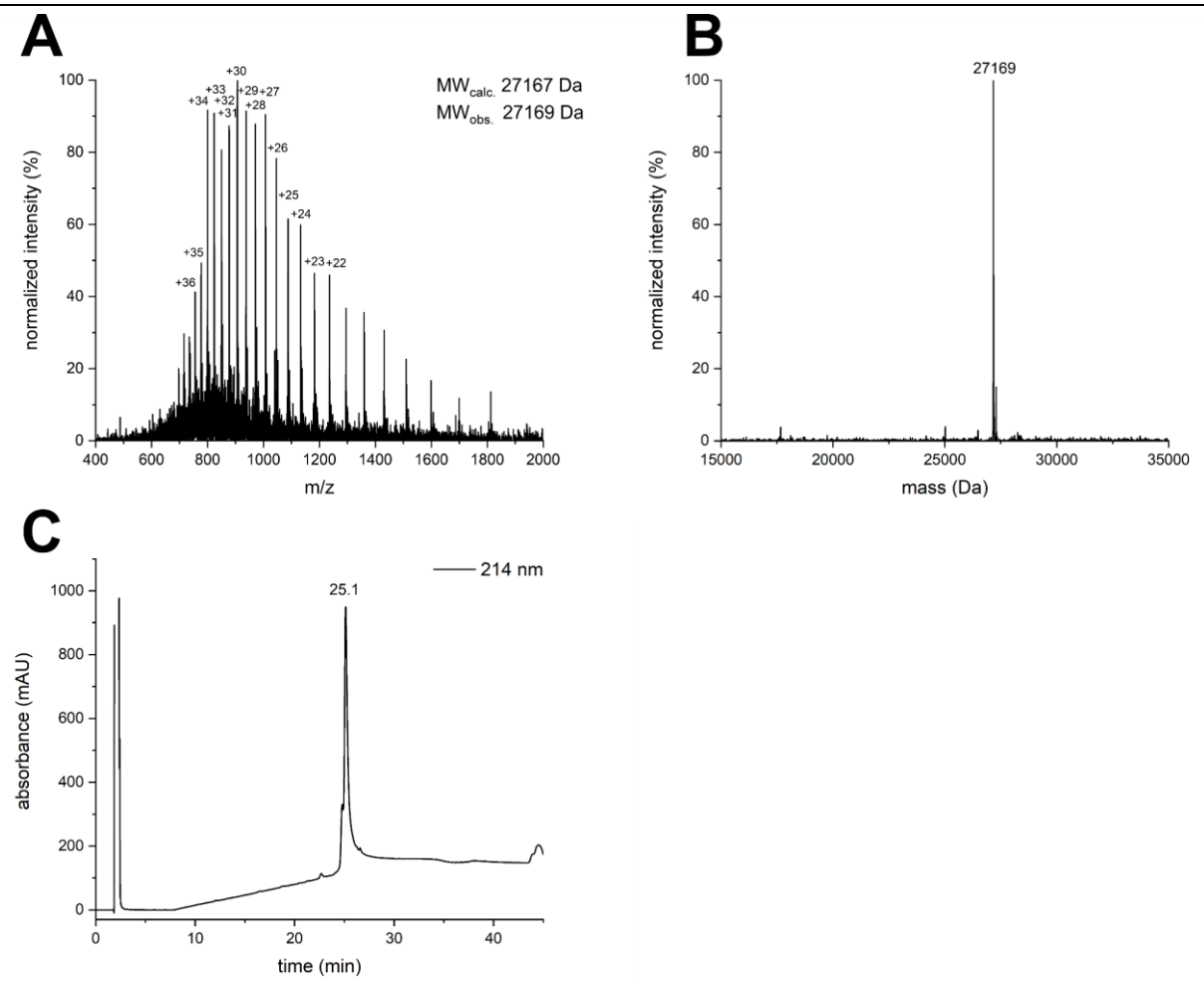

Figure 16 Characterization of semi-synthetic unmodified Hsp90 2-235 (A21C) NTD **A:** ESI-MS characterization. **B:** Deconvoluted ESI-MS spectrum. **C:** RP-HPLC analysis.

However, unfortunately, we experienced extreme difficulties concerning the purification of the ligation product indicated by late elution profiles and peak tailing that resulted in extremely low RP-HPLC recovery rates of around ~15%. All efforts trying to improve the overall purification performance and yield *via* RP-HPLC, which included focused solvent gradients, higher temperatures, different column materials, and solvent compositions, did not significantly improve the overall recovery yield of the product. This can be attributed to the high hydrophobicity and aggregation proneness of the ligation product 5. Nevertheless, we continued to perform the post-ligation desulfurization of the non-native Cys21, where it was furthermore shown that, although the reaction proceeded when using radical desulfurization *via* VA-044 chemistry, difficulties in getting the protein in solution and proper monitoring of the reaction as well as slow reaction rates require a di a different strategy for sufficiently generating the desired Hsp90 NTD semi-synthetic protein variants (**Figure 17**).

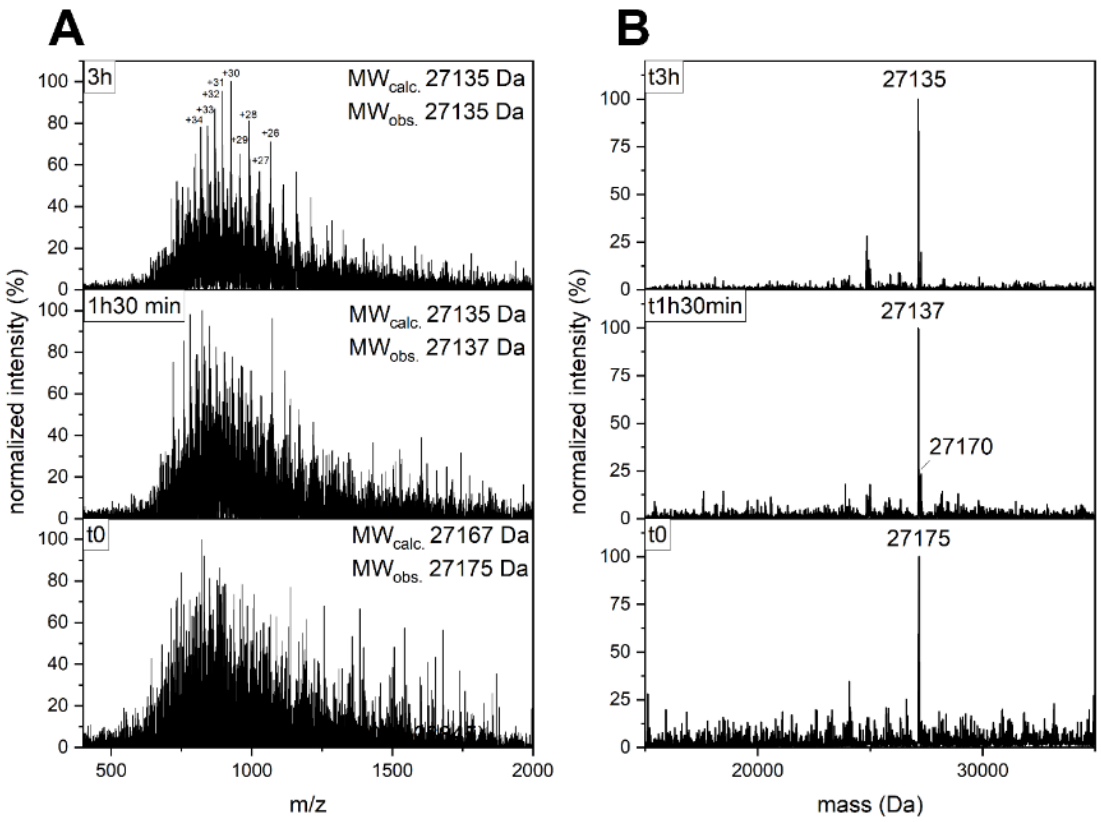

Figure 17 Monitoring of the VA-044 initiated desulfurization of unmodified semi-synthetic Hsp90 2-235 (A21C). **A:** ESI-MS spectra of the desulfurization monitoring. **B:** Deconvoluted ESI-MS spectra.

We concluded that switching to a one-pot ligation-desulfurization strategy would enable the generation of semi-synthetic NTD variants in low milligram amounts, which would still be sufficient for protein characterization despite the necessary final purification via RP-HPLC, which causes significant yield loss. Despite its known strong radical quencher properties, we decided to continue using MPAA as a catalyst due to our successful first ligations. However, we aimed for an extraction using ethyl acetate before continuing with the desulfurization, as this should remove large amounts of the aryl thiol.[110] We also switched to the novel and presumably more potent "add-and-done" radical desulfurization methodology using $NaBEt_4$ according to the procedure of Li and coworkers, which has been reported to desulfurize up to seven Cys residues in model protein within minutes and work in the presence of minor amounts of MPAA.[116] Furthermore, to ensure optimal reaction conditions and avoid any unwanted quenching or side reactions, we performed the thioester conversion of the doubly phosphorylated peptide hydrazide **1d** previous to the ligation and used RP-HPLC purified C-terminal mesna thioester peptide as starting material. As expected, the ligation proceeded successfully, and the doubly phosphorylated Hsp90 2-235 (A21C) product was formed, as depicted in **Figure 18**.

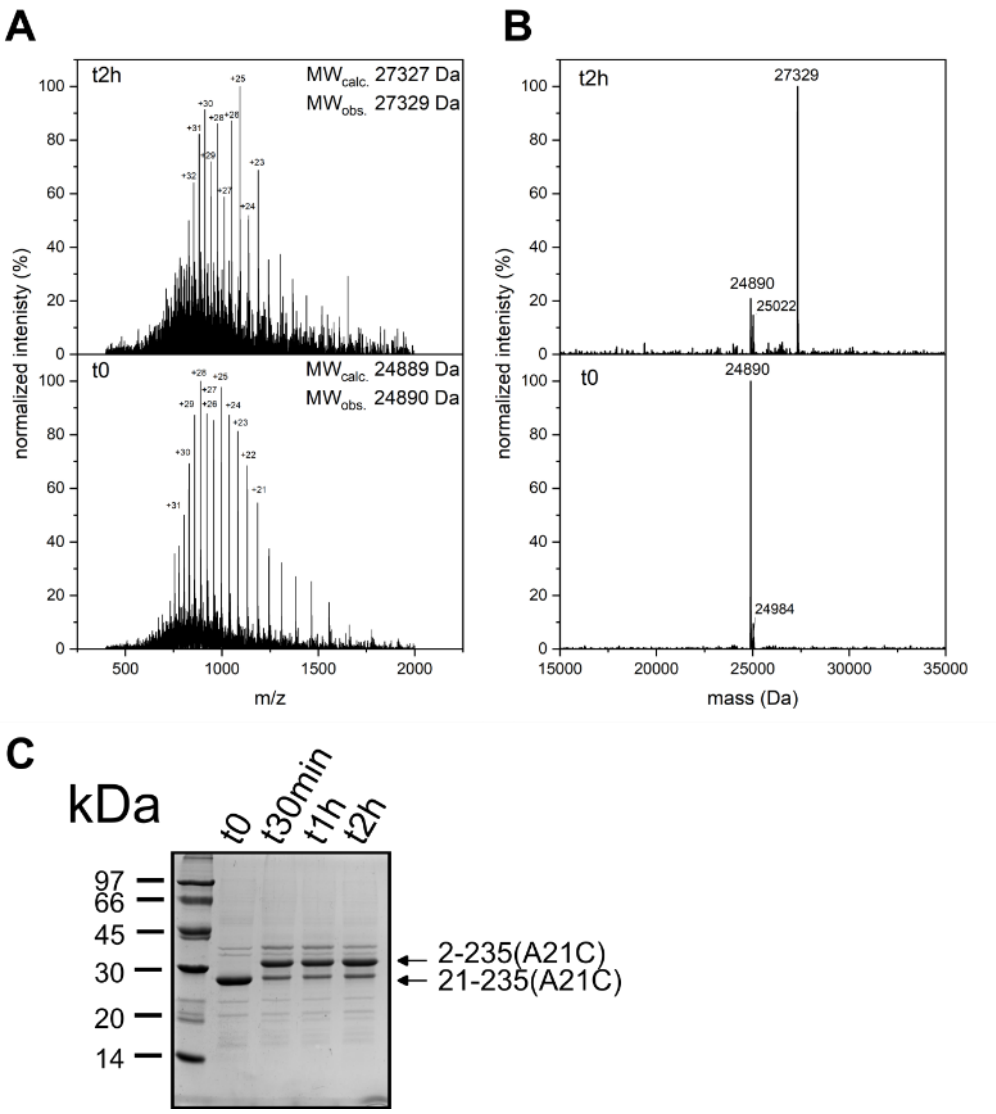

Figure 18 Ligation monitoring of doubly phosphorylated synthetic peptide thioester **1d** to recombinant segment Hsp90 2-235 (A21C). **A:** ESI-MS spectra of the ligation monitoring. **B:** Deconvoluted ESI-MS spectra. **C:** Ligation monitoring *via* SDS-PAGE.

However, after MPAA extraction and continuing with the desulfurization, only minor desulfurization product formation was observable (**Figure 19**), and overall bad signal intensity pointed towards aggregation and possibly precipitation of the ligation product.

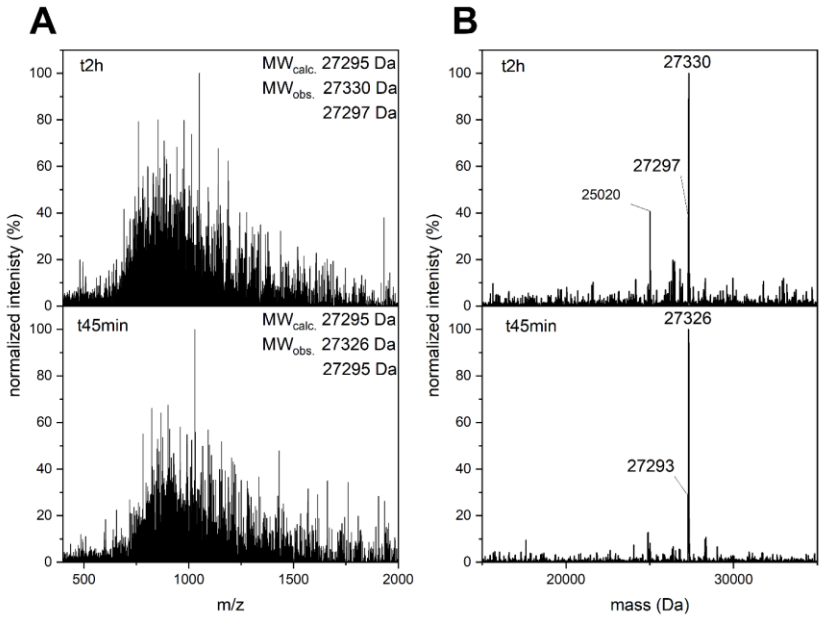

Figure 19 Monitoring of one-pot desulfurization of doubly phosphorylated semi-synthetic Hsp90 2-235 (A21C) using radical mediated $NaBEt_4$ chemistry. **A:** Monitoring *via* ESI-MS**. B:** Deconvoluted ESI-MS spectra**.**

Further attempts to use different thiols as ligation mediating catalysts, including mesna (**Figure A 6** and **A 7)** or MTG (**Figure A 8**), also did not prove successful since the ligation performed poorly compared to when MPAA was applied or showed no improvement of the desulfurization reaction. Until now, we have not seen visible alternatives to overcome this obstacle. Therefore, the desired upscaling and sufficient generation of semi-synthetic NTD to the point of biophysical characterization could not be achieved.

4. Conclusion

This chapter describes the first-ever attempt to study the impact of Hsp90 NTD phosphorylation through semi-synthetically generated site-selectively phosphorylated Hsp90 NTD variants. Although initial preparation of the required synthetic phospho-peptide variants, as well as the generation of the recombinant segment, turned out to be successful, the overall high hydrophobicity and aggregation proneness of the NTD have proven to be an insurmountable obstacle for obtaining sufficient semi-synthetic NTD. Large efforts were undertaken to optimize the synthetic and recombinant segment generation, and ligation conditions were optimized to good conversion (~80-90%). However, the final step requiring desulfurizing the non-native Cys21 has proven an enormous hurdle in producing the final NTD variants in adequate yields. Although we can show that the underlying chemistry can, in principle, be applied, including the desulfurization of purified ligation products, all our attempts did not provide access to semi-synthetic NTD in at least low milligram amounts. Desulfurization only proceeds when using purified and desalted ligation products, but high losses due to low RP-HPLC recovery rates make such an in-between purification technically inefficient. However, not all commonly known thiol catalysts, e.g., 2,2,2-Trifluoroethanol (TFE), or radical initiators for desulfurization, e.g., tetrahydroxydiboron ($B_2(OH)_4$), have been tested and could provide the highly desired one-pot procedure for the generation of semi-synthetic NTD.[33,34] Also, a visible-light-induced desulfurization methodology using the photosensitizer Rose Bengal was recently reported and could be tested on the NTD as a post-ligation desulfurization strategy.[119] Future attempts could focus on different methods for a simple and quick desalting or buffer exchange between the ligation and desulfurization steps, thereby minimizing product loss while performing the radical-based desulfurization efficiently. Nevertheless, the observation that time-dependent product aggregation and precipitation provide another significant hurdle. To overcome this obstacle, the recombinant segment 21-235 could be C-terminally fused to a cleavable solubility tag, e.g., a SUMO-tag or GST tag, which would improve the solubility properties of the protein during the ligation and desulfurization reaction. However, since these tags are of significant size (SUMO-tag: ~ 12 kDa, GST: ~ 26 kDa) and can influence the function of the NTD, enzymatic post-ligation cleavage would be necessary, thereby adding another potentially problematic step.[120] As an alternative, synthetic solubility tags can be introduced into the peptide 2-20 during SPPS, giving the advantage of a chemical-based cleavage after ligation.[121] For example, Liu and coworkers developed a removable backbone modification that allows the introduction of a solubilizing Arg4-tagg into any primary amino acid using Fmoc-SPPS.[122] After

the ligation and desulfurization, TFA treatment can remove the tag to afford the desired native protein.

In summary, the efforts undertaken here showcase how, although in principle possible, the application of semi-synthesis to novel protein targets like Hsp90 can be challenging in unpredicted ways as the required generation of synthetic and recombinant starting material, as well as the ligation and post-ligation chemistry can entail unforeseeable obstacles. This makes the development of novel ligation chemistry and strategies for building tailored proteins desirable. Nonetheless, the progress in the preparation of homogenous and site-selectively modified Hsp90 NTD presented here should be taken as an initiator to pursue this goal further, as success could provide unique insights into the regulatory mechanisms of this vital target protein.

CHAPTER TWO

Semi-synthesis of Hsp90α C-terminal Domain

Most of the results presented in the following chapter can be found in the accepted and released publication:

Site-specifically Phosphorylated Hsp90C-terminal Domain Variants Provide Access to Deciphering the Chaperone Code.

Oliver Gajsek, Christian F. W. Becker,* and Anne C. Conibear*

* Corresponding authors

Chemistry—A European Journal **2025**, *31* (6), e202403676.

I contributed to this work by designing and executing the synthesis strategy for the generation of semisynthetic Hsp90 CTD variants and expression of WT Hsp90 CTD. Furthermore, I conducted all the analytical and biophysical characterizations except for NMR spectroscopy. I also visualized and interpreted the data (except NMR data) under the supervision of the corresponding authors and prepared the manuscript and supplementary information.

1. Introduction

This chapter discusses the semi-synthesis of C-terminally modified human Hsp90alpha CTD. The detailed ligation strategy is depicted in **Figure 20**, showing the EPSL-based generation of four different CTD variants, bearing either no modifications as reference protein, single phosphorylation at either position pThr725 or pSer726 or doubly modified (pThr725 + pSer726). The ligation site at position 711 was chosen so that the synthetic peptide segments 711-732 (A711U) can be prepared using standard Fmoc SPPS while the recombinant segment 572-710 is prepared as intein-fusion to yield the desired peptide hydrazide which can be transformed to the desired selenoester for DSL *in situ*.

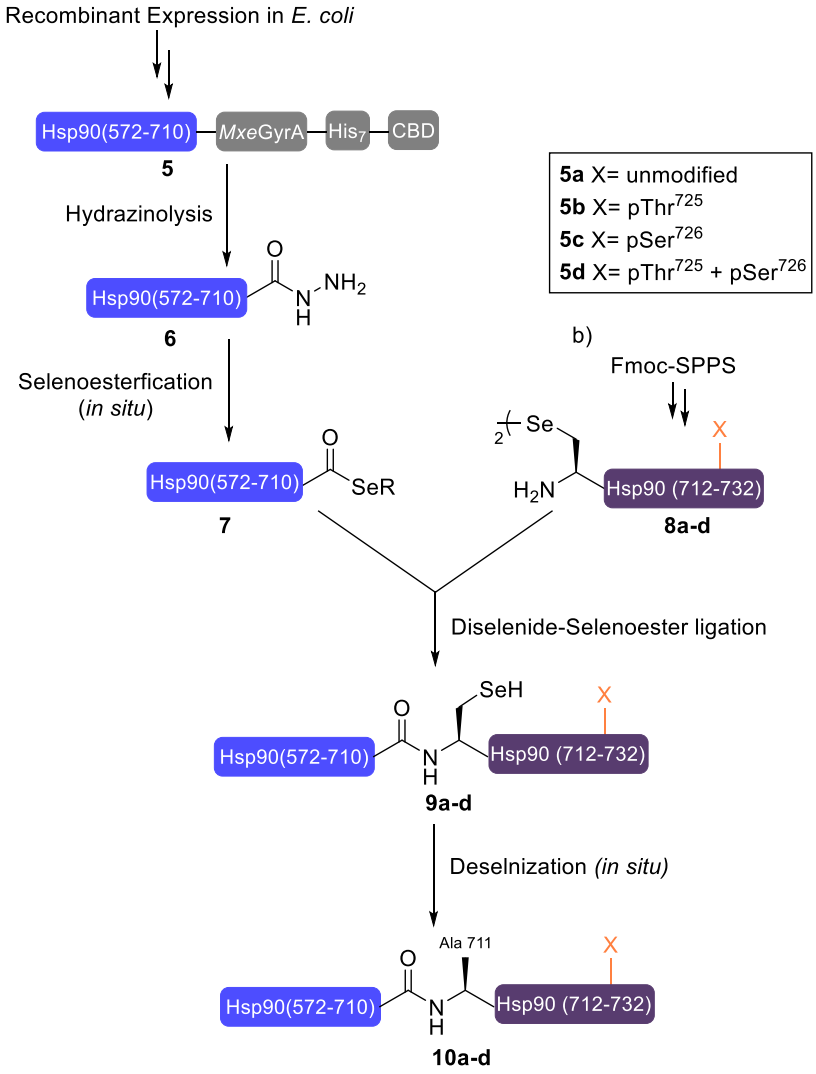

Figure 20 Schematic overview of site selectively modified Hsp90 CTD generation *via* Expressed Protein Selenoester Ligation**.**

2. Materials and Methods

2.1 Solid phase peptide synthesis of Hsp90 711-732 (U711A) peptide variants

Synthetic Hsp90 segments 711-732 (U711A) were prepared manually *via* 9-fluorenylmethoxycarbonyl (Fmoc) based solid phase peptide synthesis (SPPS) in 10 mL syringes equipped with a polypropylene filter on preloaded Tentagel resin (loading: 0.19 mmol/g). Fmoc-protected amino acid building blocks (2.5 equiv.) were double-coupled for 20-30 min at room temperature using *O*-(7-Azabenzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium-hexafluorphosphate (HATU, 2.4 equiv., 0.5 M in DMF) as activator and diisopropylethylamine (DIPEA, 5 equiv.) as base. Fmoc-deprotection was achieved *via* incubation in 20% piperidine (3 x 3 min). Phosphorylated serine/threonine residues were introduced *via* Fmoc-protected Ser(PO(OBzl)OH)-OH or Thr(PO(OBzl)OH)-OH (CEM) respectively and the N-terminal selenocysteine as Fmoc-L-Sec(Mob)-OH (Iris Biotech). The selenocysteine building block was double-coupled using OxymaPure (0.5 M in DMF) and diisopropyl carbodiimide (DIC, 0.25 M in DMF) with a 30-45 min coupling time. After final coupling and deprotection, peptides were washed with DCM and dried under vacuum overnight and cleaved from the resin with trifluoroacetic acid (TFA)/dimethyl disulfide (DMDS)/dimethyl sulfide (DMS)/ triisopropylsilane (TIPS)/H_2O 90:5:2.5:2.5 for 2.5 h followed by ether precipitation and pelleting *via* centrifugation (4000 g, 5 min). The crude peptide pellet was dissolved in acetonitrile/ water 1:1 and lyophilized prior to the final deprotection of the DMDS trans-protected N-terminal selenocysteine *via* dissolving in 6 M guanidinium hydrochloride (Gdn.HCl) 20 mM DTT in ddH_20 for 30 min. The resulting diselenide peptide dimer was purified using RP-HPLC on a preparative scale C18 column (Kromasil) using a gradient of ACN in water with 0.1% TFA. The purity of the generated peptides **8a-d** was assessed *via* electrospray ionization mass spectrometry in positive ion mode and *via* analytical RP-HPLC.

2.2 Recombinant expression and purification of Hsp90 572-710-hydrazide

For the generation of the recombinant segment Hsp90 572-710-hydrazide (**6**), a synthetic oligonucleotide comprising human Hsp90alpha 572-710-*Mxe* GyrA-His_7-CBD was ordered from Eurofins Genomics in a pET-21a (+) vector and heat shock transformed into chemically competent *E. coli* strains XL1-Blue, BL21 (DE3) and Rosetta 2 (DE3) from Invitrogen. *E. coli* Rosetta 2(DE3) was used for expression, and cells were grown in LB medium (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 10 g L^{-1} NaCl) containing 100 µg mL^{-1} ampicillin and 30 µg mL^{-1} chloramphenicol. Overnight cultures were diluted to a starting OD of ~ 0.2 and were grown at 37 °C with shaking at 170 rpm to an OD of 0.7-0.8 before induction of expression with 1 mM IPTG. After 4 h at 37 °C (170 rpm), cells were harvested *via* centrifugation (12000 g, 20 min), and cell pellets were stored at -80 °C before further workup. After thawing, 1 × TBS, pH 7.8 buffer was added to the cell pellet (50 mL buffer / 1 L expression pellet). The solution was homogenized using an ultra turax before cell lysis *via* sonification on ice (60% amplitude, 15 sec ON, 45 sec OFF time, total ON time 10 min). After clearing the supernatant *via* centrifugation (21000 rpm, 53000 g, 45 min at 4 °C), the collected inclusion bodies were solubilized by incubation in 6M guanidinium hydrochloride, TBS buffer at pH 7.5 while stirring overnight at rt. After removal of insoluble cell debris (21000 rpm, 53000 g, 45 min), the solution was directly dialyzed against 4M Urea, TBS pH 8 at 4°C overnight before inducing intein-mediated hydrazinolysis *via* dropwise addition of 10% hydrazine, 100 mM DTT in 1 TBS pH 8 (1:1 dilution) (final concentration: 5% Hydrazine, 50 mM DTT, 2 M urea in TBS). After 2 h the solution was diluted further 1:1 with 6M Gdn.HCl, filtered and purified *via* RP-HPLC on a preparative scale C4 column (Kromasil) with a gradient of 30-60% ACN in water (+0.1% TFA) at a flow rate of 10 ml/min at 60 °C. Peak fractions were pooled and lyophilized, yielding compound **6** at ~10 mg/L culture.

2.3 Selenoester formation, diselenide-selenoester ligation and deselenisation

The purified and lyophilized recombinant segment **6** was dissolved in freshly prepared and degassed 6M Gdn HCl, 100 mM TCEP, 50 mM DPDS, and 200 mM HEPES buffer at a concentration of 1-1.3 mM, and the pH adjusted to 1.5. Selenoester formation was initiated by the addition of ~30 equiv. acetylacetone and the reaction monitored *via* electrospray ionization mass 15-30 min. Synthetic peptide variants Hsp90 **8a-d** were dissolved in 6M Gdn.HCl, 5 mM DPDS, 200 mM HEPES buffer (1.5 equiv. based on recombinant segment) and directly added to Hsp90 (572-710)-SeR *via* 1:2 dilution (final TCEP concentration in the ligation mixture 33mM) after successful selenoester formation was confirmed *via* ESI-MS. The pH was adjusted to 3.5-4 using NaOH, and the Eppendorf tube was flushed with argon before incubation of the reaction mixture at r.t. under stirring. The ligation reaction was monitored *via* ESI-MS and SDS-PAGE, and after 1 h, excess phenylselenol (DPDS) was removed *via* hexane extraction (5 x, 1:1 vol/vol) before continuing with the deselenization. Deselenization was performed by diluting the reaction mixture 1:1 with freshly prepared and degassed deselenization buffer (6M Gdn.HCl, 250 mM TCEP, 200 mM HEPES, 25 mM DTT). The pH was adjusted to 7.4, and the reaction was stirred at r.t. for 4-5 h under an argon shower before freezing at -80 °C. After thawing and dilution with 6M Gdn.HCl in ddH_2O, the ligation product was purified on a semipreparative scale C4 column (Kromasil) using a linear gradient from 30-65% of solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) at a flow rate of 5 mL/min at 60 °C. Fractions containing the deselenized ligation product were analyzed by ESI-MS, combined, and lyophilized. Final yields of lyophilized ligation products ranged between 45-50% based on segment **6** as limiting starting material.

2.4 Recombinant expression and purification of Hsp90 WT CTD

The synthetic oligonucleotide comprising N-terminally His_6-tagged human Hsp90alpha CTD was ordered from Eurofins Genomics in a pET-21a (+) vector, and heat shock transformed into chemically competent *E. coli* strains XL1-Blue, BL21 (DE3) and Rosetta 2 (DE3) from Invitrogen. Expression using *E. coli* Rosetta 2(DE3) strain and subsequent cell lysis was performed analogously to the fusion construct **5**. After cell lysis and separation of cell debris *via* centrifugation (21000 rpm, 53000 g, 45 min at 4 °C), the cleared supernatant, which contained the soluble fraction of the overexpressed protein, was loaded onto a preequilibrated Ni-NTA resin on an Aekta Prime FPLC system. The protein was eluted using a 25-500 mM imidazole gradient, and the fractions were analyzed using SDS-PAGE. Pooled fractions were further purified *via* RP-HPLC on a preparative scale C4 column (Kromasil) with a gradient of 30-75% ACN in water (+0.1% TFA) in 40 min at a flow rate of 10 ml/min at 60 °C. Peak fractions were combined and lyophilized, yielding the desired protein at ~15 mg/L culture (**Figure A 2**).

2.5 Protein refolding and circular dichroism spectroscopy

Lyophilized semi-synthetic Hsp90 variants and recombinant WT CTD were dissolved in 6M Gdn.HCl TBS buffer and dialyzed towards assay buffer (50 mM KPO_4 buffer pH 7.6) at 4 °C overnight. Circular dichroism spectroscopy was performed in 50 mM KPO4 buffer pH 7.6 at a protein concentration of 0.2 mg/mL using a JASCO J-815/150S spectropolarimeter at 20 °C and a path length of 1 mm. Scans were recorded in the range of 195 to 260 nm, and the spectra shown are the average of five measurements after signal correction *via* buffer subtraction.

2.6 Thermal stability determination by nano differential scanning fluorimetry

Thermal stability assessments were performed using a Prometheus NT.48 (Nanotemper) in 50 mM NaPi, 150 mM NaCl pH 7.5 buffer at 2 mg/mL protein concentration. Protein denaturation was monitored by measuring the ratio of the tryptophan fluorescence intensities at 350 nm and 330 nm while increasing the temperature from 20 to 95 °C at 1 °C/ min. Triplicates were measured, averaged, and corrected *via* buffer subtraction. Depicted curves were normalized between values 0.0 and 1.0 for better visual comparison.

2.7 Aggregation assays

The anti-aggregation (chaperoning) capabilities of Hsp90 variants **10a-d** were measured using alcohol dehydrogenase (ADH) and Insulin aggregation assay similar to previously reported methods.[123], [124] Equine ADH (Sigma) was directly dissolved in assay buffer (50 mM NaPi, 100 mM NaCl pH 7.0), and the concentration was determined *via* BCA assay before dilution to the assay concentration (3.1 µM). CTD variants **10a-d** were dialyzed towards assay buffer, the concentration determined (BCA), and added to ADH to a final concentration of 500 nM (final volume 1400 µL). ADH aggregation was thermally induced at 55 °C and monitored over 60 min. by measuring the absorbance at 360 nm in a SAFAS UVmc2 spectrophotometer equipped with a temperature-controlled multicell holder (SAFAS, Monaco) in 700 µL quartz cuvettes (Hellma Analytics). Samples were measured as triplicates, and the endpoint of ADH aggregation readout was set to 100% to compare chaperoning efficiency.

To determine insulin aggregation inhibition, bovine pancreas insulin (Sigma) was dissolved in 1% acetic acid in ddH_2O, and the stock concentration was determined using a NanoDrop 2000 (molar extinction coefficient used: 5,734 $cm^{-1}M^{-1}$). The insulin stock was diluted to 40 µM using assay buffer (50 mM NaPi, 100 mM NaCl pH 7.4) before adding Hsp90 CTD variants **10a-d** (4 µM) dialyzed into the assay buffer. The protein solution was incubated for 15 minutes at rt before adding 20 mM DTT induced Insulin aggregation. Insulin aggregation was measured by monitoring the absorbance at 360 nm in a SAFAS UVmc2 spectrophotometer equipped with a temperature-controlled multicell holder (SAFAS, Monaco) in 700 µL quartz cuvettes (Hellma Analytics) for 45 minutes. Triplicates were measured, baseline subtracted and averaged, and the endpoint of Insulin aggregation was set to 100% to compare chaperoning efficiency.

2.8 Dimerization assessment *via* Crosslinking and western blotting

Dimerization of Hsp90 CTD variants **10a-d** was assessed by incubation of the protein with the amine-reactive crosslinker bis(sulfosuccinimidyl) suberate (BS3) (Pierce) as previously described. The Protein was dialyzed towards the assay buffer (50 mM NaPi pH 7.4) and diluted to 2 µM before adding BS3 from a freshly prepared stock in a total volume of 25 µL (final BS_3 cons.:20 µM). After incubation for 1 h at rt the reaction was quenched by the addition of 50 mM Tris pH 7.4 for another 15 min before adding 2x SDS sample buffer. Samples were boiled at 95 °C for 5 min before separating them on a precast 12% SDS-PAGE gel (BioRad) at 200V for 35 min following western blotting. Semi-dry western blotting was performed by transferring on a PVDF membrane using a Turbo Plotter (30 min, 200V) followed by blocking with 5% w/v nonfat milk in TBST (0.05% v/v Tween 20) and probed *via* anti-Hsp90 AC88 (Abcam) primary antibody.

2.9 Characterization of variants 10a-d *via* Size Exclusion Chromatography

Proteins were characterized *via* size exclusion chromatography on a Superdex 200 (10x300 mm) column equilibrated with 10 mM MES, 200 mM KCl, 1 mM EDTA, 1% Glycerol pH 6 on a ÄKTA Purifier system (both GE Healthcare) with a flow rate of 0.5 mL/min at 4°C. Samples were centrifuged for 15 min, 14.000xg at 4 °C before being loaded on the column, and the following proteins were used as MW calibration: Aproptinin,6.5 kDa, Ribonuclease, 13.7 kDa; Carbanhydrase, 29kDa;Ovalbumin, 43 kDa; Conalbumin, 75 kDa, Aldolase,158 kDa (GE Healthcare).

3. Results and Discussion

3.1 Synthesis route for accessing site-selectively phosphorylated Hsp90 CTD

While planning our ligation strategy for installing the desired PTMs at positions Thr725 and Ser726, we quickly realized that a classic NCL strategy with the C-terminal cysteine-containing part being synthetically made *via* SPPS and the N-terminal part being generated *via* recombinant expression is not suitable due to the location of all three native occurring cysteines (Cys572/597/598) close to the amino-terminal end of the domain (sequence in appendix 5.1.5).[65,93] Using an Alanine residue as a position for installing a Cysteine or Thiol-modified ligation handle as a surrogate was also excluded since it would require a desulfurization step to gain access to the native residue but would, therefore, cause the loss of the other essential native cysteine residues and potentially alter the folding and/or function of the CTD. Therefore, we decided to utilize a synthesis strategy based on the relatively recently developed Expressed Protein Selenoester Ligation Protocol, where a C-terminal selenoester reacts with an N-terminal Selenocysteine-containing peptide to form the peptide bond (**Figure 21**).[125,71]

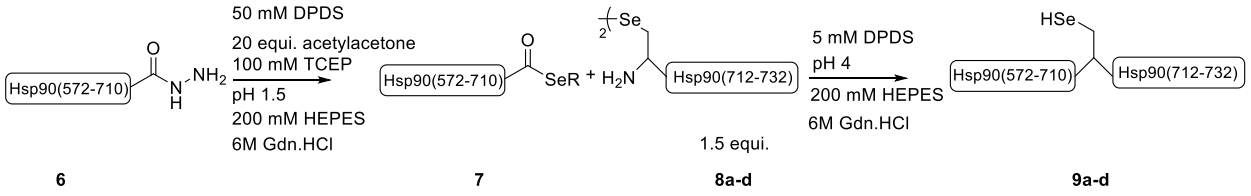

Figure 21 Schematic depiction of the *in situ* selenoester conversion and ligation of recombinant peptide hydrazide **6** to synthetic segments **8a-d** using DSL.

Exchanging the thiol-reactive moieties with selenol groups would not only yield beneficial reaction kinetics and allow for lower concentrations of peptide segments in the reaction mixture but also should enable a selective post-ligation deselenization step to yield an alanine at the ligation site without losing native thiol functionalities.[82] Doing so, we picked an alanine Residue (Ala 711) as the ligation site, enabling us to install the selenocysteine ligation handle on the synthetic peptide segment of modest length (711-732, 21 residues), where we can introduce either single phosphorylations (pThr725 or pSer726) or the double modified variants (pThr725 and pSer726) as well as a reference unmodified variant.

The N-terminal segment 572-710 was designed as a C-terminally fused *Mxe*GyrA-His_7-CBD protein **5** and accessed *via* recombinant expression using *E. coli* as a host organism (**Figure 22**).

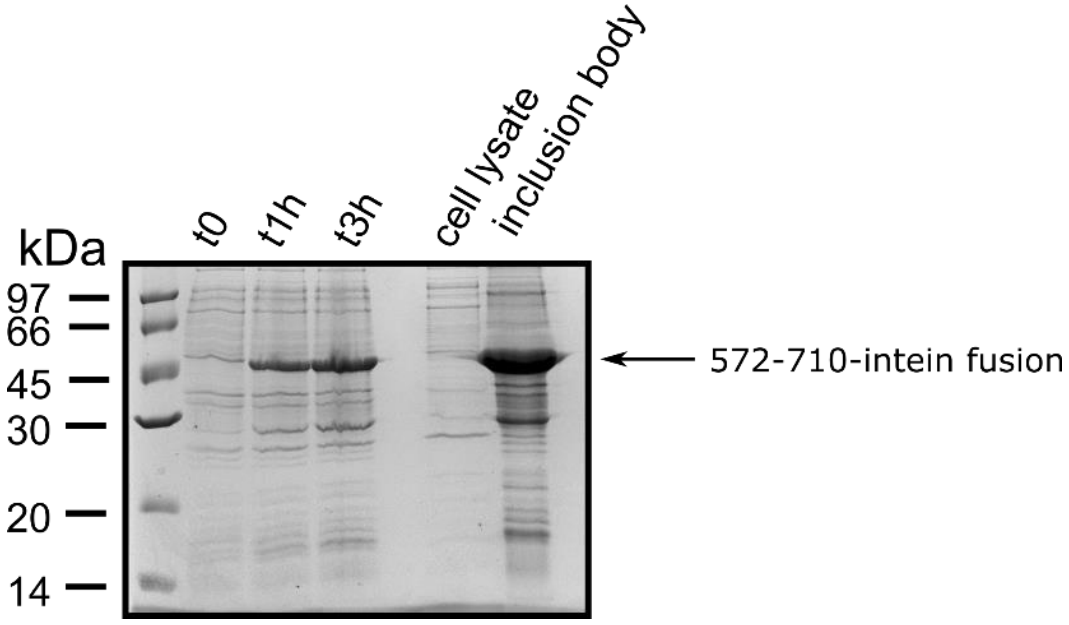

Figure 22 Monitoring of expression of Hsp90 572-710-MxeGyrA-His7-CBD fusion protein **5** *via* SDS-PAGE. The protein appeared as a strong overexpressed band at the ~45 kDa marker height.

Notably, we observed that the translation-initiating methionine of our designed fusion protein was quantitatively cleaved by endogenous methionine aminopeptidase of *E. coli* as we only observed traces of the Met-572-710-intein fusion protein when analyzing the inclusion body pellet (**Figure 23**).

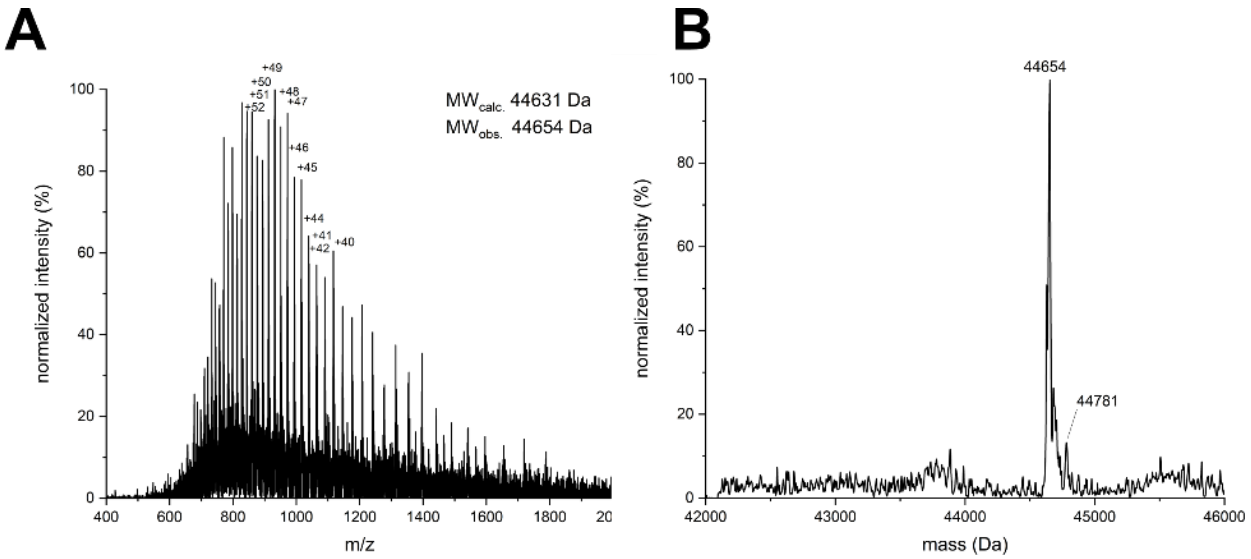

Figure 23 A: Characterization of Hsp90 572-710-intein fusion construct *via* ESI-MS. The de*via*-tion of the observed mass of 44654 Da from the calculated mass of 44631 (+23 Da) can be assigned to the formation of a N-terminal thiazolidine involving the free thiol group of Cys572 and acetaldehyde from the cellular background of *E. coli*. **B:** Deconvoluted ESI-MS spectrum of Hsp90 572-710-intein fusion construct.

The additional +24 Da in the mass we observed for the fusion protein **5** can be explained by the formation of thiazolidine through the reaction of the released N-terminal cysteine with the metabolite acetaldehyde, as this is commonly observed in the production of N-terminal cysteine-containing proteins in *E. coli*.. [126] After solubilization of the fusion protein, which was expressed as inclusion bodies, the intein-affinity tag was cleaved *via* hydrazinolysis (**Figure 24**).

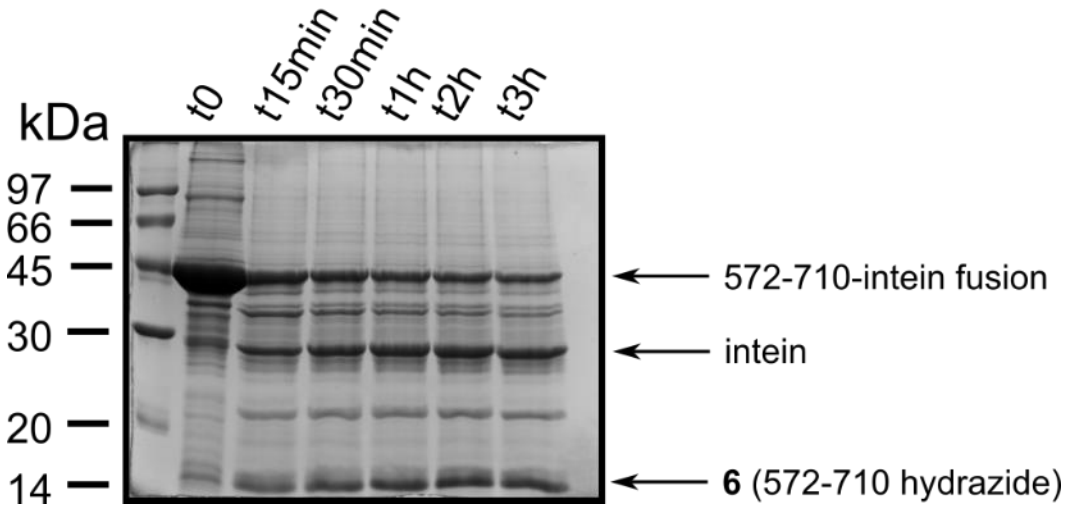

Figure 24 Monitoring of Hydrazinolysis of fusion protein **5** *via* SDS-PAGE. The desired hydrazide **6** was separated from uncleaved fusion protein **5** and cleaved intein-affinity tag *via* RP-HPLC.

This yielded us the C-terminal hydrazide segment **6** after RP-HPLC purification as a precursor for the desired selenoester, which we aimed to produce *in situ* for the ligation to the synthetic peptides 711-732 (A711U) **8a-d**. Mass characterization of the purified hydrazide segment **6** revealed the hydrolysis of the N-terminal thiazolidine, presumably caused by the acidic RP-HPLC purification conditions, as the observed mass of 6 is in excellent agreement with the calculated mass for the free N-terminal cysteine (**Figure 25**).

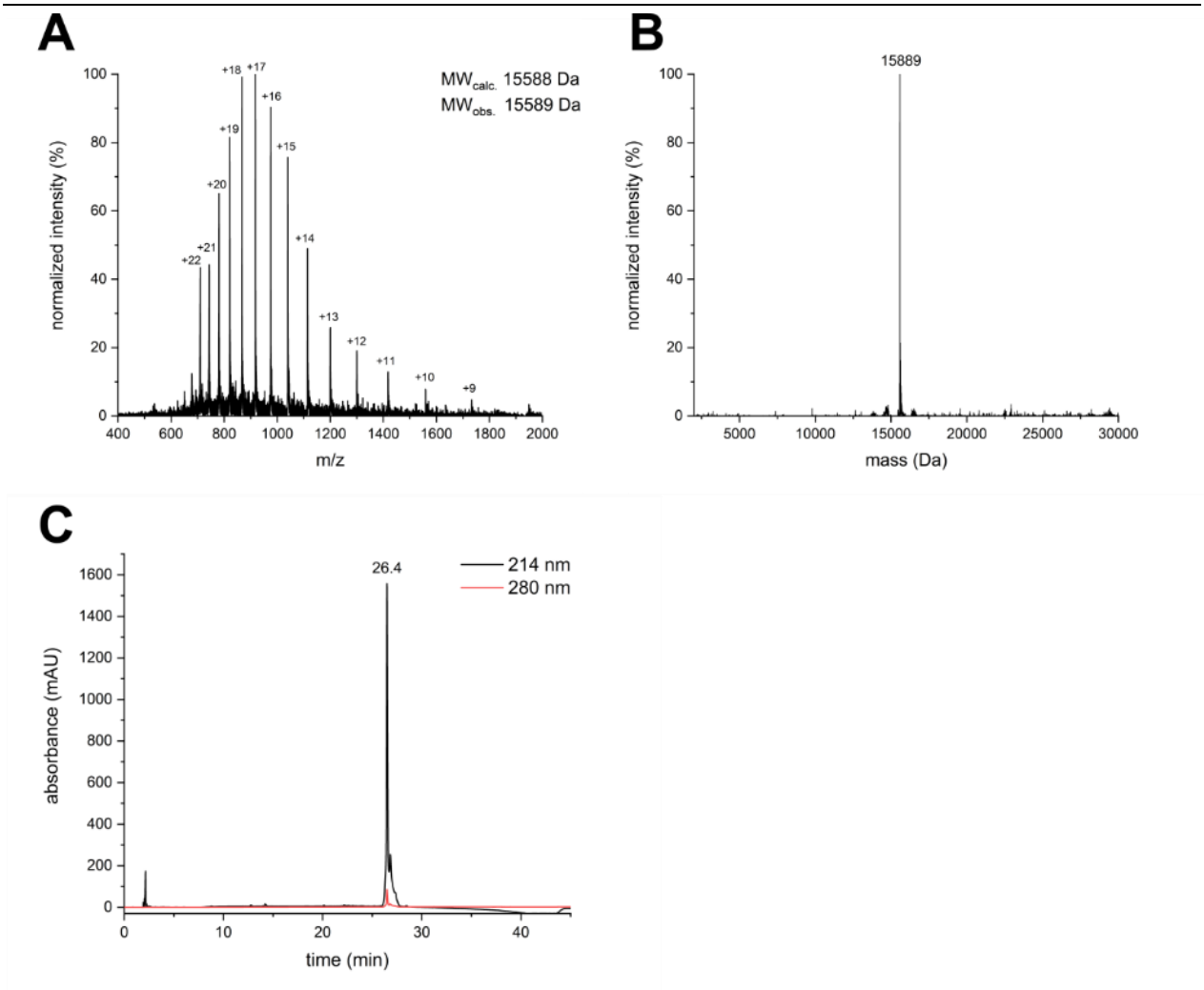

Figure 25 Characterization of Hsp90 572-710 hydrazide (**6**). **A:** ESI-MS characterization of **6**. **B:** Deconvoluted ESI-MS spectrum of **6**. **C:** RP-HPLC characterization with indicated product peak at 26.4 min.

This made us excited to see as it would allow for the potential use of the free thiol moiety of N-terminal cysteine 572 as an additional ligation handle for future reactions but would also potentially require an additional orthogonal thiol protection step to avoid intermolecular ligation of the recombinant segment after the selenoester conversion. However, we continued to generate the necessary synthetic peptide segments **8a-d** before aiming for the selenoester conversion and ligation. These relatively short peptide segments were prepared using standard Fmoc-based SPPS with *O*-benzyl side chain protected phospho-amino acid building blocks (Fmoc-Thr(PO(OBzl)OH)-OH/ Fmoc-Ser(PO(OBzl)OH)-OH) and a 4-methoxybenzyl (Mob) protected Selenocysteine on Tentagel resin. After deprotection and RP-HPLC purification, this yielded the four different variants of segment 711-732 (A711U) either unmodified, singly modified (pThr725 or pSer726), or double modified (pThr725 and pSer726) (**Figure 26**).

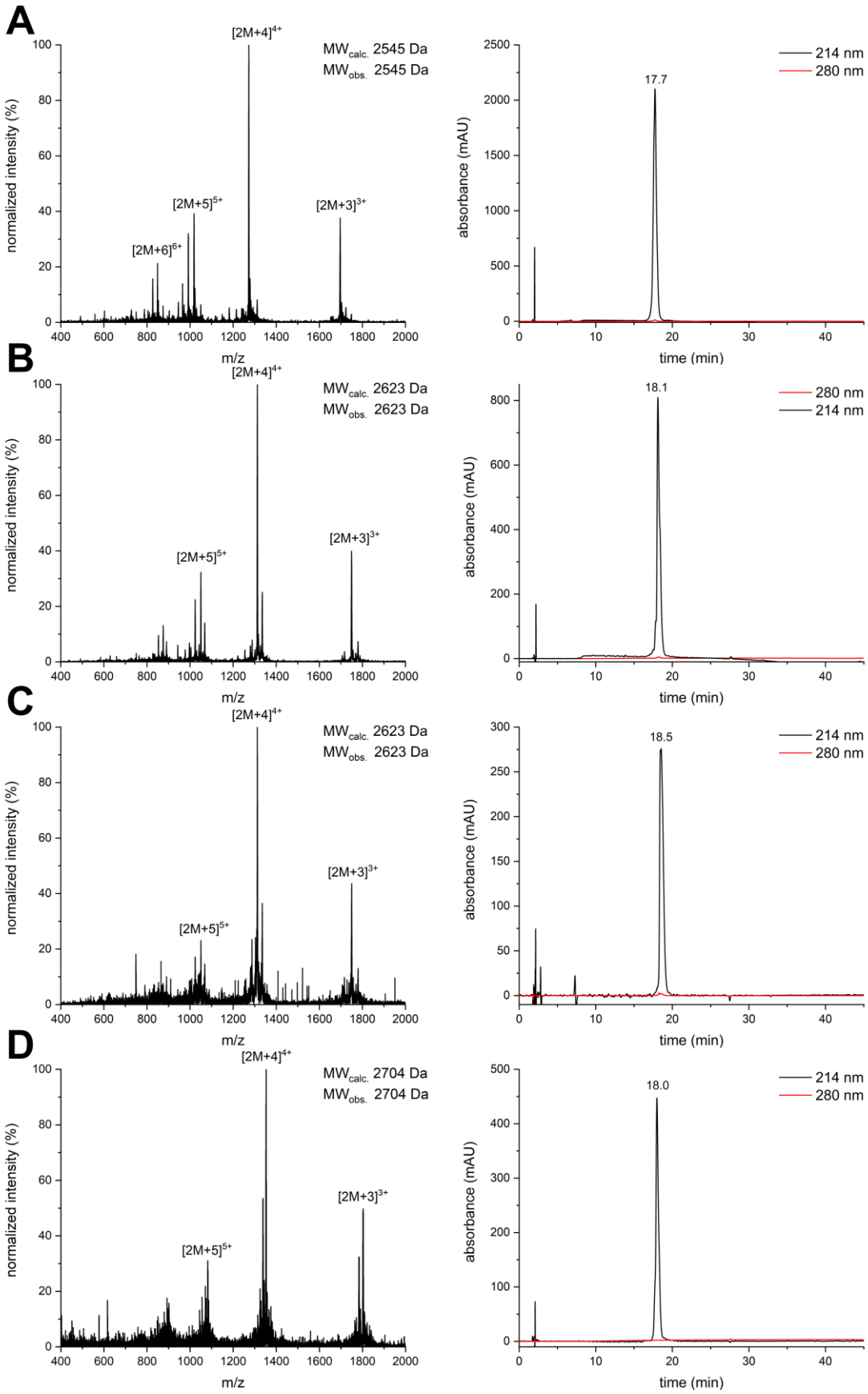

Figure 26 Characterization of synthetic peptides Hsp90 711-732(A711U) **8a-d** *via* ESI-MS (left panel) and RP-HPLC (right panel). **A:** characterization of unmodified peptide **8a**. **B:** Characterization of phosphothreonine peptide **8b** (pThr725). **C:** Characterization of phosphoserine peptide **8c** (pSer726). **D:** Characterization of double phosphorylated peptide **8d** (pThr725/pSer726).

Having the synthetic peptides and the recombinant segment at hand, we focused on optimizing the Selenoester formation of the recombinant hydrazide. We were delighted to observe that the hydrazide cleanly converts to the respective aryl selenoester **6** using acetylacetone to form the active pyrazole intermediate, followed by diphenyl diselenide as a selenol donor (**Figure 27**).

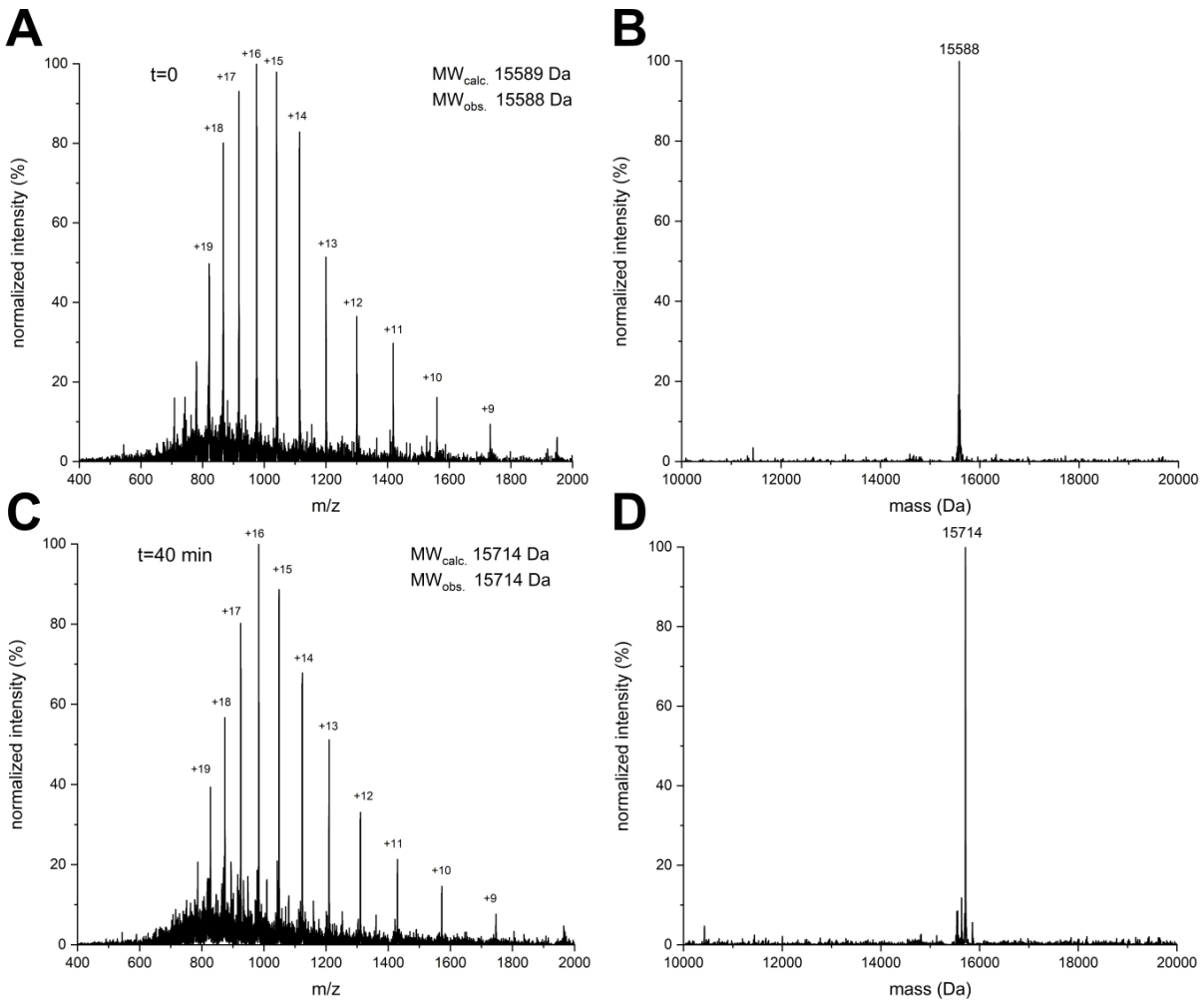

Figure 27 Monitoring of Selenoesterfication of Hsp90 572-710 peptide hydrazide segment **6** *via* LC-MS. **A:** LC-MS spectrum of selenoesterfication at t=0 before the addition of acetyl acetone. **B:** Deconvoluted mass spectrum of t=0. **C:** MS spectrum of selenoesterciation at t=40 min. **D:** Deconvoluted mass spectrum of selenoesterfication at t=40min. The observed mass (MWobs. 15714 Da) agrees with the calculated mass (MWcalc. 15714 Da).

This further motivated us to aim for an in situ generation of the early seleneoester directly before the addition of the synthetic peptide, rather than isolating the freshly formed selenoester through an additional RP-HPLC step as it was done in the original work by Kulkarni *et al*.. We added a slight excess (1.2 eq) of our synthetic peptide segment dissolved in TCEP-free ligation buffer to the freshly formed selenoester, thereby diluting the TCEP concentration to ~ 33mM. This suppressed the premature deselenization of the N-terminal selenocysteine of our synthetic peptide segments through radical-induced TCEP reduction. We were excited to observe the quantitative formation of the ligation products **9a-d** in a matter of 30 minutes to 1 hour (**Figure 28**).

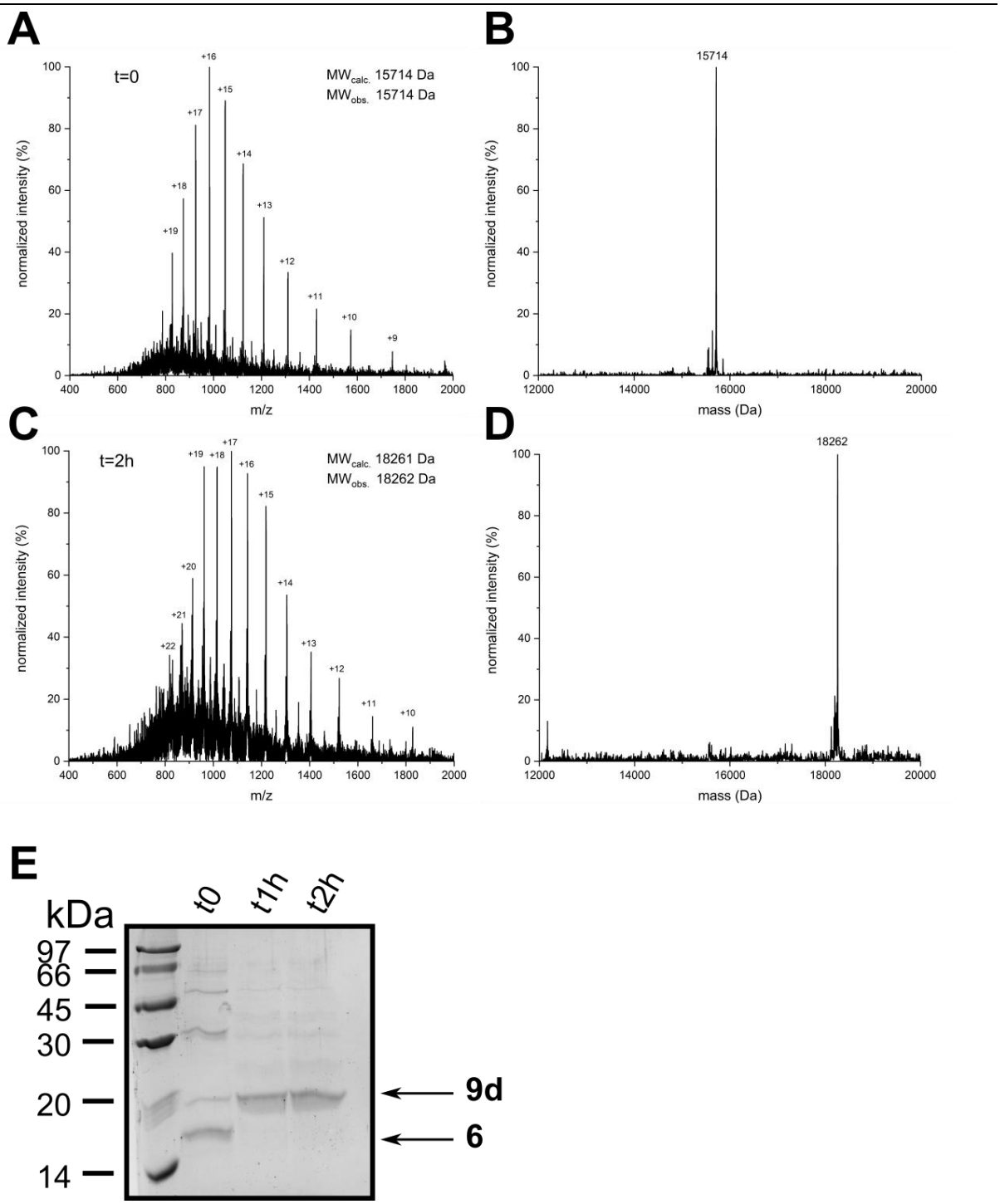

Figure 28 Ligation monitoring of in situ generated aryl selenoester segment **6** with double phosphorylated synthetic peptide **8d**. **A:** LC-MS ligation monitoring at t=0. **B:** Deconvoluted mass spectrum of ligation at t=0. **C:** LC-MS ligation monitoring at t=2 h. **D:** Deconvoluted Mas spectrum of ligation at t=4 h. The observed mass (MWobs. 18261 Da) is in excellent agreement with the calculated mass (MWcalc. 18261 Da) of the double phosphorylated ligation product **9d** (Hsp90 572-732, A710U). **E:** Ligation monitoring *via* SDS-PAGE. The ligation product **9d** appears as a single band at the height of the 20 kDa marker.

Especially noteworthy in this case would be that a potential intermolecular ligation of a freshly generated arylselenoester with the unprotected N-terminal cysteine of another selenoester segment was not observed. Finally, we continued with the deselenization procedure in a one-pot manner to

maximize the potential yield and convenience of the ligation setup. To do so, we removed the radical quenching excess of free phenylselenol *via* hexane extraction and continued by adding deselenization buffer (v/v 1:1, final TCEP concentration 100-150 mM). After initial slow reaction rates and only a limited amount of deselenized product formation, we optimized this step by performing the reaction at a slightly basic pH (7.4), which led to good conversion rates (between 70-80%) and the formation of final semi-synthetic CTD products **10a-d** in 4 hours (**Figure A 13** and **Figure 29**). Interestingly, after freezing the ligation solution at -80°C and subsequent thawing and workup by 1:1 dilution with 6M Gdn.HCl no or only traces of Selenocysteine containing ligation product could be observed in the collected peaks after RP-HPLC purification. However, careful reaction monitoring was needed since, in the presence of high amounts of TCEP (>100 mM) and after prolonged reaction times (>4 hours), we were able to detect the formation of a -32 Da side product after a few hours which would be a strong indicator for the undesired desulfurization of the most likely accessible N-terminal free cysteine of the ligation product. With final yields ranging between 40 and 50 % based on the recombinant starting material as a limiting factor, we consider this a successful, highly optimized ligation strategy for efficiently generating site-selectively modified Hsp90 CTD variants.

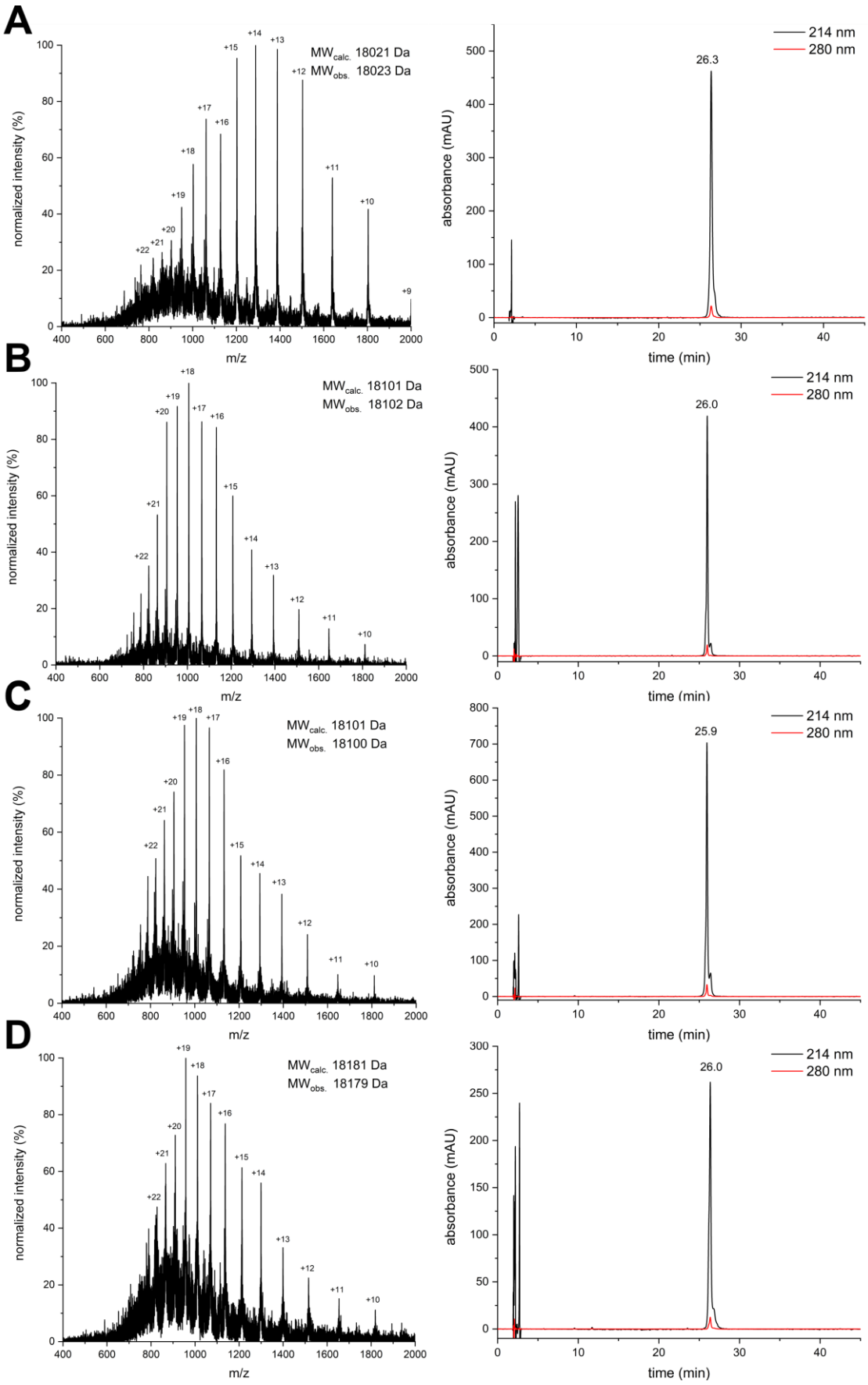

Figure 29 Characterization of semi-synthetic Hsp90 CTD variants **10a-d** *via* ESI-MS (left panel) and RP-HPLC (right panel). **A:** Characterization of unmodified CTD **10a**. **B:** Characterization of phosphothreonine CTD variants **10b** (pThr725). **C:** Characterization of phosphoserine CTD **10c** (pSer726). **D:** Characterization of double phosphorylated CTD **10d** (pThr725/pSer726).

3.2 PTM impact on secondary structure elements and thermal unfolding

After successfully generating the four different CTD variants, we were first interested in the secondary structure exhibited by the semisynthetic proteins. Attempting to dissolve the lyophilized protein in the desired refolding buffer directly failed due to the high hydrophobicity of the domain, so we initially dissolved them in high chaotropic salt buffer (6M Gdn.HCl, TBS pH 7.5) before refolding *via* dialysis into the desired buffer for characterization. To determine successful refolding and the potential impact of the introduced PTMs on secondary structure elements, we characterized our semi-synthetic variants *via* circular dichroism (CD) spectroscopy. CD spectra of the unmodified variant, as well as the single and double modified variants, showed the characteristic minima around 207 and 225 nm, which are well in accordance with the literature and the CD spectra we recorded of the recombinantly expressed WT CTD **(Figure A 14**).[127] This indicates the well-defined primary alpha-helical structure of the CTD variants and, therefore, suggests successful refolding; however, it is difficult to determine from the CD experiments whether the Hsp90 CTD variants also form dimers during refolding. Interestingly, we observed slightly more pronounced minima at 207 nm and 225 nm for the double-modified variant, compared to the unmodified variant at the same concentration, indicating a more defined alpha-helical secondary structure upon introducing the two phosphate groups (**Figure 30**).

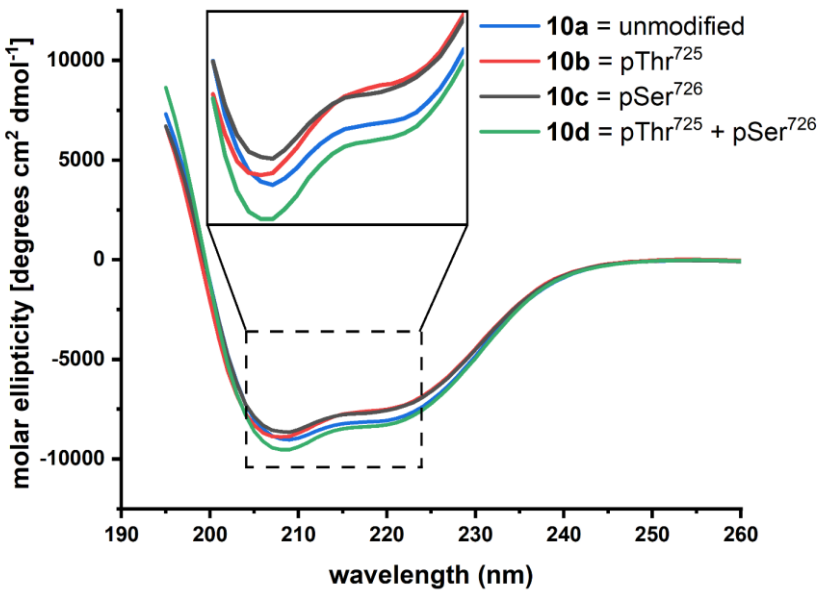

Figure 30 Structural characterization of CTD variants *via* Circular dichroism spectroscopy. Zoom was inserted on the alpha-helical region between 205 and 220 nm. Circular dichroism spectra of CTD analogs were measured in 50 mM KPO4 buffer at pH 7.6 (0,2 mg/ml).

The impact of the introduced PTMs on the secondary structure of the CTD variants sparked our interest if we could also observe and influence the thermal stability of the semi-synthetic proteins. Therefore, we aimed to examine the thermal stabilities of the semi-synthetic CTD variants using Nano differential scanning fluorimetry (NanoDSF) (**Figure 31**). This method uses intrinsic tryptophan or tyrosine fluorescence to monitor protein unfolding. Since the fluorescence intensity and the fluorescence maximum depend on the local surroundings of Trp and Tyr residues, structural changes and protein unfolding can be monitored through the ratio of fluorescence intensities at 350 nm and 330 nm. This would theoretically allow for determining the Protein Melting Temperature Tm defined as the midpoint of the unfolding transition. Additionally, a qualitative comparison of the melting curves can indicate the difference in thermal stability caused by the PTMs. Previous analysis of Hsp90 CTD using thermofluor assays and CD showed that it is most thermally stable at mild neutral conditions. Therefore, all variants were screened in NaPi Buffer at pH 7.4.[127] Unfortunately, signal intensity was relatively low, presumably because only one Tryptophan residue was present in the CTD (Y689), resulting in melting curves, making precise Tm determination impossible. However, a significant qualitative difference in the melting behavior of the double-modified variant 10d compared to the single-modified and unmodified control is observable (**Figure 31**, dotted lines). This indicates a lowering of the Tm value of ~3-5°C caused by introducing the two phosphorylations and an overall lower thermal stability of the double-modified variant. In addition to the left-shifted melting curve of **10d** compared to the other variants, the F350/F330 ratio curve for **10d** reaches a maximum at about 70 degrees and then decreases at higher

temperatures. We speculate that this could indicate unfolding followed by aggregation, which would alter the environment of the Trp/Tyr residues. Combined with the CD data, these results suggest that the doubly modified Hsp90 CTD has a slightly more defined alpha-helical structure than the other variants but with lower thermal stability.

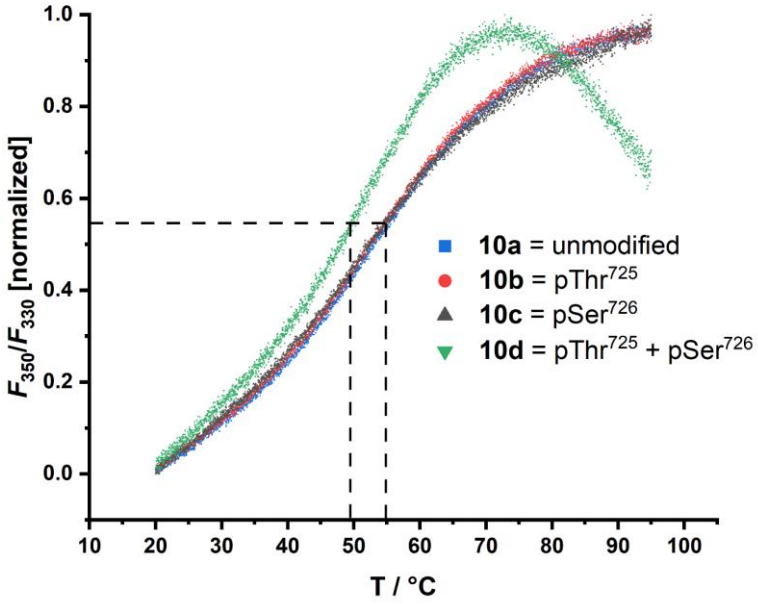

Figure 31 Nano differential scanning fluorimetry of semi-synthetic Hsp90 CTD with an indicated shift towards lower thermal stability of the doubly phosphorylated CTD compared to the unmodified and singly modified variants

3.3 Impact of phosphorylations on CTD chaperone activity *in vitro*

After observing the altered secondary structure and thermal stability of our double phosphorylated CTD variant, we aimed to investigate the effect of the PTMs on the chaperoning ability of the CTD. Previous studies have shown that the isolated CTD of Hsp90 exhibits chaperoning capabilities in various *in vitro* aggregation assays with model substrates.[127,123,128] We examine the anti-aggregation function of our CTD variants using an ADH and Insulin aggregation assay as previously used to study novel CTD dimerization inhibitors as the chaperone activity correlated with the ability of the CTD to form dimers in previous studies. Both the aggregation of ADH and Insulin can be monitored by the detection of light absorption at 360 nm, while ADH aggregation is solely thermally induced by increasing the temperature to 55°C, the aggregation of insulin is based on structural disintegration through internal disulfide reduction *via* the addition of DTT. In both assays, all of our Hsp90 variants show a strong chaperone activity *via* the depletion of aggregates to about ~60% using ADH and ~ 55% using Insulin when comparing endpoint mean values (**Figure 32**). Since there are no significant differences between the activity of our variants, we conclude that the introduced PTMs do not impact the chaperone activity, indicating successful active dimer formation of all your Hsp90 variants.

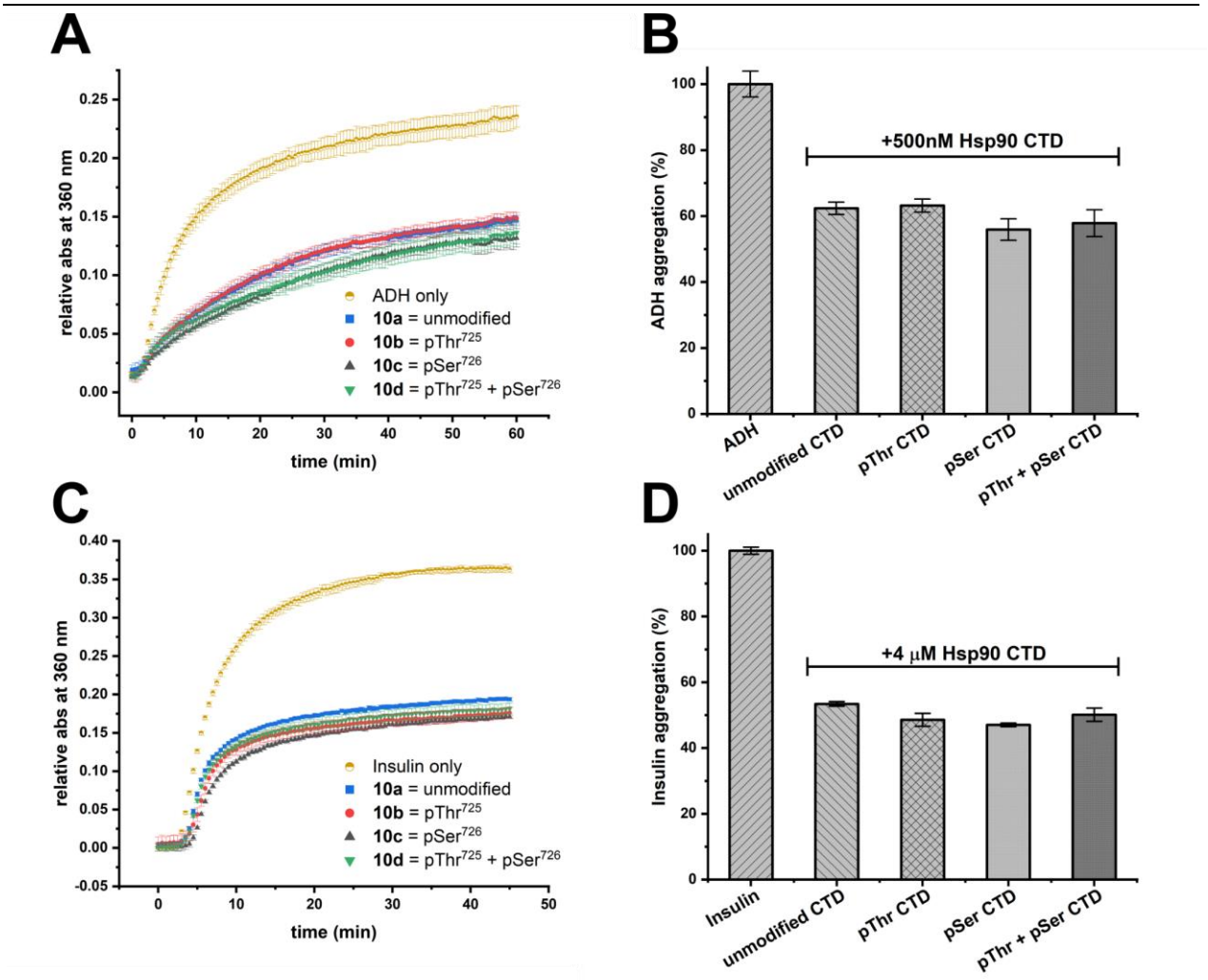

Figure 32 Assessing in vitro chaperone activity of phosphorylated and unmodified Hsp90 CTD variants **10a-d** using ADH and Insulin aggregation assay. **A:** ADH aggregation (3.1 µM) was induced by incubation at 55 °C in the presence and absence of 500 nM **10a-d** and monitored by measuring the absorbance at 360 nm over 60 minutes. Displayed are mean curves of a triplicate measurement. **B:** Endpoint mean values (60 min) of ADH aggregation without the addition of CTD were set to 100% and compared with mean values at 60 min from samples with additional Hsp90 CTD incubation. **C:** Insulin aggregation (40 µM) was induced by adding 20 mM DTT in the presence and absence of 4 µM **10a-d** and monitored by measuring the absorbance at 360 nm over 45 min at 37°C. **D:** Endpoint mean values (45 min) of Insulin aggregation (40 µM) without the addition of CTD were set to 100% and compared with mean values at 45 min from samples with additional Hsp90 CTD incubation (4 µM).

3.4 Assessment of PTM impact on the oligomeric status *via* cross-linking and SEC characterization

One of the major tasks specifically attributed to the CTD is its facilitation of Hsp90 dimerization, an essential property since previous studies have shown that the protein only functions properly as an active homodimer.[129] Furthermore, it has been demonstrated that the CTD itself can dimerize and form tetramers and higher-order oligomers. Therefore, we aimed to assess phosphorylation's effect on its oligomeric state by using amine-reactive covalent crosslinking *via* BS^3, as previously described.[123] The amine-reactive crosslinker allows for capturing the protein's dominant oligomeric state, which is presented by the slower migrating species compared to the control, where no crosslinker was added (**Figure 33**).[130] Although faint bands that correspond to the monomeric (~23 kDa) species as well as more clear bands that could represent the dimeric (~46 kDa) CTD were visible in two of our attempted assays, higher-order oligomers (>90 kDa) where strongly visible for all variants suggestions that either the used dimerization conditions including the applied crosslinker or the process of membrane transfer for western blotting strongly supported this higher-order oligomerization of the CTD. Furthermore, we could not generate reproducible results, making a relative quantification of the present monomer to dimer species for the different CTD variants impossible (**Figure A 12**). Many factors might cause this since the procedure is highly sensitive and error-prone due to the multiple steps with intrinsic error margins, including protein refolding and concentration determination, quantitative western blot transfer and antibody binding, as well as variants in chemiluminescent western blotting substrate signal intensity.

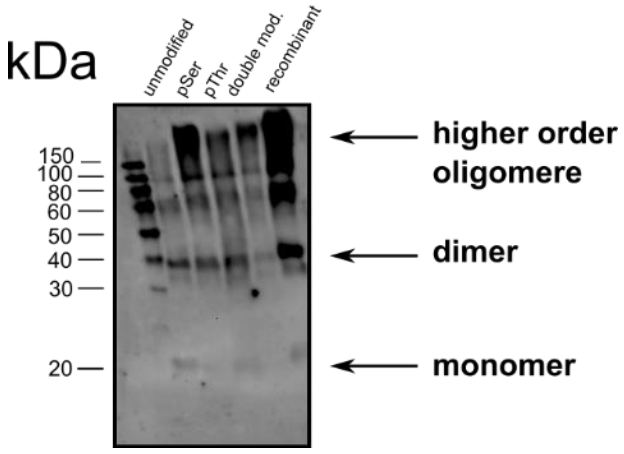

Figure 33 Amine reactive crosslinking and Western blotting for assessing the oligomeric state of semi-synthetic Hsp90 CTD variants **10a-d**.

We decided that this method might not be suitable for accessing the presumably minor differences of oligomeric state difference of the different CTD variants and aimed for an estimation of the molecular weight after refolding using size exclusion chromatography (SEC) as was reported previously.[127] Unfortunately, the observed retention profiles indicated only the predominant presence of higher-order oligomers (>100 kDa), and therefore, no prediction about the specific monomeric to dimeric, oligomeric state ratio under native circumstances could be made (**Figure A 16**).

4. Conclusion

Here, we present the first semisynthetic approach for deciphering the Hsp90 chaperone code by efficiently generating homogenous side-selective modified human Hsp90 C-terminal domain (CTD). Novel Expressed Protein Selenoester ligation methodology allowed access to four different variants of Hsp90 CTD with two different phosphorylation sites close to the C-terminus of the domain. Selenochemistry efficiently enabled us to generate semisynthetic proteins without additional thiol protection chemistry. Using an *in-situ* approach for the selenoester generation and a one-pot ligation and selective deselenization strategy in the presence of three natively occurring cysteines enabled the efficient generation of four different Hsp90 CTD variants in multi-milligram amounts. Our characterization of the semi-synthetic proteins shows that, after successful refolding, they resemble the overall native secondary structure of WT recombinant expressed Hsp90 CTD. However, the introduced phosphorylation impacts the alpha-helical content of the C-terminal domain. Furthermore, the presence of both modifications showed reduced thermal stability of the CTD, hinting at an overall destabilizing effect of the introduced phosphorylations. Our assessment of the chaperone activity of the CTD showed that the modifications do not impact the anti-aggregation activity of the CTD in an *in vitro* ADH and Insulin aggregation assay. Although a more detailed study of the oligomeric state *via* crosslinking and western blotting or through molecular weight determination *via* SEC was not successful, these results suggest that the dimerization of the CTD should not be significantly impaired since studies of dimerization inhibitors have shown to reduce the chaperone activity of Hsp90 CTD severely.[131]

CHAPTER THREE

Semi-synthesis of full-length Hsp90α

1. Introduction

This chapter describes the semi-synthesis of full-length human Hsp90α utilizing a sequential *C*-to-*N* ligation strategy based on the successful generation of semi-synthetic human Hsp90α CTD, as discussed in the previous chapter. We aimed to utilize other natively occurring cysteines of the MD for assembling full-length Hsp90 while introducing other PTMs with indicated functional relevance based on prior studies. Therefore, we envisioned continuing with the ligation of synthetic peptide segment 529-571 to semi-synthetic CTD (572-732) utilizing the natively occurring cysteine at the N-terminus of the CTD (Cys572). This would enable us to incorporate acetylated Lys at position 546 *via* SPPS. In a previous study by Yang *et al.*, it was shown that this residue is one of six Lys residues in the Hsp90 MD, which get hyper-acetylated among the knockdown of histone deacetylase 6 (HDAC 6). Using site-specific Glu mutants as acetylation-mimetic, they were able to indicate that acetylation of these specific residues resulted in diminished binding affinities of Hsp90 towards ATP, co-chaperone p23, and client protein binding. Furthermore, using the same acetylation-mimicking mutants, they linked Hsp90α MD lysin acetylation with cellular location, indicating that these PTMs promote export and extra-cellular location of Hsp90α compared to acetylation-deficient mutants with Arg instead of Lys residues at the corresponding position.[132] However, the specific effects of single lysin modifications without the use of mimetics have not yet been revealed, nor has the potential cross-talk between those and other PTMs, e.g., the phosphorylation present in the CTD.

Therefore, our synthesis strategy would not only involve a new synthetic segment of reasonable length (42aa) that would enable us to introduce acetylated lysine at position 546 but would also provide another ligation handle due to the native occurrence of the Cys529. We aimed for the synthesis of a peptide hydrazide as a precursor for the C-terminal thioester of segment 529-571 while orthogonally Acm protecting the thiol group of its N-terminus to avoid intramolecular ligation or ligation with another synthetic 529-571 segment. Finally, after successful NCL of the synthetic segment 529-571 to semi-synthetic 572-732 (Hsp90 CTD) and removal of the N-terminal thiol Acm protection group, full-length Hsp90 can be generated *via* another EPL of Hsp90 1-528 to the novel 529-732 protein. Therefore, the last required segment 1-528 will be generated from an intein-fusion construct to yield Hsp90 1-528-SR bearing C-terminal α-thioester necessary for ligation (**Figure 34**).

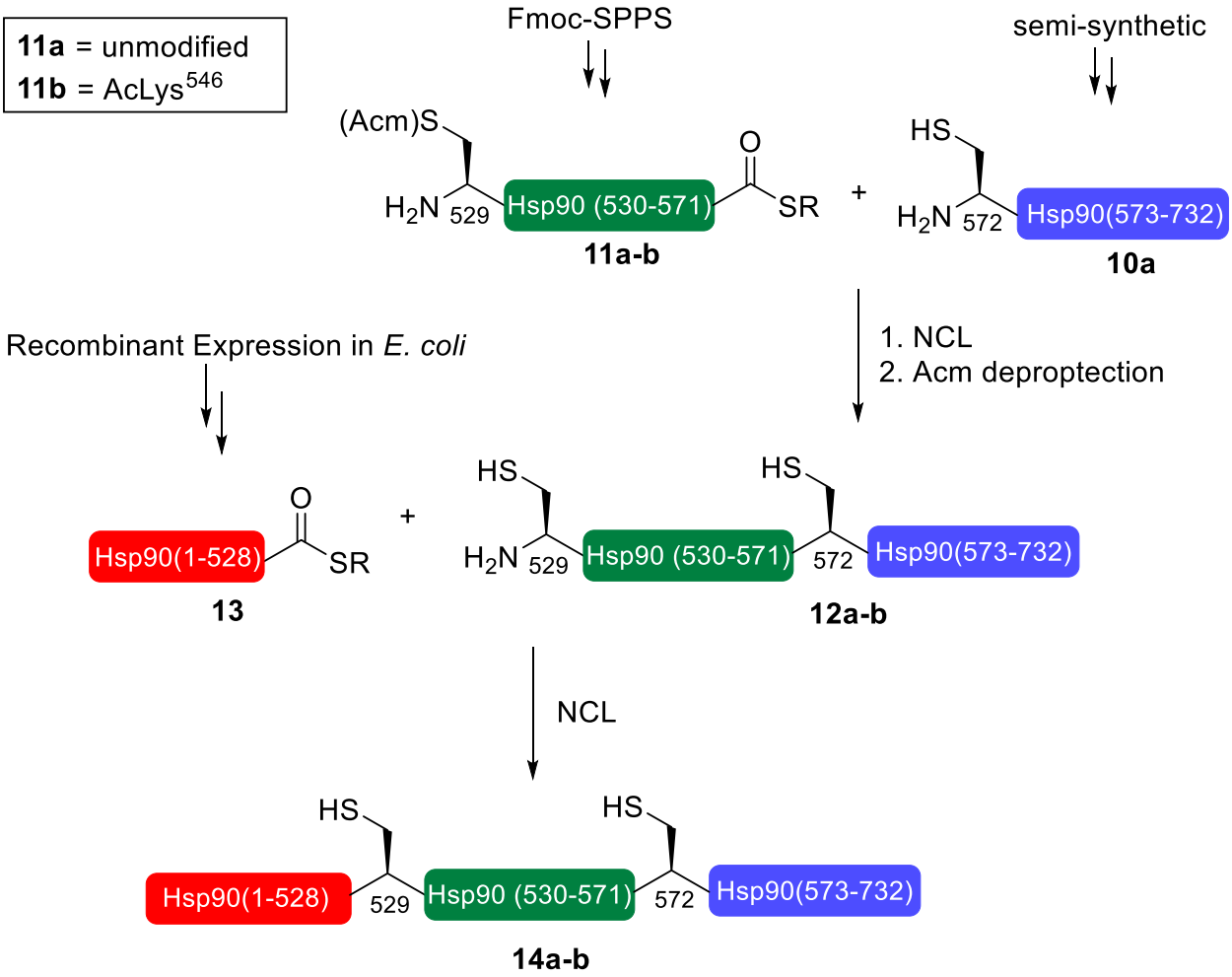

Figure 34 Schematic overview of the generation of site selectively modified full-length human Hsp90α using a sequential *C*-to-*N* terminal EPSL and EPL strategy**.**

2. Materials and Methods

2.1 Solid phase peptide synthesis of Hsp90 529-571-SR peptides

The synthesis of Hsp90 MD peptides 529-571-SR **11a** and **11b** was performed on hydrazine-modified 2-chlorotrityl resin according to the protocol of Liu *et al*..[109] After the manual coupling of the first amino acid, the loading was determined, and the residual amino acids were coupled using microwave-assisted automated peptide synthesis on a Liberty Prime synthesizer (CEM). For the activation of amino acids, 5 equivalents of Oxyma and 10 equivalents of DIC were used, and coupling was performed at 90°C for 2 minutes. Fmoc-deprotection was performed by the addition of 20% piperidine in DMF at 90°C for 1 minute to the resin. Besides Standard Fmoc- building blocks, Fmoc-Lys(Ac)-OH was used for introducing the acetylation in modified peptide **11b** and Fmoc-Cys(Acm)-OH for the coupling of the N-terminal Cys on both peptides. Furthermore, the following pseudo-proline dipeptides were used to decrease probabilities of on-resin aggregation during the automated synthesis: Thr-Gly, Arg-Thr, Val-Ser, and Asp Thr. After the final Fmoc-deprotection step, the resin was removed from the synthesizer, washed with DMF and DMC, and dried under vacuum overnight before cleavage was performed using TFA/TIPS/H_2O 95:2.5:2.5 for 2.5 h. The cleaved peptides were precipitated using cold ether, pelleted *via* centrifugation (4000 g, 5 min), dissolved in 50:50 ACN/H_2O, and dried via lyophilization, which yielded the crude peptide hydrazides. Subsequently, the lyophilized peptides hydrazides were dissolved in 6M Gdn.HCl NaPi pH 3, cooled to -15°C and activated by adding 0.5M NaNO2 for 15 min under stirring. Following that, thioester conversion was achieved by the addition of 100 eq. mesna and pH adjustment to 6.8. After 20 min incubation at rt, the reaction was diluted 1:1 with 6M Gdn.HCl in H_20 and directly purified on a preparative C18 column (Kromasil) using a linear gradient from 20-60 % of solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 40 min at a flow rate of 9 mL/min. Peptide fractions were pooled and lyophilized, which produced the desired peptide **11a** at a yield of mg (16%) and **11b** at a yield of 16 mg (20%) based on the crude peptide yield.

2.2 Ligation of unmodified Hsp90 529-571-SR to semi-synthetic CTD (572-732)

For the generation of the unmodified semi-synthetic Hsp90 529-732 segment (**12a**), lyophilized unmodified semi-synthetic Hsp90 CTD (572-732) was dissolved in freshly prepared and degassed 6M Gdn.HCl, 100 mM TCEP, 100 mM MPAA, 150 mM NaPi pH 7.2 buffer to a final concentration of 1mM. The mixture was added to 1.5 eq of lyophilized unmodified synthetic peptide

thioester **11a**, the pH readjusted to 7.2, and the reaction was incubated at rt for 2 h. The reaction was monitored by LC-MS and SDS-PAGE and, after completion, quenched through 1:1 dilution with 20% B-Me in 6M Gdn.HCl in H_20 for 20 minutes. Product purification was performed on semi-preparative C4 column (Kromasil) using a linear gradient from 20-55 % of solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 40 min at a flow rate of 5 mL/min at 60°C. Product fractions were identified *via* ESI-MS, pooled, and lyophilized, which generated the desired segment at a yield of 23%.

2.3 Acm deprotection of unmodified semi-synthetic Hsp90 529-732

For N-terminal Cys(Acm) deprotection, purified and lyophilized **12a** was dissolved in freshly degassed 85% Acidic acid / 15% H_20 to a final concentration of 1 mM. Subsequently, 140 equivalents of AgOAc were added to the solution, and the reaction was incubated at rt under vigorous shaking. After completion (1 h 30 min), the deprotection was monitored via LC-MS and quenched via dilution with an equal volume of 6 M Gdn.HCl, 1 M DTT in H_20 for 20 min. Purification was performed on a semi-preparative C4 column (Kromasil) using a linear gradient from 5-60% of solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 20 min at a flow rate of 5 mL/min at 60°C. Peak fractions were pooled and lyophilized, and the yield was determined to be ~90%.

2.4 Expression intein-cleavage and purification of Hsp90 1-528-SR

For the generation of the recombinant thioester segment 1-528-SR (13), a synthetic oligonucleotide comprising human Hsp90alpha 1-528-Mxe GyrA-His7-CBD was ordered from Eurofins Genomics in a pET-21a (+) vector and heat shock transformed into chemically competent *E. coli* strains XL1-Blue, BL21 (DE3 Gold) and Rosetta 2 (DE3) from Invitrogen. Large-scale expression was performed using the E. coli Rosetta 2(DE3) strain in 2YT medium (16 g L-1 tryptone, 10 g L-1 yeast extract, 5 g L-1 NaCl) containing 100 µg mL-1 ampicillin. Overnight cultures were diluted to a starting OD of ~ 0.2 and were grown at 37 °C with shaking at 170 rpm to an OD of 0.7-0.8 before induction of expression with 1 mM IPTG. After 4 h at 37 °C (170 rpm), cells were harvested *via* centrifugation (12000 g, 20 min), and cell pellets were stored at -80 °C before further workup. After thawing, 1 × TBS, pH 7.8 buffer was added to the cell pellet (50 mL buffer / 1 L expression pellet). The solution was homogenized using an ultra turax before cell lysis *via* sonification on ice (60% amplitude, 15 sec ON, 45 sec OFF time, total ON time 10 min). After clearing the supernatant *via* centrifugation (21000 rpm, 53000 g, 45 min at 4 °C), the collected inclusion bodies were solubilized by incubation in 6M Gdn.HCl, TBS buffer at pH 7.5 while stirring overnight at rt. After

removal of insoluble cell debris (21000 rpm, 53000 g, 45 min), the solution was loaded on Ni-NTA resin slurry (5mL resin slurry/ 2L culture) in syringes equipped with polypropylene membranes and incubated for 1 h30 min at a rotator at 35 rpm. The loading solution was discarded, and the resin was washed two times with 6M Gdn.HCl, 25 mM Imidazole TBS buffer at pH 7.5 before the elution of the target protein using 6M Gdn.HCl, 500 mM Imidazole TBS buff-er pH 7.5. The eluted protein solution was dialyzed against 2M Urea, TBS, pH 7.5 overnight at 4°C, and intein cleavage was performed by the addition of 500 mM mesna at pH 7.4. The cleavage solution was incubated at rt and stirred overnight before being quenched by adding 25 mM TCEP for 15 min, followed by the 1:1 dilution with 6M Gdn.HCl in H_2O. Purification was performed on a C4 preparative column (Kromasil) using a linear gradient from 30-65% of solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 40 min at a flow rate of 9 mL/min at 60°C. Peak fractions were pooled and lyophilized, which yielded the desired protein **13** at 2 mg/L culture.

2.5 Ligation of recombinant Hsp90 1-528-SR to semi-synthetic 529-732

Ligation of recombinant thioester segment **13** to semi-synthetic segment **12a** was performed by dissolving segment 529-732 (1.5 equi.) in freshly prepared and degassed 6M Gdn.HCl, 50 mM TCEP, 200 mM NaPi pH 7.2 buffer followed by the addition of segment 1-528-SR (0.5mM final concentration). The pH was readjusted to 7.2, and the reaction was incubated under argon at rt for 3 h. Reaction monitoring was performed *via* LC-MS, SDS-PAGE, and analytical HPLC, and the reaction was quenched for 15 min by adding an equal volume of 20% β-Me, 6M Gdn.HCl in H_2O. Purification was performed on an analytical C4 column (C4 Kromasil) and a linear gradient of 35-65% of solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 30 min at a flow rate of 1 mL/min at 60°C. Peaks were manually collected, product fractions combined, lyophilized, and the yield was determined to be ~20%.

3. Results and Discussion

3.1 Solid phase peptide synthesis of Hsp90 529-571-SR peptides

When designing a semi-synthetic route for accessing full-length Hsp90, the natural occurrence of native cysteines at the N-terminus of the CTD immediately offered an opportunity. Therefore, the middle domain (MD) peptide comprising aa 529-571 (**11**) was chosen since it would directly extend the route developed for semi-synthetic CTD variants (see Chapter 2) through ligation to the free N-terminal Cys 572. This would allow the introduction of various PTMs into the MD, including the reported modifications of lysine residue 546, where acetylation and succinylation have been identified and connected to the cellular location of Hsp90.[63,64] Furthermore, the naturally occurring Cys at position 529 would mark the next ligation site to attach the remaining NTD-MD residues 1-528 without surrogates and post-ligation desulfurization/deselenization procedures. However, the simultaneous presence of an N-terminal Cys residue and an acyl donor after C-terminal hydrazide activation and thioester conversion on the same peptide segment requires an N-terminal thiol-protecting group to prevent cyclization and/or oligomerization. In this case, we decided to use acetamidomethyl (Acm) as a suitable thiol-protecting group since it can withstand the hydrazide activation and ligation conditions and can readily be introduced as a building block.[134] Furthermore, this strategy is also suitable for the generation of the peptide *via* automated microwave-assisted synthesis, which can be considered as another benefit considering the length of the sequence (43 AS). Therefore, we opted for using hydrazine-modified 2-CTC generating the C-terminal peptide hydrazide as thioester precursor through microwave-assisted SPPS when and thereby producing the unmodified control peptide 529-571(**11a**) as well as the acetylated variant 529-571 (**11b**).

We were delighted to see the unmodified peptide successfully synthesized and converted to thioesters through this methodology, as demonstrated by analytical data of the Hsp90 peptide 529-571-SR (**Figure 34**), which showed satisfactory final yields of 21 mg (16%) for the unmodified variant.

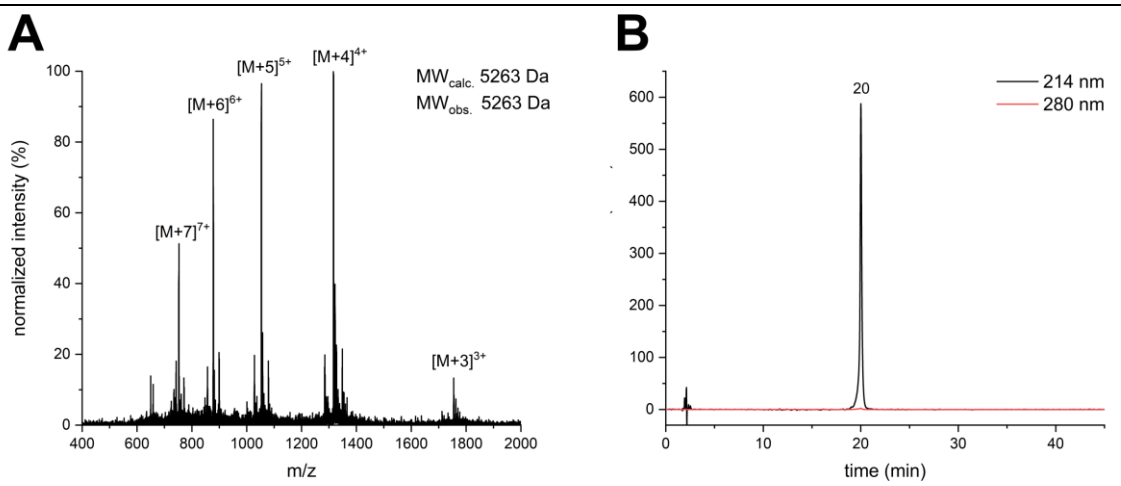

Figure 35 A: ESI-MS characterization of unmodified Hsp90 529-571-SR mesna thioester peptide. **B:** RP-HPLC characterization of unmodified Hsp90 529-571-SR mesna thioester peptide.

Similarly, we aimed to generate the AcLys(546) variant of this peptide by applying the established synthesis protocol only differing in exchanging the standard Fmoc-Lys(Boc)-OH building block in position 546 to the ε-acetylamino-lysine variant. We successfully generated the modified peptide thioester variant 529-571-SR (AcLys546) with a final yield of 20% based on the crude peptide before thioester conversion and purification (**Figure 36**). For both peptides, only minor side products with mass differences of +/−128 Da were observed, which we believe to originate from Glu (exact mass 129) and/or Lys (exact mass 128) deletions as well as addition products as those are frequently occurring residues in the peptide, especially in the region of aa 552-571. An optimized synthesis protocol with capping steps in the critical regions where most of the Lys and Glu residues are coupled (aa 552-571) could further improve the quality of the final peptide, although the purity of the final peptides was good enough for the first ligations attempt and therefore, we decided to continue with the ligation to semi-synthetic 572-732.

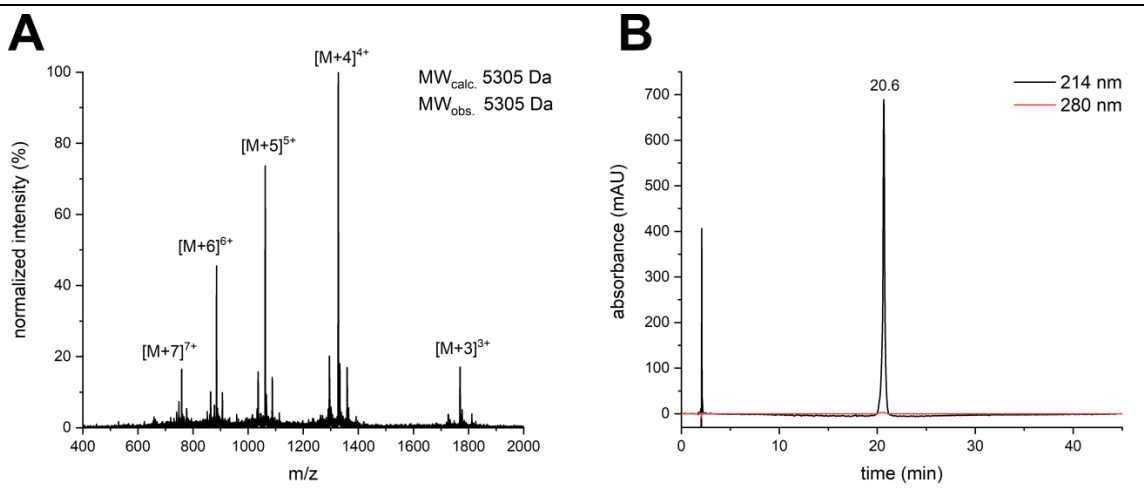

Figure 36 A: ESI-MS characterization of Hsp90 529-571-SR (AcLys546) mesna thioester peptide. **B:** RP-HPLC characterization of Hsp90 529-571-SR(AcLys546) mesna thioester peptide.

3.2 Generation of unmodified semi-synthetic Hsp90 529-732 MD-CTD

For the generation of the semi-synthetic Hsp90 (aa529-732) segment **12a**, we aimed for an NCL-based strategy using an aryl thiol catalyst since a natively occurring cysteine at position 572 was used as a ligation junction, and, therefore, no post-ligation desulfurization was required (**Figure 37**).

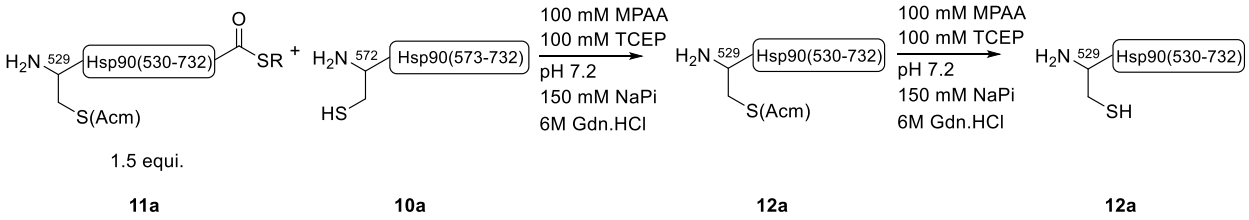

Figure 37 Schematic showing the reaction conditions for the generation of semi-synthetic Hsp90 529-732 *via* the ligation of unmodified synthetic peptide **11a** to semi-synthetic segment **10a** followed by Acm deprotection of the Cys529 thiol group.

Using the unmodified peptide 529-571-SR **11** and the unmodified semi-synthetic CTD 572-732 **10a**, we observed product formation through LC-MS and SDS-PAGE monitoring, as depicted in **Figure 38**.

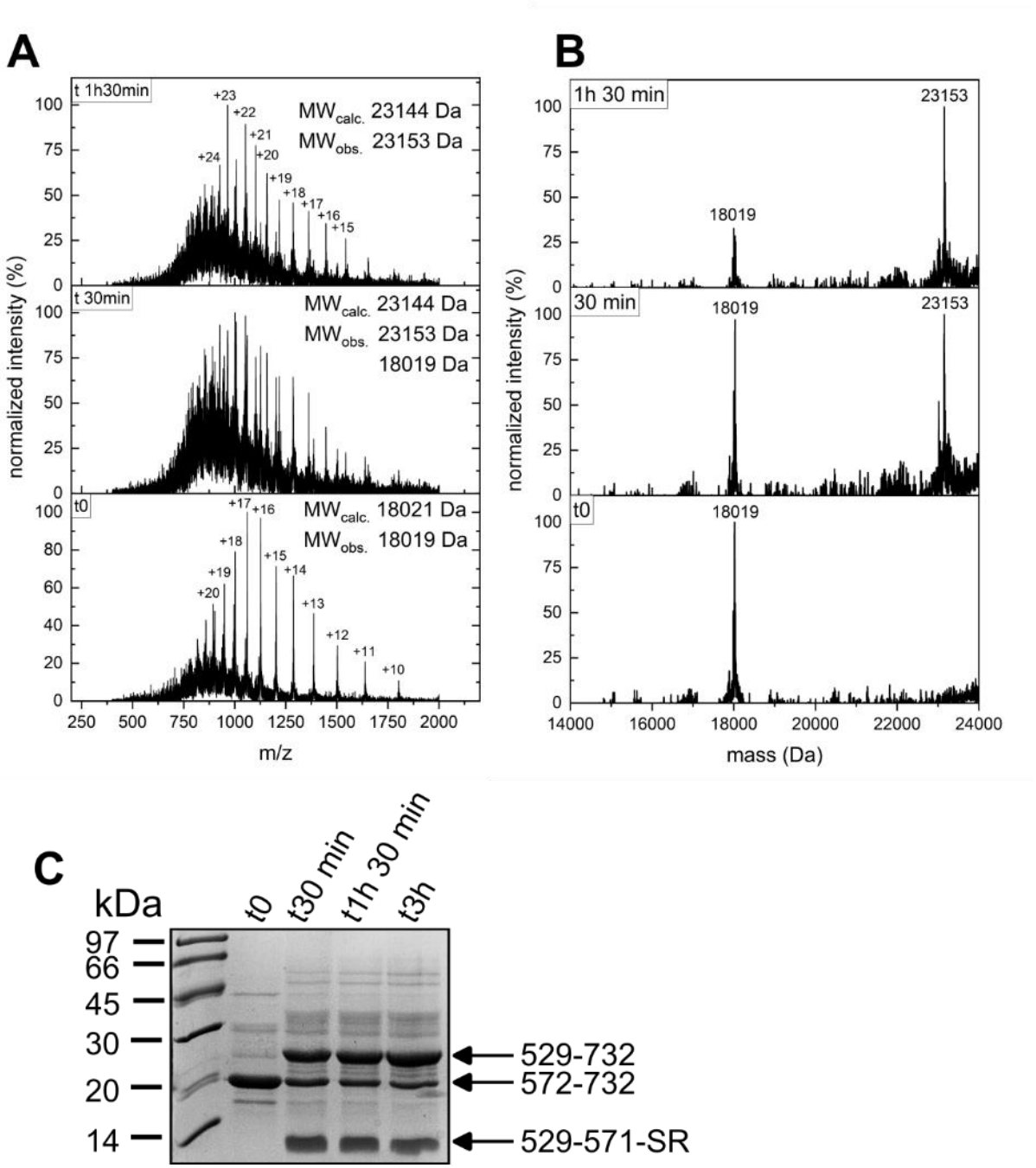

Figure 38 A: LC-MS Monitoring of ligation of synthetic Hsp90 520-571-SR peptide to semi-synthetic Hsp90 CTD (572-732). **B:** Deconvoluted MS spectra of ligation of synthetic Hsp90 520-571-SR peptide to semi-synthetic Hsp90 CTD (572-732). **C:** Ligation monitoring of synthetic Hsp90 520-571-SR peptide to semi-synthetic Hsp90 CTD (572-732*) via* SDS-PAGE. Semi-synthetic unmodified CTD (572 -732) is visible as a strong band at an apparent MW of ~20 kDa, and the ligation product 529-732 at an apparent MW of ~25 kDa. The faint band at ~17 kDa can be attributed to small amounts of hydrolyzed 572-710 recombinant protein, which was not separatable after assembling semi-synthetic CTD *via* EPSL.

Notably, we observed a decreasing signal-to-noise ratio, which might originate from an increase in hydrophobicity after ligation and, therefore, lead to a decrease in the ionizability of the ligation product. However, no visible precipitation during the reaction or detectable side-product formation was observed. Monitoring *via* SDS-PAGE allowed the estimation of a conversion of ~80%. The remaining starting material can be partially explained by small amounts of N-terminally desulfurized semi-synthetic CTD, as this side-product was observed during post-ligation deselenization of the CTD and could not be fully separated during purification. The presumably high hydrophobicity of the ligation product, in combination with a rather shallow HPLC gradient, caused losses through low HPLC recovery. The overall low yield of ~20 % left room for improvement. Since N-terminal Acm deprotection was also required before the next ligation to segment 1-528 for full-length Hsp90 generation, we aimed for a "one-pot" post-ligation Acm deprotection. Unfortunately, we did not observe the desired conversion when using AgOAc for cleavage in a one-pot manner. However, we were delighted to observe a quantitative conversion to the deprotected product when the HPLC-purified ligation product was applied. For the final purification of the unprotected segment, we achieved good yields (~90%) by using a steep HPLC gradient and higher temperatures (60°C), which yielded the desired unprotected semi-synthetic Hsp90 segment 520-732 (**12a**) in satisfactory purity (**Figure 39**).

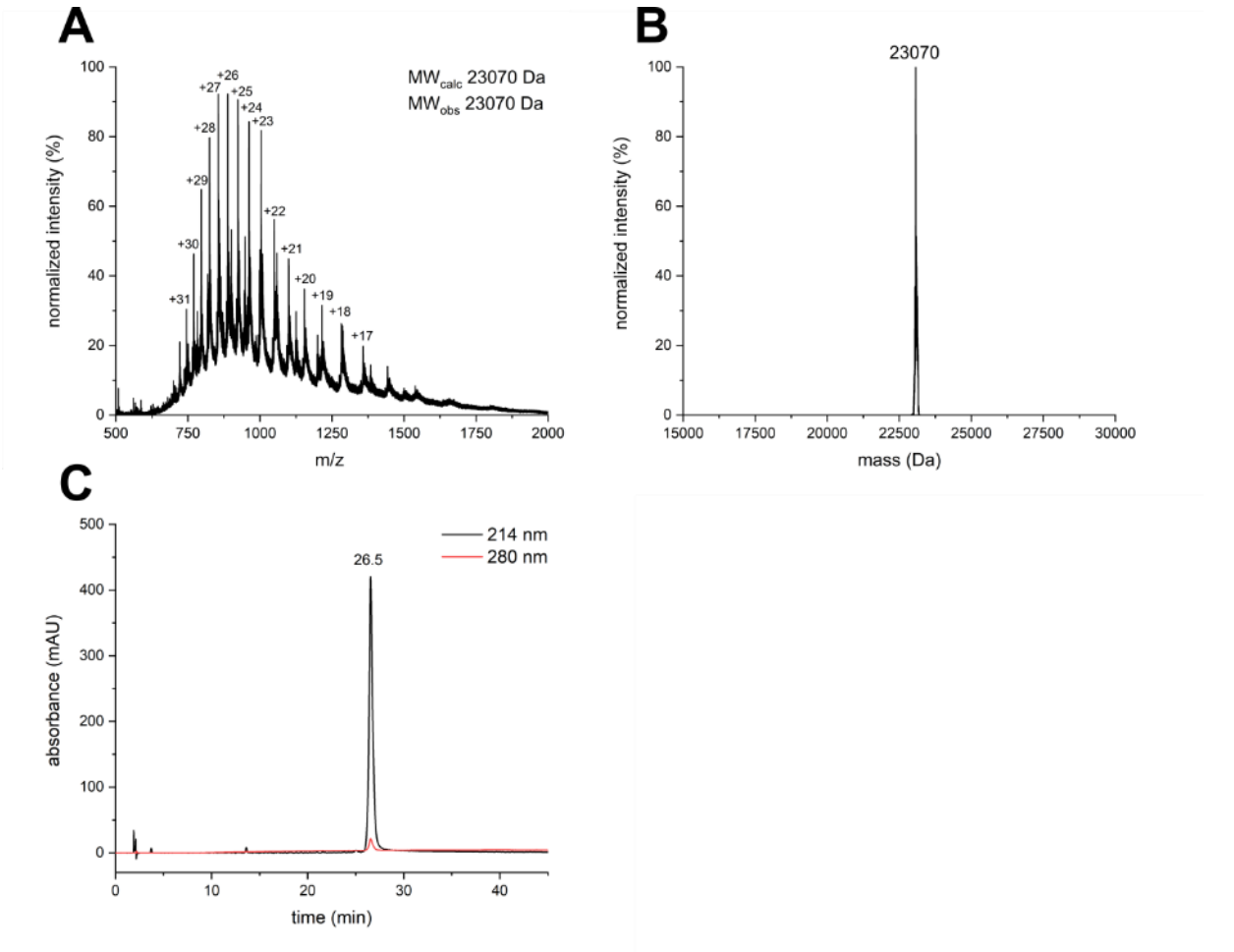

Figure 39 Characterization of semi-synthetic Hsp90 529-732. **A:** ESI-MS characterization. **B:** Deconvoluted ESI-MS spectrum. **C:** RP-HPLC analysis of Hsp90 529-732.

3.3 Expression, intein cleavage, and purification of Hsp90 1-528-SR

Having established a semi-synthetic route for accessing Hsp90 529-732, we next aimed to generate the final segment necessary to obtain full-length human Hsp90α. This N-terminal segment had to comprise residues 1-528 and required a C-terminal α-thio- or selenoester or a suitable activatable precursor to undergo ligation with Hsp90 segment 529-732 with an N-terminal cysteine residue. Therefore, we aimed for recombinant expression of Hsp90 (aa 1-528) as a C-terminal intein-fusion construct. However, when designing the designated construct, we realized that due to the large size of the N-terminal segment (~ 62 kDa) in addition to the size of the intein and the required His_6-CBD affinity tag (~ 30 kDa), this could represent a burden for the *E. coli* expression machinery and lower expression yields were expected. Nevertheless, the plasmid of the construct Hsp90 1-528-Mxe GyrA-His6-CBD was ordered in pET-21a vector and cloned into two different expression strains previously described (section 2.4). Indeed, when screening different strains and expression conditions, only a relatively low overexpression of the desired construct in insoluble inclusion bodies was observed under all conditions (**Figure A 3**). Nevertheless, *E. coli.* BL21 (Gold) strain yielded a clean expression of the protein with no significant premature intein-cleavage or other proteolytic side products, and we decided to continue with larger scale (1L) expression for optimization of the cell workup and purification using this strain (**Figure 40**).

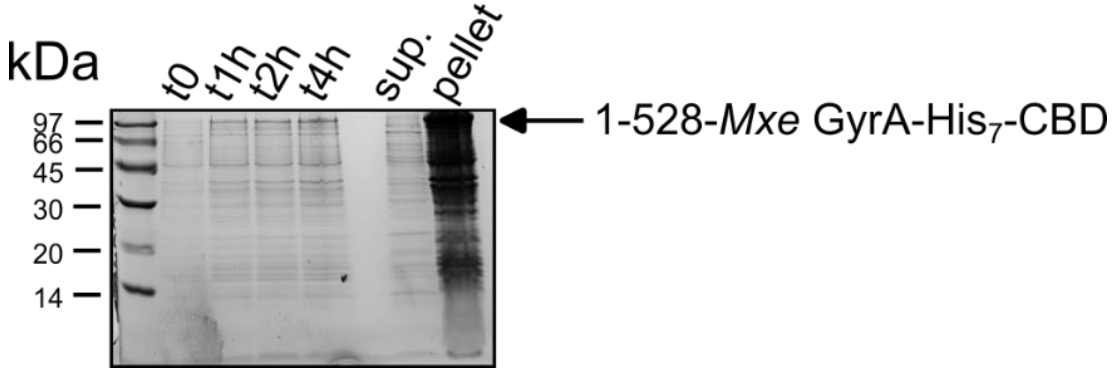

Figure 40 Expression monitoring of Hsp90 1-528-Mxe GyrA-His7-CBD. Sample time points after induction with 1mM IPTG are labelled as t0-t4 h. Cell lysate supernatant after clearance through centrifugation is labelled with sup.. Inclusion body pellet after the removal of the cell lysate supernatant is labelled as pellet. Optimized expression was performed in *E. coli.* BL21 (Gold) strain for 4 h at 37°C using 1mM IPTG for induction.

We were delighted to achieve relatively good purity of the 1-528-intein fusion protein **13** already after a first affinity-based purification using IMAC (**Figure 41**).

Figure 41 Monitoring of Ni-NTA based IMAC purification of Hsp90 1-528-Mxe GyrA-His7-CBD. L: Solubilized inclusion body (IB) pellet. FT: Supernatant of solubilized IB after incubation on Ni-NTA resin. The desired protein is indicated as a strong band at 97 kDa protein marker height eluting at ~150 mM Imidazole.

The optimization of the following step includes intein cleavage and purification of the desired Hsp90 segment 1-528-SR carrying a C-terminal thioester. Separation of the Hsp90 segment and intein presented a challenge in this case. Cleavage conditions were screened using various chaotropic salt concentrations, as well as buffer compositions (**Table 1**). Mesna cleavages were generally performed at pH 7.4, while hydrazinolysis was between pH 8-9. Furthermore, reactions were performed either in solution by adding the cleavage initiator to the dialyzed protein or after immobilization on chitin resin, which would ideally provide a first purification step. However, only minor differences were observed for all the screened conditions, and cleavage did not exceed ~60% based on densitometry quantification *via* SDS-PAGE (**Figure A 18 - Figure A 21**).

Table 1 Screened conditions for intein cleavage of Hsp90 1-528-fusion protein. Cleavage was performed after purification *via* Ni-NTA IMAC in solution or the addition of cleavage buffer to the immobilized protein on chitin resin.

Buffer	*Initiator*	*Additives*	*Reducing Agents*	
TBS	500 mM mesna			in solution
2M Urea TBS	500 mM mesna			in solution
2M Urea TBS	500 mM mesna		+ 50 mM DTT	in solution
2M Urea TBS	500 mM mesna	+0,1%Tween		on resin
2M Urea TBS	500 mM mesna	+0,1% Triton		on resin
2M Urea TBS	500 mM mesna	+10% glycerol		on resin
2M Urea TBS	500 mM mesna	+10% glycerol		in solution
3M Urea TBS	500 mM mesna			in solution
4M Urea	500 mM mesna			in solution
TBS	5% Hydrazine		+100 mM DTT	in solution
TBS	5% Hydrazine		+100 mM DTT	on resin
TBS 2M Urea	5% Hydrazine		+100 mM DTT	in solution
TBS	5% Hydrazine		+100 mM DTT	on resin

Since the protein was expressed almost exclusively in inclusion bodies, highly concentrated chaotropic salt buffers were used for solubilization, which needed to be exchanged or diluted for (partial) intein refolding and subsequent cleavage. Interestingly, this did not result in any precipitation or signs of aggregation. When observing that mesna-induced intein cleavage would not go to completion under the screened conditions, we also attempted a hydrazine-mediated cleavage, which often displayed faster reaction rates and overall higher conversions but did not significantly improve the cleavage in our case (**Figure A 21**). We decided to continue with the mesna cleavage of the construct in solution and continued the post-cleavage purification optimization (**Figure 42**).

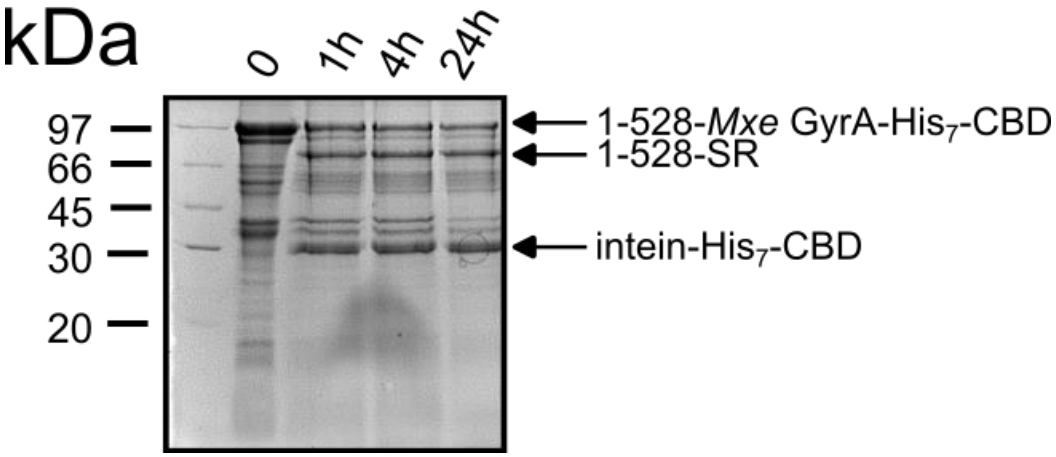

Figure 42 SDS-PAGE monitoring of mesna-induced intein cleavage. Cleavage was performed using 500 mM mesna and 2M Urea in TBS buffer, pH 7.4, for 24 h.

This constituted another hurdle as separating the desired 1-528-SR segment from the non-cleaved 1-528-intein fusion protein was highly challenging. Using a combination of different chromatographic methods, including size exclusion chromatography (SEC), affinity purification on Chitin Beads, and affinity purification on Ni-NTA, did not result in sufficient separation and/or yield (**Figure A 22** and **Figure A 23**). Finally, we decided to apply a single purification step *via* RP-HPLC where, upon optimization of sample preparation and purification conditions, we achieved the best results concerning yield (2 mg/L culture) and purity (~65%). However, analysis of the final product *via* analytical RP-HPLC as well as SDS-PAGE indicates the remaining impurities in the final product through smaller peaks with earlier retention time (between 23-29 min) and multiple bands in the lower apparent molecular weight range (~48kDa and 30 kDa). However, the most abundant impurity has an apparent molecular weight of ~97 kDa (**Figure 43**). Based on our previous observations, we can assume that the protein corresponding to the band of ~97 kDa corresponds to the uncleaved Hsp90 1-528-intein fusion protein since their apparent molecular weight match (**Figure 42**). Furthermore, our struggles concerning the optimization of the intein cleavage and the following difficulties when trying to separate the cleaved protein from the fusion protein using different affinity-based separation techniques (see above) already indicated the extremely strong affinity of the two proteins towards each other. Initially, this might not be too surprising since it is known that Hsp90 has a high dimerization affinity (under native conditions). However, in this case, strong hydrophobic interactions might be the more plausible cause due to the fact that the dimerization-initiating domain, the CTD, is missing in this construct.[129] Under the applied conditions (6M Gdn.HCl), it is to assume that the protein is unfolded and, therefore, exposing a large hydrophobic surface for unspecific interactions resulting in the carryover of residual cleaved intein segment, which would match the protein band at an apparent MW of ~30 kDa.[135] The protein corresponding to the band with an apparent MW of ~48 kDa remains unidentified, and we can only speculate that it represents a protein cleavage side product, presumably originating from

the acidic RP-HPLC purification, as there seems to be no matching protein band present before the purification (**Figure 42**).

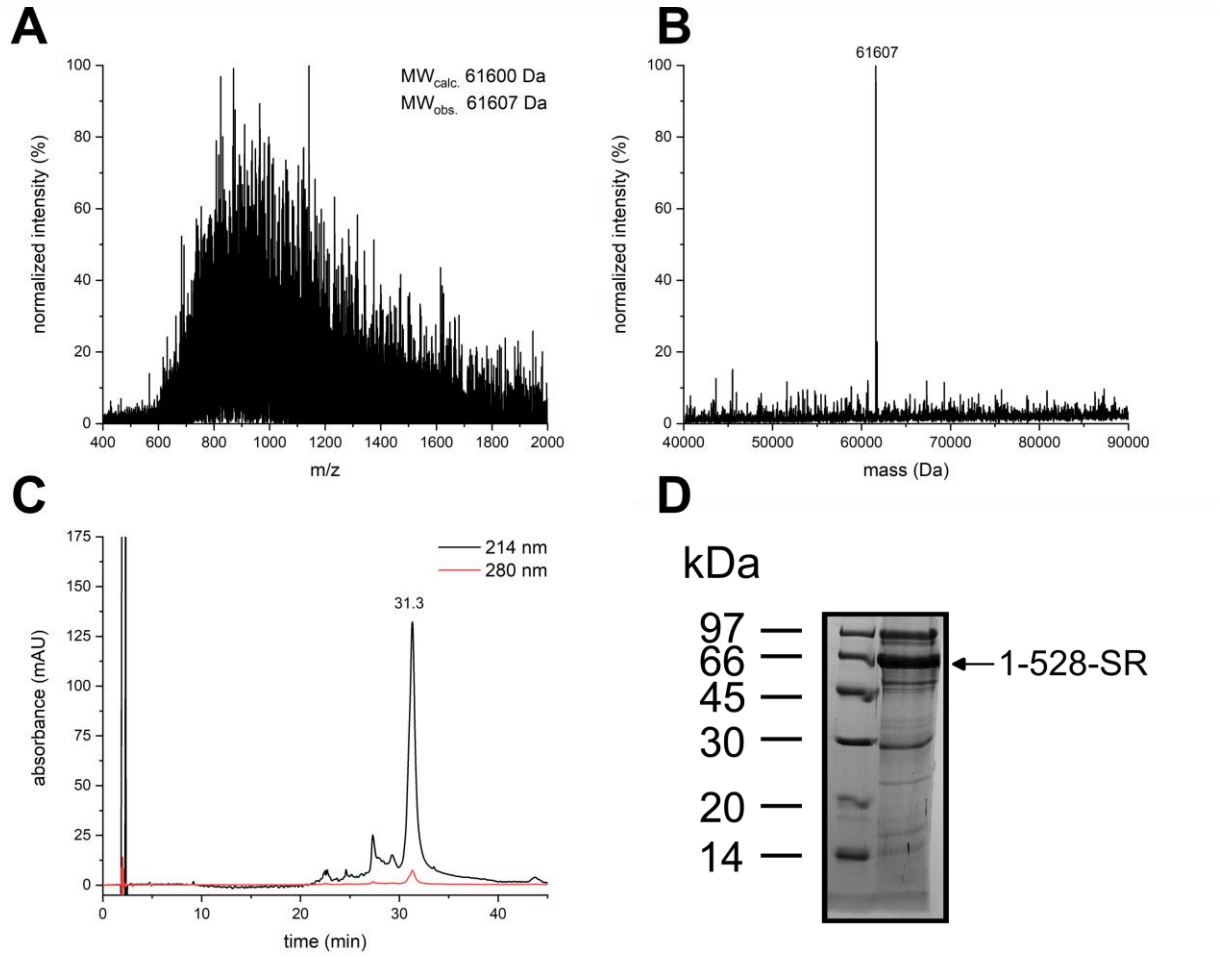

Figure 43 Characterization of recombinantly expressed Hsp90 1-528-SR. **A:** ESI-MS characterization. **B**: Deconvoluted ESI-MS spectrum. **C:** Characterization by RP-HPLC. **D:** Characterization *via* SDS-PAGE.

These strong, presumably hydrophobically based protein interactions might also prevent intein cleavage completion since they are potentially sterically hindering proper intein folding for thiolytic cleavage. Taken together, further efforts should be undertaken to increase the intein cleavage, for example, performing the reaction at higher temperatures or screening further additives to prevent the unspecific interactions between non-cleaved and cleaved protein. However, careful reaction monitoring has to be performed as the competing hydrolysis of the ester intermediate needs to be avoided, or a C-terminal carboxyl group instead of the thioester is formed.[136] A more fundamental change in the generation of segment **13** could also be approached by genetically reengineering the fusion protein to apply a trans-splicing strategy. Protein trans-splicing (PTS) utilizes split inteins made of two separate peptides: IntN for the N-terminal fragment and Int^{C} for the C-

terminal fragment. Isolated split intein fragments are unreactive, but when they are combined, they form a heterodimeric complex of Int^N and Int^C and regain their functionally folded conformation, restoring the splicing activity. Muir and co-workers reported superior kinetics of an engineered version of the naturally occurring DnaE split intein while generating recombinant C-terminal thioesters to prepare modified histone and monoclonal antibodies *via* EPL.[137] Interestingly, the same naturally occurring split intein (DnaE) has been used to generate recombinant ubiquitin al C-terminal thioester but as a fused split intein construct.[138]

3.3 Ligation of recombinant Hsp90 1-528-SR to semi-synthetic Hsp90 529-732.

After the challenging generation of the recombinant segment **13**, we continued our path to synthesizing semi-synthetic full-length Hsp90 by using small-scale test ligations of **13** to the previously generated semi-synthetic **12a** *via* NCL (**Figure 44**).

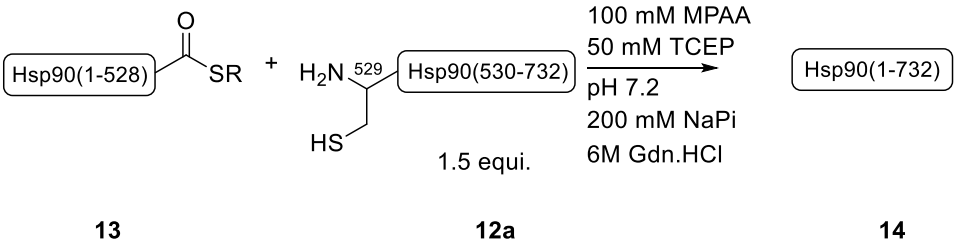

Figure 44 Schematic displaying the reaction conditions for generating full-length Hsp90 through the ligation of recombinant segment **13** to semi-synthetic segment **12a** *via* NCL.

When monitoring the reaction via SDS-PAGE and analytical RP-HPLC (**Figure 45**), we were delighted to clearly see indications of the 85 kDa ligation product formation.

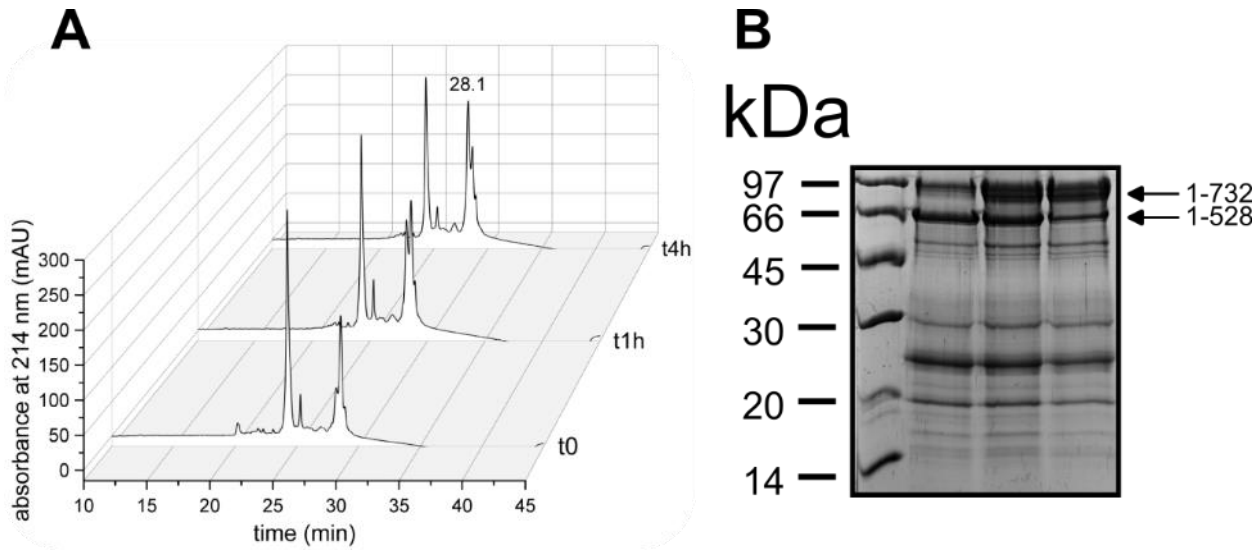

Figure 45 Monitoring of the NCL reaction between recombinant Hsp90 1-528-SR (**13**) and semi-synthetic Hsp90 529-732 (**12a**). **A:** RP-HPLC monitoring with indicated product peak emerging at a retention time of 28.1 min. The strong peak at a retention time of ~25.6 min can be assigned to semi-synthetic segment 12a, while the small sharp peak at ~26.0 min originates from intein impurities. **B:** SDS-PAGE monitoring indicates product formation by reducing recombinant starting material 1-528-SR and emerging a new band between the 66 and 97 kDa marker bands.

Unfortunately, a mass determination of the product using ESI-MS was not possible for us, which was to be expected since the recombinant segment **13** already proved to ionize badly when applied for ESI-MS (see section 3.3). Furthermore, we could not characterize the starting material or product using MALDI-MS, a more suitable method for high molecular weight proteins.[139] Notably, literature frequently reports that Hsp90 MALDI TOF characterization is done using a tryptic digest to gather a peptide mass fingerprint, thereby indicating the difficulties when applying the intact full-length protein for MS characterization.[140,141] However, also chemically crosslinked Hsp90 oligomers have been reportedly characterized *via* MALDI TOF using sinapinic acid as a sample matrix.[142]

Overall, we assume that higher final protein purity and total amount of the final product would enable the proper mass characterization of semi-synthetic Hsp90 in the future.

Nevertheless, a final ligation on a low mg scale (1.2 mg) was performed and purified *via* analytical RP-HPLC, which yielded the highly desired full-length semi-synthetic Hsp90 (~200 ug, 16%) as characterized in **Figure 46**.

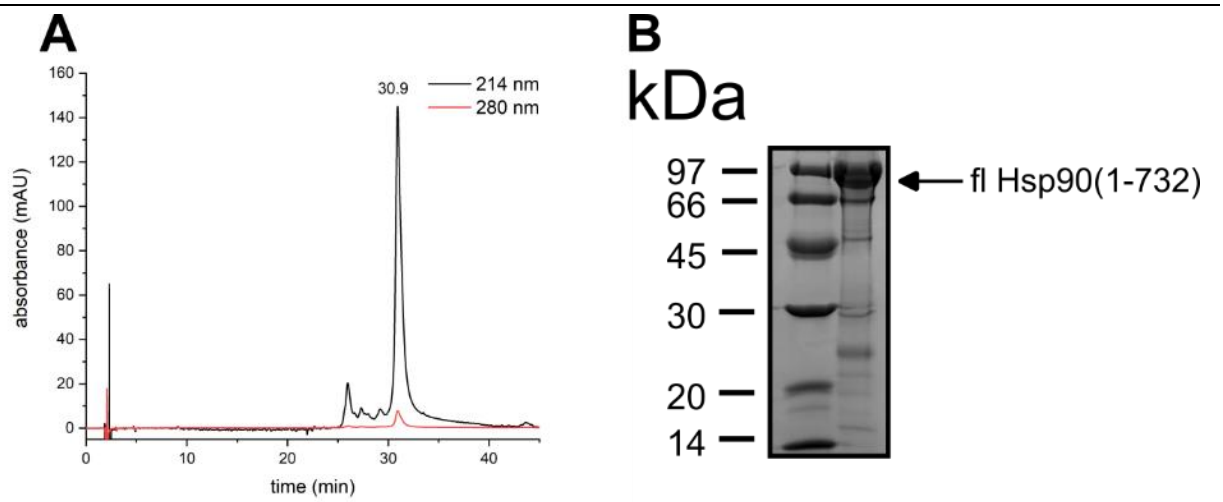

Figure 46 Characterization of semi-synthetic full-length Hsp90α (1-732). **A:** RP-HPLC characterization with indicated product peak at 30.9 min. **B:** SDS-PAGE characterization on purified semi-synthetic Hsp90 1-732.

Unfortunately, the challenges regarding the purification of larger recombinant Hsp90 constructs are also affecting the separation after the successful ligation of the full-length protein. SDS-PAGE and RP-HPLC characterization revealed that impurities in the form of the large 1-528-intein fusion protein, which were carried over from the recombinant starting material and unreacted 1-528-SR are still present after purification and presumably elute at the same retention time as the product. Nevertheless, although the final confirmation *via* high-resolution MS is missing, our results strongly imply the successful semi-synthetic generation of full-length Hsp90, representing a milestone in the project. Starting material generation and purity of the final ligation remains to be optimized and represent a not inconsiderable hurdle for the generation of pure semi-synthetic Hsp90. In this regard, optimization of the generation of the recombinant segment 1-528-SR represents a key step in generating full-length Hsp90, as discussed in the previous section. An alternative approach for achieving better separability of the starting material and /or other impurities could be ligation and purification under non-denaturing conditions. Notably, both the uncleaved fusion protein and the thioester segment 13 were stable in buffers without adding chaotropic salts once the inclusion bodies were solubilized and the buffer was exchanged *via* dialysis (see **Table 1** for buffer conditions). This indicates that at least partial refolding of the protein would be supported by reports of Hsp90 unfolding and refolding being mostly reversible.[143] Changes in the protein properties upon refolding could yield overall improved separation and recovery of the final product.[144]

4. Conclusion

This chapter explores the access of full-length human Hsp90 alpha via a sequential C- to N-terminal semi-synthesis strategy. As described in Chapter Two, the successful semi-synthetic generation of Hsp90 CTD with free N-terminal Cys residue enabled the domain's extension *via* the ligation of 42mer MD peptide. We successfully generated the synthetic MD peptide segment as an unmodified control and modified variant bearing an acetylated lysin and ligated the unmodified variant to the already semi-synthetically CTD. Although a strategy that would involve recombinant generation of the whole CTD as a single segment, e.g., including an N-terminal affinity tag that includes a protease cleavage site for the release of the Cys residues, would provide a more straightforward and possibly easier to upscale approach, the successful generation of semi-synthetic CTD enables the incorporation of PTMs on two distinctly different domains, the CTD and MD, of Hsp90. This enables us to study PTM crosstalk, which is highly desired for studying the chaperone code. Nevertheless, the final N-terminal segment piece required for the full-length production of Hsp90 provided similar challenges when it comes to the yield and purification as encountered when working on the NTD segment 20-235. Expression monitoring shows that the fusion protein, presumably due to its relatively large size of ~90 kDa, provides a burden to the E. coli expression machinery indicated by a relatively low overexpression resulting in only modest expression yields (~50 mg per Liter culture) before further processing. Furthermore, the incomplete intein cleavage and the low HPLC recovery rates limited the ability to upscale the generation of this segment. This hindered the possibility of a larger-scale final ligation and required further optimization of starting material generation. However, the access to small amounts of this segment and the data indicating the successful ligation and generation of full-length Hsp90 can be considered a major achievement. It exemplifies the possibility of accessing this challenging protein, which is among the largest proteins semi-synthetically accessed up to this point, enabled by recent achievements in seleno-ligation chemistry.

3. Final conclusion and outlook

This thesis explores several protein semi-synthesis approaches for studying the chaperone code oh Hsp90, which describes the complex modulation of a pivotal chaperone by post-translation modifications. Although protein semi-synthesis has already proven to be an invaluable tool for providing access to site-selectively modified proteins and thereby enabling unique insights into the specific impact of PTMs, e.g., for the neurodegenerative disease-related proteins Tau and alpha-synuclein, larger and more structurally conserved proteins like Hsp90 are usually omitted since they usually provide a much more challenging target due to their complicated handling. Here, we can show that by carefully designing a multi-step semi-synthesis strategy and by utilizing novel seleno-based chemo selective ligation strategies (EPSL), we are able to gain access to modified Hsp90 CTD variants and full-length Hsp90 with the potential to introduce site-selectively C-terminal and middle domain modifications. Although the impact of the introduced PTMs on CTD structure and function was only minor when studying the generated variants, further studies, e.g., CTD and co-chaperone interactions, could provide new insights into the regulatory role on protein-protein interactions. Furthermore, the sequential C- to N- terminal elongation of the CTD using a combination of EPSL and a classic NCL-based semi-synthesis route allows the study of PTM crosstalk by providing access to a variety of PTM combinations located at the C-terminus of the protein and in the middle domain. This would also enable us to study rarer PTMs of Hsp90, e.g., lysine succinylation as reported to be observed on MD residue Lys546, which would be theoretically accessible through the synthetic peptide segment 529-571 used in this synthesis route. In addition, great efforts were undertaken also to provide access to N-terminally modified Hsp90 NTD variants. However, although the underlying chemistry, involving a nonnative Cysteine mediated NCL followed by a selective desulfurization, is shown to be possible, the obstacles when it comes to handling the protein due to its high hydrophobicity and aggregation proneness require further optimization before the sufficient generation of the desired modified Hsp90 NTD variants is possible. This could be achieved by combining different factors, including optimizing the ligation and desulfurization step and different purification strategies, or *via* a more ground-up approach involving the redesign of the recombinant segment for improved solubility. In summary, this thesis contributes to the exploration of Hsp90's complex regulation by PTMs by providing access to modified protein variants through a combination of well-established NCL-based and more recent selenochemistry-based protein semi-synthesis strategies. This allows a uniquely detailed inside into the chaperone code and represents an exciting foundation for further studies of this pivotal protein.

4 References

[1] L. Le Breton, M. P. Mayer, *eLife* **2016**, *5*, e22850.

[2] K. Richter, M. Haslbeck, J. Buchner, *Molecular Cell* **2010**, *40*, 253–266.

[3] M. K. Kurop, C. M. Huyen, J. H. Kelly, B. S. J. Blagg, *Eur J Med Chem* **2021**, *226*, 113846.

[4] F. Ritossa, *Experientia* **1962**, *18*, 571–573.

[5] R. I. Morimoto, *Cold Spring Harb. Symp. Quant. Biol.* **2011**, *76*, 91–99.

[6] R. Gomez-Pastor, E. T. Burchfiel, D. J. Thiele, *Nat Rev Mol Cell Biol* **2018**, *19*, 4–19.

[7] M. K. Singh, Y. Shin, S. Ju, S. Han, W. Choe, K.-S. Yoon, S. S. Kim, I. Kang, *International Journal of Molecular Sciences* **2024**, *25*, 4209.

[8] R. J. Ellis, *Semin Cell Biol* **1990**, *1*, 1–9.

[9] C. Jantschitsch, F. Trautinger, *Photochem Photobiol Sci* **2003**, *2*, 899–903.

[10] A. Al Refaii, J.-H. Alix, *Mol Microbiol* **2009**, *71*, 748–762.

[11] C. Hu, J. Yang, Z. Qi, H. Wu, B. Wang, F. Zou, H. Mei, J. Liu, W. Wang, Q. Liu, *MedComm (2020)* **2022**, *3*, e161.

[12] F. Di Domenico, R. Sultana, G. F. Tiu, N. N. Scheff, M. Perluigi, C. Cini, D. A. Butterfield, *Brain Res* **2010**, *1333*, 72–81.

[13] R. A. Laskey, B. M. Honda, A. D. Mills, J. T. Finch, *Nature* **1978**, *275*, 416–420.

[14] L. Tutar, Y. Tutar, *Curr Pharm Biotechnol* **2010**, *11*, 216–222.

[15] A. M. de Graff, D. E. Mosedale, T. Sharp, K. A. Dill, D. J. Grainger, *PLoS Comput Biol* **2020**, *16*, e1008460.

[16] B. S. Rutledge, W.-Y. Choy, M. L. Duennwald, *J Biol Chem* **2022**, *298*, 101905.

[17] M. Haslbeck, S. Weinkauf, J. Buchner, *J Biol Chem* **2019**, *294*, 2121–2132.

[18] M. Haslbeck, T. Franzmann, D. Weinfurtner, J. Buchner, *Nat Struct Mol Biol* **2005**, *12*, 842–846.

[19] M. K. Singh, Y. Shin, S. Han, J. Ha, P. K. Tiwari, S. S. Kim, I. Kang, *Int J Mol Sci* **2024**, *25*, 5483.

[20] R. Rosenzweig, N. B. Nillegoda, M. P. Mayer, B. Bukau, *Nat Rev Mol Cell Biol* **2019**, *20*, 665–680.

[21] D. P. Easton, Y. Kaneko, J. R. Subjeck, *Cell Stress Chaperones* **2000**, *5*, 276–290.

[22] D. Zuo, J. Subjeck, X.-Y. Wang, *Front Immunol* **2016**, *7*, 75.

[23] P. Csermely, T. Schnaider, C. Soti, Z. Prohászka, G. Nardai, *Pharmacol Ther* **1998**, *79*, 129–168.

[24] P. Csermely, T. Schnaider, C. So″ti, Z. Prohászka, G. Nardai, *Pharmacology & Therapeutics* **1998**, *79*, 129–168.

[25] L. H. Pearl, C. Prodromou, *Annual Review of Biochemistry* **2006**, *75*, 271–294.

[26] M. Radli, S. G. D. Rüdiger, *J. Mol. Biol.* **2018**, *430*, 3029–3040.

[27] X. Ren, T. Li, W. Zhang, X. Yang, *Cells* **2022**, *11*, 2556.

[28] B. J. Lang, T. L. Prince, Y. Okusha, H. Bunch, S. K. Calderwood, *Biochim Biophys Acta Mol Cell Res* **2021**, 119187.

[29] A. M. Jaeger, L. Whitesell, *Annu. Rev. Cancer Biol.* **2018**, DOI 10.1146/annurev-cancer-bio-030518-055533.

[30] A. Hoter, M. E. El-Sabban, H. Y. Naim, *International Journal of Molecular Sciences* **2018**, *19*, 2560.

[31] T. Doi, N. Yamamoto, S. Ohkubo, *Future Oncology* **2024**, *20*, 507–519.

[32] J. Liu, H. Shu, Q. Xia, Q. You, L. Wang, *Expert Opinion on Therapeutic Patents* **2024**, *34*, 1–15.

[33] K. Ono, T. Eguchi, *Methods Mol Biol* **2023**, *2693*, 141–161.

[34] S. Maiti, D. Picard, *Biomolecules* **2022**, *12*, 1166.

[35] C. Chang, X. Tang, D. T. Woodley, M. Chen, W. Li, *Cells* **2023**, *12*, 277.

[36] V. A. Assimon, D. R. Southworth, J. E. Gestwicki, *Biochemistry* **2015**, *54*, 7120–7131.

[37] E. M. Clerico, L. M. Gierasch, *Mol Cell* **2022**, *82*, 1403–1404.

[38] A. K. Shiau, S. F. Harris, D. R. Southworth, D. A. Agard, *Cell* **2006**, *127*, 329–340.

[39] V. Dahiya, D. A. Rutz, P. Moessmer, M. Mühlhofer, J. Lawatscheck, M. Rief, J. Buchner, *Molecular Cell* **2022**, *82*, 1543-1556.e6.

[40] S. M. Roe, M. M. U. Ali, P. Meyer, C. K. Vaughan, B. Panaretou, P. W. Piper, C. Prodromou, L. H. Pearl, *Cell* **2004**, *116*, 87–98.

[41] A. Wolmarans, B. Lee, L. Spyracopoulos, P. LaPointe, *Scientific Reports* **2016**, *6*, 33179.

[42] M. M. U. Ali, S. M. Roe, C. K. Vaughan, P. Meyer, B. Panaretou, P. W. Piper, C. Prodromou, L. H. Pearl, *Nature* **2006**, *440*, 1013–1017.

[43] R. A. Sager, M. R. Woodford, L. Neckers, M. Mollapour, *Methods Mol Biol* **2018**, *1709*, 209–219.

[44] J. J. Kovacs, P. J. M. Murphy, S. Gaillard, X. Zhao, J.-T. Wu, C. V. Nicchitta, M. Yoshida, D. O. Toft, W. B. Pratt, T.-P. Yao, *Molecular Cell* **2005**, *18*, 601–607.

[45] E. G. Mimnaugh, P. J. Worland, L. Whitesell, L. M. Neckers, *Journal of Biological Chemistry* **1995**, *270*, 28654–28659.

[46] P. Muller, E. Ruckova, P. Halada, P. J. Coates, R. Hrstka, D. P. Lane, B. Vojtesek, *Oncogene* **2013**, *32*, 3101–3110.

[47] S. J. Backe, R. A. Sager, J. A. Heritz, L. A. Wengert, K. A. Meluni, X. Aran-Guiu, B. Panaretou, M. R. Woodford, C. Prodromou, D. Bourboulia, M. Mollapour, *Cell Reports* **2023**, *42*, 112807.

[48] M. Mollapour, S. Tsutsumi, A. W. Truman, W. Xu, C. K. Vaughan, K. Beebe, A. Konstantinova, S. Vourganti, B. Panaretou, P. W. Piper, J. B. Trepel, C. Prodromou, L. H. Pearl, L. Neckers, *Mol. Cell* **2011**, *41*, 672–681.

[49] M. Mollapour, S. Tsutsumi, Y. S. Kim, J. Trepel, L. Neckers, *Oncotarget* **2011**, *2*, 407–417.

[50] B. T. Scroggins, K. Robzyk, D. Wang, M. G. Marcu, S. Tsutsumi, K. Beebe, R. J. Cotter, S. Felts, D. Toft, L. Karnitz, N. Rosen, L. Neckers, *Mol. Cell* **2007**, *25*, 151–159.

[51] S. P. Lees-Miller, C. W. Anderson, *J. Biol. Chem.* **1989**, *264*, 17275–17280.

[52] M. Quanz, A. Herbette, M. Sayarath, L. de Koning, T. Dubois, J.-S. Sun, M. Dutreix, *Journal of Biological Chemistry* **2012**, *287*, 8803–8815.

[53] S. Solier, K. W. Kohn, B. Scroggins, W. Xu, J. Trepel, L. Neckers, Y. Pommier, *Proceedings of the National Academy of Sciences* **2012**, *109*, 12866–12872.

[54] K. E. Paleologou, A. W. Schmid, C. C. Rospigliosi, H.-Y. Kim, G. R. Lamberto, R. A. Fredenburg, P. T. Lansbury, C. O. Fernandez, D. Eliezer, M. Zweckstetter, H. A. Lashuel, *J Biol Chem* **2008**, *283*, 16895–16905.

[55] D. Rabuka, *Curr Opin Chem Biol* **2010**, *14*, 790–796.

[56] P. Cloutier, B. Coulombe, *Biochim Biophys Acta* **2013**, *1829*, 443–454.

[57] M. J. Suskiewicz, *BioEssays* **2024**, *46*, 2300178.

[58] Q. Zhong, X. Xiao, Y. Qiu, Z. Xu, C. Chen, B. Chong, X. Zhao, S. Hai, S. Li, Z. An, L. Dai, *MedComm (2020)* **2023**, *4*, e261.

[59] J. Hermann, L. Schurgers, V. Jankowski, *Molecular Aspects of Medicine* **2022**, *86*, 101066.

[60] A. S. Venne, L. Kollipara, R. P. Zahedi, *Proteomics* **2014**, *14*, 513–524.

[61] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

[62] C.-D. Chang, J. Meienhofer, *International Journal of Peptide and Protein Research* **1978**, *11*, 246–249.

[63] S. Pomplun, M. Jbara, C. K. Schissel, S. Wilson Hawken, A. Boija, C. Li, I. Klein, B. L. Pentelute, *ACS Cent. Sci.* **2021**, *7*, 1408–1418.

[64] N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo, X. Ye, A. J. Mijalis, M. D. Simon, A. Loas, S. Liu, C. Jessen, T. E. Nielsen, B. L. Pentelute, *Science* **2020**, *368*, 980–987.

[65] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.

[66] B. L. Nilsson, M. B. Soellner, R. T. Raines, *Annu Rev Biophys Biomol Struct* **2005**, *34*, 91–118.

[67] S. C. Stanhope, T. Brandwine-Shemmer, H. R. Blum, E. H. Doud, A. Jannasch, A. L. Mosley, B. Minke, V. M. Weake, *Redox Biol* **2023**, *63*, 102723.

[68] J. Sánchez, *Biochem Biophys Res Commun* **2008**, *373*, 589–592.

[69] L. R. Malins, R. J. Payne, *Australian Journal of Chemistry* **2015**, *68*, 521–537.

[70] H. Liu, X. Li, *Acc. Chem. Res.* **2018**, *51*, 1643–1655.

[71] S. S. Kulkarni, E. E. Watson, B. Premdjee, K. W. Conde-Frieboes, R. J. Payne, *Nature Protocols* **2019**, *14*, 2229–2257.

[72] J. W. Bode, R. M. Fox, K. D. Baucom, *Angewandte Chemie International Edition* **2006**, *45*, 1248–1252.

[73] M. Patil, *Org. Biomol. Chem.* **2017**, *15*, 416–425.

[74] M. Patil, *Org. Biomol. Chem.* **2017**, *15*, 416–425.

[75] J. Farnung, H. Song, J. W. Bode, *Methods Mol Biol* **2021**, *2355*, 151–162.

[76] S. Baldauf, A. O. Ogunkoya, G. N. Boross, J. W. Bode, *J. Org. Chem.* **2020**, *85*, 1352–1364.

[77] W. H. H. Guenther, *J. Org. Chem.* **1967**, *32*, 3931–3933.

[78] M. J. Bedding, S. S. Kulkarni, R. J. Payne, *Methods Enzymol* **2022**, *662*, 363–399.

[79] S. B. Pollock, S. B. H. Kent, *Chem. Commun.* **2011**, *47*, 2342–2344.

[80] S. Dery, P. S. Reddy, L. Dery, R. Mousa, R. N. Dardashti, N. Metanis, *Chem. Sci.* **2015**, *6*, 6207–6212.

[81] J. Ma, J. Zeng, Q. Wan, *Top Curr Chem* **2015**, *363*, 57–101.

[82] N. J. Mitchell, S. S. Kulkarni, L. R. Malins, S. Wang, R. J. Payne, *Chemistry – A European Journal* **2017**, *23*, 946–952.

[83] O. Fuchs, S. Trunschke, H. Hanebrink, M. Reimann, O. Seitz, *Angew Chem Int Ed Engl* **2021**, *60*, 19483–19490.

[84] S. Trunschke, E. Piemontese, O. Fuchs, S. Abboud, O. Seitz, *Chemistry – A European Journal* **2022**, *28*, e202202065.

[85] A. K. Alexander, S. I. Elshahawi, *ChemBioChem* **2023**, *24*, e202300372.

[86] M. A. Shandell, Z. Tan, V. W. Cornish, *Biochemistry* **2021**, *60*, 3455–3469.

[87] R. E. Thompson, T. W. Muir, *Chem. Rev.* **2020**, *120*, 3051–3126.

[88] Z. A. Wang, P. A. Cole, *Methods Mol Biol* **2020**, *2133*, 1–13.

[89] P. M. Kane, C. T. Yamashiro, D. F. Wolczyk, N. Neff, M. Goebl, T. H. Stevens, *Science* **1990**, *250*, 651–657.

[90] R. Hirata, Y. Ohsumk, A. Nakano, H. Kawasaki, K. Suzuki, Y. Anraku, *J Biol Chem* **1990**, *265*, 6726–6733.

[91] F. B. Perler, *IUBMB Life* **2005**, *57*, 469–476.

[92] T. W. Muir, D. Sondhi, P. A. Cole, *Proceedings of the National Academy of Sciences* **1998**, *95*, 6705–6710.

[93] T. W. Muir, *Annual Review of Biochemistry* **2003**, *72*, 249–289.

[94] S. S. Kulkarni, E. E. Watson, J. W. C. Maxwell, G. Niederacher, J. Johansen-Leete, S. Huhmann, S. Mukherjee, A. R. Norman, J. Kriegesmann, C. F. W. Becker, R. J. Payne, *Angewandte Chemie International Edition* **2022**, *61*, e202200163.

[95] J. Hu, W. Xia, S. Zeng, Y.-J. Lim, Y. Tao, Y. Sun, L. Zhao, H. Wang, W. Le, D. Li, S. Zhang, C. Liu, Y.-M. Li, *Nat Commun* **2024**, *15*, 2677.

[96] M. Haj-Yahya, H. A. Lashuel, *J. Am. Chem. Soc.* **2018**, *140*, 6611–6621.

[97] G. Niederacher, D. Urwin, Y. Dijkwel, D. J. Tremethick, K. J. Rosengren, C. F. W. Becker, A. C. Conibear, *RSC Chem Biol* **n.d.**, *2*, 537–550.

[98] J. C. Shimko, C. J. Howard, M. G. Poirier, J. J. Ottesen, *Methods Mol Biol* **2013**, *981*, 10.1007/978-1-62703-305–3_14.

[99] G. K. T. Nguyen, S. Wang, Y. Qiu, X. Hemu, Y. Lian, J. P. Tam, *Nat Chem Biol* **2014**, *10*, 732–738.

[100] H. Mao, S. A. Hart, A. Schink, B. A. Pollok, *J. Am. Chem. Soc.* **2004**, *126*, 2670–2671.

[101] S. Riedl, E. Bilgen, G. Agam, V. Hirvonen, A. Jussupow, F. Tippl, M. Riedl, A. Maier, C. F. W. Becker, V. R. I. Kaila, D. C. Lamb, J. Buchner, *Nat Commun* **2024**, *15*, 8627.

[102] O. R. Lorenz, L. Freiburger, D. A. Rutz, M. Krause, B. K. Zierer, S. Alvira, J. Cuéllar, J. M. Valpuesta, T. Madl, M. Sattler, J. Buchner, *Molecular Cell* **2014**, *53*, 941–953.

[103] J. Lee, J. McIntosh, B. J. Hathaway, E. W. Schmidt, *J Am Chem Soc* **2009**, *131*, 2122–2124.

[104] G. K. T. Nguyen, S. Wang, Y. Qiu, X. Hemu, Y. Lian, J. P. Tam, *Nat Chem Biol* **2014**, *10*, 732–738.

[105] G. K. T. Nguyen, Y. Cao, W. Wang, C. F. Liu, J. P. Tam, *Angew Chem Int Ed Engl* **2015**, *54*, 15694–15698.

[106] X. Hemu, X. Zhang, X. Bi, C.-F. Liu, J. P. Tam, *Methods Mol Biol* **2019**, *2012*, 83–109.

[107] A. K. Singh, A. Antonenko, A. Kocyła, A. Krężel, *Microbial Cell Factories* **2024**, *23*, 325.

[108] S. P. Lees-Miller, C. W. Anderson, *J. Biol. Chem.* **1989**, *264*, 2431–2437.

[109] J.-S. Zheng, S. Tang, Y.-K. Qi, Z.-P. Wang, L. Liu, *Nature Protocols* **2013**, *8*, 2483–2495.

[110] Z. Sun, W. Ma, Y. Cao, T. Wei, X. Mo, H. Y. Chow, Y. Tan, C. H. P. Cheung, J. Liu, H. K. Lee, E. C. M. Tse, H. Liu, X. Li, *Chem* **2022**, *8*, 2542–2557.

[111] J.-S. Zheng, S. Tang, Y.-K. Qi, Z.-P. Wang, L. Liu, *Nature Protocols* **2013**, *8*, 2483–2495.

[112] R. B. Kapust, J. Tözsér, J. D. Fox, D. E. Anderson, S. Cherry, T. D. Copeland, D. S. Waugh, *Protein Engineering, Design and Selection* **2001**, *14*, 993–1000.

[113] C. M. Hickey, N. R. Wilson, M. Hochstrasser, *Nat Rev Mol Cell Biol* **2012**, *13*, 755–766.

[114] D. A. Polayes, T. D. Parks, S. A. Johnston, W. G. Dougherty, *Methods Mol Med* **1998**, *13*, 169–183.

[115] P. Wingfield, *Curr Protoc Protein Sci* **2017**, *88*, 6.14.1-6.14.3.

[116] Z. Sun, W. Ma, Y. Cao, T. Wei, X. Mo, H. Y. Chow, Y. Tan, C. H. P. Cheung, J. Liu, H. K. Lee, E. C. M. Tse, H. Liu, X. Li, *Chem* **2022**, *8*, 2542–2557.

[117] R. E. Thompson, X. Liu, N. Alonso-García, P. J. B. Pereira, K. A. Jolliffe, R. J. Payne, *J Am Chem Soc* **2014**, *136*, 8161–8164.

[118] R. Jing, M. A. Walczak, *Org. Lett.* **2024**, *26*, 2590–2595.

[119] M. Waliczek, P. Stefanowicz, *Org. Lett.* **2025**, DOI 10.1021/acs.orglett.4c04671.

[120] D. Walls, S. T. Loughran, *Methods Mol Biol* **2011**, *681*, 151–175.

[121] *Chemical Science* **2021**, *12*, 3194–3201.

[122] S. Tang, C. Zuo, D.-L. Huang, X.-Y. Cai, L.-H. Zhang, C.-L. Tian, J.-S. Zheng, L. Liu, *Nat Protoc* **2017**, *12*, 2554–2569.

[123] K. M. Goode, D. P. Petrov, R. E. Vickman, S. A. Crist, P. E. Pascuzzi, T. L. Ratliff, V. J. Davisson, T. R. Hazbun, *Biochim Biophys Acta Gen Subj* **2017**, *1861*, 1992–2006.

[124] E. V. Mymrikov, M. Daake, B. Richter, M. Haslbeck, J. Buchner, *J Biol Chem* **2017**, *292*, 672–684.

[125] S. S. Kulkarni, E. E. Watson, J. W. C. Maxwell, G. Niederacher, J. Johansen-Leete, S. Huhmann, S. Mukherjee, A. R. Norman, J. Kriegesmann, C. F. W. Becker, R. J. Payne, *Angew Chem Int Ed Engl* **2022**, *61*, e202200163.

[126] I. E. Gentle, D. P. De Souza, M. Baca, *Bioconjugate Chem.* **2004**, *15*, 658–663.

[127] E. Ciglia, J. Vergin, S. Reimann, S. H. J. Smits, L. Schmitt, G. Groth, H. Gohlke, *PLoS One* **2014**, *9*, e96031.

[128] R. K. Allan, D. Mok, B. K. Ward, T. Ratajczak, *Journal of Biological Chemistry* **2006**, *281*, 7161–7171.

[129] Y. Minami, Y. Kimura, H. Kawasaki, K. Suzuki, I. Yahara, *Mol Cell Biol* **1994**, *14*, 1459–1464.

[130] S. Peng, J. Woodruff, P. K. Pathak, R. L. Matts, J. Deng, *Acta Crystallogr D Struct Biol* **2022**, *78*, 571–585.

[131] S. Bhatia, L. Spanier, D. Bickel, N. Dienstbier, V. Woloschin, M. Vogt, H. Pols, B. Lungerich, J. Reiners, N. Aghaallaei, D. Diedrich, B. Frieg, J. Schliehe-Diecks, B. Bopp, F. Lang, M. Gopalswamy, J. Loschwitz, B. Bajohgli, J. Skokowa, A. Borkhardt, J. Hauer, F. K. Hansen, S. H. J. Smits, J. Jose, H. Gohlke, T. Kurz, *ACS Cent. Sci.* **2022**, *8*, 636–655.

[132] Y. Yang, R. Rao, J. Shen, Y. Tang, W. Fiskus, J. Nechtman, P. Atadja, K. Bhalla, *Cancer Res* **2008**, *68*, 4833–4842.

[133] B. T. Weinert, C. Schölz, S. A. Wagner, V. Iesmantavicius, D. Su, J. A. Daniel, C. Choudhary, *Cell Rep* **2013**, *4*, 842–851.

[134] R. J. Spears, C. McMahon, V. Chudasama, *Chemical Society Reviews* **2021**, *50*, 11098–11155.

[135] E. P. O'Brien, R. I. Dima, B. Brooks, D. Thirumalai, *J Am Chem Soc* **2007**, *129*, 7346–7353.

[136] G. Volkmann, W. Sun, X.-Q. Liu, *Protein Sci* **2009**, *18*, 2393–2402.

[137] M. Vila-Perelló, Z. Liu, N. H. Shah, J. A. Willis, J. Idoyaga, T. W. Muir, *J Am Chem Soc* **2013**, *135*, 286–292.

[138] B. J. Lim, R. F. Berkeley, G. T. Debelouchina, *Methods Mol Biol* **2020**, *2133*, 163–181.

[139] L. Signor, E. Boeri Erba, *J Vis Exp* **2013**, 50635.

[140] T. Shibata, Y. Kimura, A. Mukai, H. Mori, S. Ito, Y. Asaka, S. Oe, H. Tanaka, T. Takahashi, K. Uchida, *J Biol Chem* **2011**, *286*, 42150–42161.

[141] A. H. Sosalegowda, R. R. Kundapur, M. H. Boregowda, *Proteomics* **2010**, *10*, 2734–2745.

[142] L. Moullintraffort, M. Bruneaux, A. Nazabal, D. Allegro, E. Giudice, F. Zal, V. Peyrot, P. Barbier, D. Thomas, C. Garnier, *J Biol Chem* **2010**, *285*, 15100–15110.

[143] M. Jahn, J. Buchner, T. Hugel, M. Rief, *Proceedings of the National Academy of Sciences* **2016**, *113*, 1232–1237.

[144] L. Shiwen, B. L. Karger, *Journal of Chromatography A* **1990**, *499*, 89–102.

5 Appendix

5.1 Gene and peptide sequences

5.1.1 Gene sequence of Hsp90α 21-235-His_6 (A21C)

TGT TTT CAG GCG GAG ATC GCC CAA CTG ATG AGC TTA ATC ATT AAT ACG TTC TAT AGC AAC AAG GAG ATT TTT TTG CGT GAA CTT ATC TCG AAT TCT TCG GAC GCG CTT GAC AAA ATT CGT TAT GAA TCC CTG ACC GAC CCC AGC AAA TTA GAT AGT GGC AAG GAG CTT CAT ATC AAC CTG ATT CCA AAC AAG CAA GAT CGC ACC CTG ACT ATT GTA GAC ACG GGC ATC GGG ATG ACG AAA GCA GAC TTG ATT AAC AAC TTG GGT ACT ATT GCT AAA TCT GGC ACG AAG GCT TTT ATG GAG GCC CTT CAA GCC GGG GCC GAC ATC AGT ATG ATC GGA CAG TTT GGT GTT GGG TTC TAC AGT GCA TAC TTA GTC GCA GAA AAA GTA ACC GTT ATC ACA AAA CAC AAT GAC GAT GAA CAG TAC GCA TGG GAG TCA TCA GCG GGA GGA AGT TTC ACA GTA CGC ACT GAC ACT GGT GAA CCA ATG GGT CGT GGA ACC AAA GTG ATC TTG CAT TTG AAA GAA GAT CAG ACA GAG TAT CTG GAA GAA CGT CGC ATT AAA GAG ATT GTC AAA AAG CAT AGT CAA TTC ATC GGT TAC CCT ATT ACG TTA TTT GTG GAA AAA GAA CGC GAC AAA GAG GTG AGT GAT GAC GAA GCT CAC CAC CAT CAC CAT CAC

5.1.2 Translated aminoacid sequence of Hsp90α 21-235-His_6 (A21C)

CFQAEIAQLM SLIINTFYSN KEIFLRELIS NSSDALDKIR YESLTDPSKL DSGKELHINL IPNKQDRTLT IVDTGIGMTK ADLINNLGTI AKSGTKAFME ALQAGADISM IGQFGVGFYS AYLVAEKVTV ITKHNDDEQY AWESSAGGSF TVRTDTGEPM GRGTKVILHL KEDQTEYLEE RRIKEIVKKH SQFIGYPITL FVEKERDKEV SDDEA HHHHHH

5.1.3 Peptide sequence of Hsp90alpha 2-20

PEETQTQDQ PMEEEEVETF

5.1.4 Gene sequence of Hsp90α 572-710-*Mxe* GyrA-His$_7$-CBD

CAT ATG TGT AAA ATC ATG AAA GAC ATT CTT GAA AAG AAA GTC GAA AAA GTA GTC GTT TCT AAT CGT TTA GTA ACC TCT CCC TGC TGC ATT GTA ACT AGC ACT TAC GGG TGG ACT GCC AAC ATG GAA CGC ATT ATG AAG GCC CAA GCC CTG CGT GAT AAC AGT ACG ATG GGA TAC ATG GCT GCT AAG AAG CAC CTG GAG ATC AAC CCC GAT CAT TCA ATT ATT GAA ACT CTG CGT CAA AAA GCC GAA GCT GAC AAA AAT GAC AAG TCG GTG AAA GAC TTG GTC ATT CTT TTA TAC GAA ACT GCG CTT TTG TCA AGT GGA TTT AGC TTG GAG GAC CCA CAG ACT CAT GCA AAT CGT ATT TAT CGT ATG ATT AAG CTT GGT CTT GGA ATC GAT GAA GAT GAT CCG ACC GCC GAC GAC ACG TCC GCA TGT ATC ACT GGT GAC GCA CTT GTG GCT CTG CCA GAA GGC GAG TCG GTC CGT ATC GCG GAC ATC GTT CCA GGT GCG CGC CCG AAC TCC GAC AAC GCA ATC GAC TTA AAA GTC TTA GAT CGT CAC GGT AAT CCG GTC TTA GCC GAT CGT CTT TTT CAT AGC GGC GAA CAT CCT GTG TAT ACG GTC CGC ACC GTG GAG GGG TTA CGC GTG ACC GGC ACG GCC AAT CAC CCG TTG TTA TGT TTG GTA GAC GTT GCG GGG GTT CCT ACT TTG TTA TGG AAA CTT ATC GAT GAA ATT AAG CCT GGA GAC TAT GCC GTA ATC CAA CGT AGT GCG TTC AGT GTA GAT TGT GCA GGG TTC GCT CGC GGA AAG CCC GAG TTC GCC CCA ACG ACA TAT ACT GTA GGA GTG CCG GGG TTG GTT CGT TTT CTG GAG GCT CAT CAC CGT GAC CCG GAT GCT CAG GCC ATT GCA GAC GAA CTT ACG GAC GGT CGC TTT TAC TAT GCA AAA GTA GCT TCA GTA ACC GAC GCC GGA GTG CAA CCA GTT TAT AGC CTT CGT GTA GAC ACG GCG GAC CAT GCC TTC ATC ACC AAT GGC TTT GTC AGC CAC GCT ACT GGA CTT ACT GGA ATC CAT CAT CAC CAT CAC CAC CAT TCA GGT TTG AAT TCA GGA CTT ACG ACT AAT CCT GGT GTA TCC GCG TGG CAG GTC AAT ACA GCA TAT ACG GCT GGG CAG CTG GTA ACC TAC AAT GGT AAA ACG TAC AAA TGT TTG CAG CCA CAT ACG AGC CTT GCA GGC TGG GAG CCT TCC AAC GTA CCT GCT CTT TGG CAA CTT CAA TGA CTC GAG

5.1.5 Translated amino acid sequence of Hsp90 α 572-710-*Mxe* GyrA-His$_7$-CBD

CKIMKDILEKKVEKVVVSNRLVTSPCCIVTSTYGWTANMERIMKAQALRD-
NSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSVKDLVILLYETALLSSGFSLEDP
QTHANRIYRMIKLGLGIDEDDPTADDTSACITGDALVALPEGESVRIADIVPGARPNSD-
NAILKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTL
LWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRD-
PDAQAADELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGIH
HHHHHHSGLNSGLTTNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEP-
SNVPALWQLQ

5.1.6 Gene sequence of WT Hsp90α CTD (His_6-ENLYFQ-572-732) plasmid

CAC CAT CAT CAC CAC CAC GAG AAC TTA TAT TTC CAA TGT AAA ATC ATG AAA GAC ATT TTG GAA AAG AAA GTG GAG AAG GTT GTC GTG AGT AAT CGC CTT GTA ACC TCT CCG TGC TGT ATC GTC ACG AGT ACC TAC GGA TGG ACA GCG AAC ATG GAG CGC ATC ATG AAG GCT CAA GCT TTA CGC GAT AAC TCT ACG ATG GGA TAT ATG GCT GCC AAG AAA CAT CTT GAA ATT AAC CCG GAT CAT AGT ATC ATC GAA ACC CTG CGC CAA AAA GCG GAA GCA GAT AAG AAC GAT AAG TCT GTC AAG GAT CTT GTG ATC TTG TTG TAC GAG ACG GCA TTG TTG TCC AGC GGG TTC AGC TTG GAA GAT CCA CAG ACA CAC GCG AAC CGT ATC TAT CGT ATG ATT AAG CTG GGG TTG GGT ATT GAC GAA GAT GAT CCT ACT GCC GAT GAC ACG TCT GCT GCG GTT ACG GAA GAA ATG CCG CCT TTA GAA GGC GAT GAT GAC ACC AGC CGC ATG GAA GAA GTT GAC

5.1.7 Translated Aminoacid sequence of WT-Hsp90CTD

HHHHHHENLYFQ CKIMKDILE KKVEKVVVSN RLVTSPCCIV TSTYGWTANM ERIMKAQALR DNSTMGYMAA KKHLEINPDH SIIETLRQKA EADKNDKSVK DLVILLYETA LLSSGFSLED PQTHANRIYR MIKLGLGIDE DDPTADDTSA AVTEEMPPLE GDDDTSRMEE VD

5.1.8 Peptide sequence of Hsp90α 711-732 (A711U)

UVTEEMPPLE GDDDTSRMEE VD

5.1.9 Gene sequence of Hsp90*α* 1-528-*Mxe* GyrA-His$_7$-CBD

CCG GAG GAG ACG CAG ACG CAA GAC CAG CCC ATG GAA GAG GAG GAG GTT
GAA ACT TTC GCG TTC CAG GCA GAA ATC GCA CAG TTG ATG TCT CTT ATC ATC
AAT ACC TTC TAT AGT AAT AAA GAG ATC TTC TTG CGC GAG CTG ATT AGT AAC
TCG TCT GAC GCT TTA GAC AAG ATT CGT TAC GAA AGC CTG ACT GAT CCC AGT
AAG TTA GAC TCA GGG AAG GAA TTG CAC ATC AAT CTT ATC CCA AAC AAA CAG
GAT CGC ACA CTT ACG ATT GTA GAC ACA GGC ATT GGA ATG ACA AAG GCA GAT
CTG ATC AAT AAC TTG GGG ACT ATT GCA AAA AGC GGA ACA AAA GCG TTC ATG
GAA GCT CTT CAA GCA GGC GCG GAC ATC TCT ATG ATC GGC CAA TTT GGC GTC
GGT TTT TAC TCC GCT TAT CTG GTC GCA GAA AAG GTA ACA GTA ATT ACT AAG
CAC AAT GAC GAC GAG CAA TAC GCC TGG GAA TCT TCG GCA GGT GGA AGT TTT
ACT GTC CGC ACC GAC ACT GGG GAG CCG ATG GGC CGT GGC ACA AAG GTT ATC
TTA CAC TTG AAG GAG GAC CAA ACT GAA TAT TTA GAG GAG CGT CGC ATT AAA
GAA ATC GTG AAA AAG CAC TCT CAA TTC ATC GGT TAT CCT ATT ACC CTT TTT
GTG GAG AAG GAG CGT GAC AAG GAG GTT TCA GAT GAT GAA GCT GAG GAA AAG
GAG GAT AAA GAA GAA GAG AAG GAG AAA GAG GAA AAA GAG AGC GAA GAC
AAG CCC GAA ATT GAG GAT GTA GGA TCT GAC GAA GAA GAA GAA AAG AAG GAT
GGG GAC AAG AAG AAG AAG AAA AAA ATC AAA GAG AAA TAT ATC GAT CAG
GAG GAA CTG AAT AAG ACA AAA CCT ATT TGG ACA CGC AAT CCT GAC GAT ATC
ACC AAT GAA GAG TAC GGC GAA TTT TAT AAA AGC CTT ACT AAC GAC TGG GAA
GAC CAT TTG GCT GTG AAA CAT TTT TCG GTG GAG GGA CAA CTG GAA TTC CGC
GCC CTT CTG TTC GTA CCG CGC CGC GCA CCA TTT GAT CTT TTT GAG AAC CGC
AAG AAG AAG AAT AAT ATC AAG CTG TAC GTA CGC CGC GTC TTT ATC ATG GAC
AAT TGC GAG GAG CTT ATT CCT GAG TAC TTG AAC TTT ATT CGT GGC GTT GTG
GAT AGT GAA GAT TTA CCG CTG AAT ATT TCA CGC GAA ATG TTA CAA CAG TCG
AAA ATC TTA AAG GTA ATC CGC AAA AAC CTG GTC AAA AAA TGT CTG GAA TTA
TTC ACC GAA TTG GCC GAG GAC AAA GAG AAT TAT AAG AAG TTT TAC GAG CAG
TTT TCG AAG AAC ATT AAG CTG GGA ATT CAC GAA GAC TCT CAG AAC CGT AAA
AAG TTG TCG GAG CTG CTT CGT TAC TAT ACC AGT GCT TCT GGC GAC GAA ATG
GTT TCT CTG AAA GAT TAT TGT ACA CGC ATG AAG GAG AAT CAG AAA CAC ATT
TAT TAT ATC ACG GGT GAG ACG AAG GAT CAA GTT GCA AAC TCC GCC TTT GTC
GAG CGC TTG CGC AAG CAT GGC CTT GAG GTC ATT TAT ATG ATC GAA CCT ATC
GAC GAG TAC TGT ATT ACC GGG GAT GCT TTG GTC GCG CTG CCG GAA GGT GAA
TCG GTG CGT ATC GCA GAC ATT GTC CCA GGG GCG CGT CCG AAT TCG GAT AAT
GCG ATC GAC CTT AAG GTG CTT GAT CGT CAC GGA AAT CCA GTC CTT GCT GAC
CGT CTG TTC CAC TCG GGG GAA CAT CCA GTT TAC ACA GTC CGT ACG GTA GAA
GGC TTG CGT GTT ACG GGG ACA GCA AAT CAT CCC CTT CTT TGC TTA GTC GAT
GTG GCG GGG GTA CCC ACA TTA CTT TGG AAG CTG ATT GAT GAA ATT AAG CCG
GGT GAT TAT GCC GTG ATT CAA CGC TCG GCA TTT AGC GTC GAC TGC GCG GGT
TTC GCG CGC GGA AAA CCA GAA TTC GCG CCC ACA ACG TAC ACC GTT GGT GTC
CCC GGG TTA GTA CGC TTC CTG GAG GCA CAT CAC CGT GAC CCC GAC GCG CAA
GCC ATT GCA GAC GAA CTT ACA GAC GGG CGC TTC TAC TAT GCC AAG GTA GCT
TCG GTC ACA GAT GCA GGT GTA CAG CCA GTT TAT AGC TTG CGT GTG GAT ACC
GCA GAT CAC GCG TTT ATT ACG AAC GGC TTC GTC AGT CAT GCT ACA GGG CTT
ACA GGC ATT CAT CAC CAC CAC CAC CAC CAT AGT GGA TTA AAC AGT GGG TTA
ACA ACT AAC CCC GGA GTT AGT GCC TGG CAA GTT AAC ACT GCA TAT ACC GCT
GGC CAA CTT GTG ACT TAT AAC GGA AAG ACT TAC AAA TGT CTG CAA CCA CAC
ACA TCG CTT GCT GGC TGG GAG CCC TCA AAC GTG CCG GCG CTT TGG CAG TTG
CAG

5.1.10 Translated amino acid sequence of Hsp90α 1-528-*Mxe* GyrA-His$_7$-CBD

MPEETQTQDQ PMEEEEVETF AFQAEIAQLM SLIINTFYSN KEIFLRELIS
NSSDALDKIR YESLTDPSKL DSGKELHINL IPNKQDRTLT IVDTGIGMTK
ADLINNLGTI AKSGTKAFME ALQAGADISM IGQFGVGFYS AYLVAEKVTV
ITKHNDDEQY AWESSAGGSF TVRTDTGEPM GRGTKVILHL KEDQTEYLEE
RRIKEIVKKH SQFIGYPITL FVEKERDKEV SDDEAEEKED KEEEKEKEEK
ESEDKPEIED VGSDEEEEKK DGDKKKKKKI KEKYIDQEEL NKTKPIWTRN
PDDITNEEYG EFYKSLTNDW EDHLAVKHFS VEGQLEFRAL LFVPRRAPFD
LFENRKKKNN IKLYVRRVFI MDNCEELIPE YLNFIRGVVD SEDLPLNISR
EMLQQSKILK VIRKNLVKKC LELFTELAED KENYKKFYEQ FSKNIKLGIH
EDSQNRKKLS ELLRYYTSAS GDEMVSLKDY CTRMKENQKH IYYITGETKD
QVANSAFVER LRKHGLEVIY MIEPIDEY
CITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLF-
HSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDYAVIQRSAF
SVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAK-
VASVT-
DAGVQPVYSLRVDTADHAFITNGFVSHATGLTGIHHHHHHHSGLNSGLTTNPGVSAWQ
VNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ

5.1.11 Peptide sequence of Hsp90α 529-571

CVQQLKEFEGKTLVSVTKEGLELPEDEEEKKKQEEKKTKFEN

5.2 Chapter One Appendix

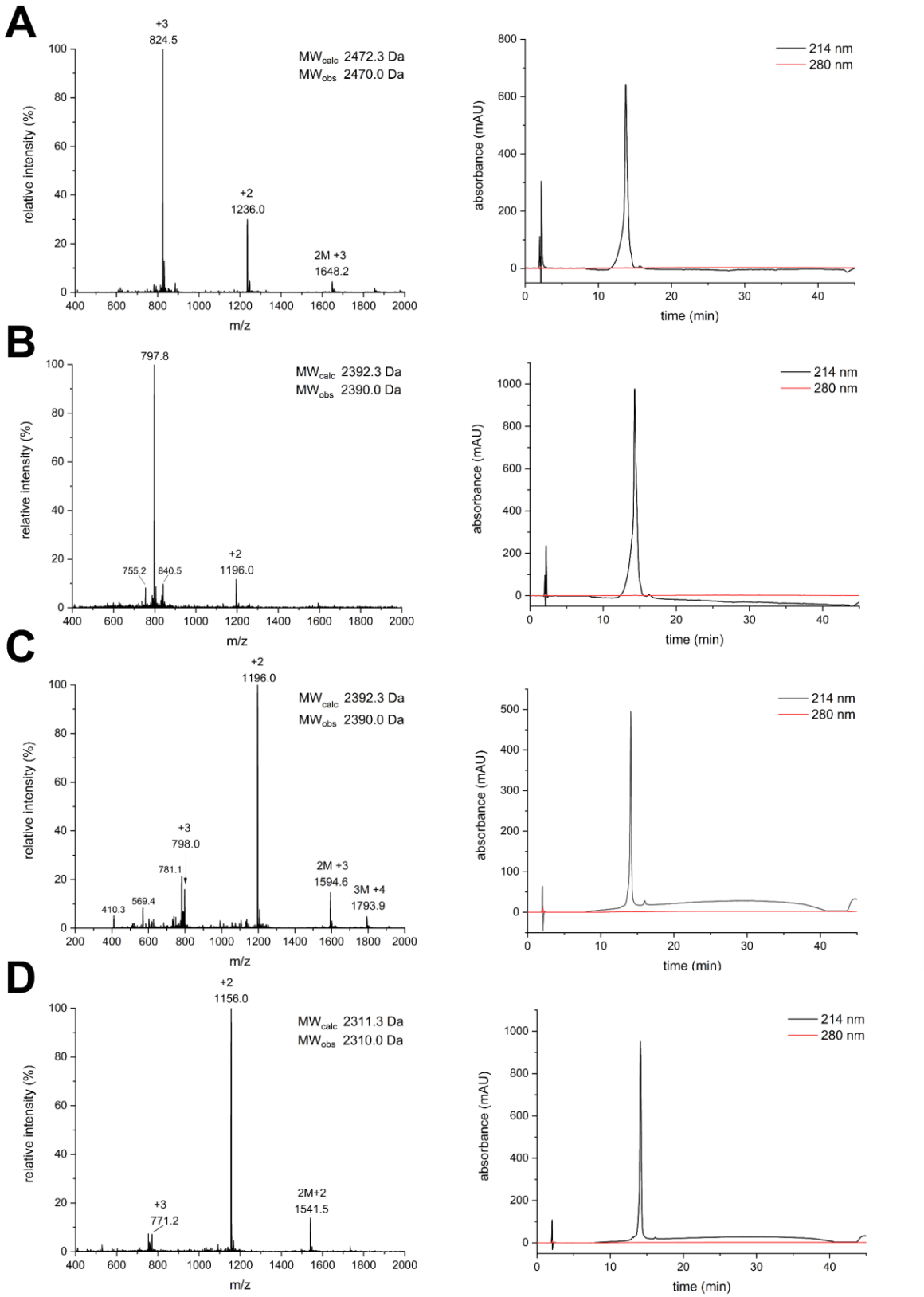

Figure A 4 Characterization of Hsp90 2-20 peptide hydrazides *via* ESI-MS (left lane) and RP-HPLC (right lane). **A:** Characterization of doubly modified (pThr5, pThr7) variant. **B:** Characterization of pThr5 variant. **C:** Characterization of pThr7 variant. **D:** Characterization of unmodified variant**.**

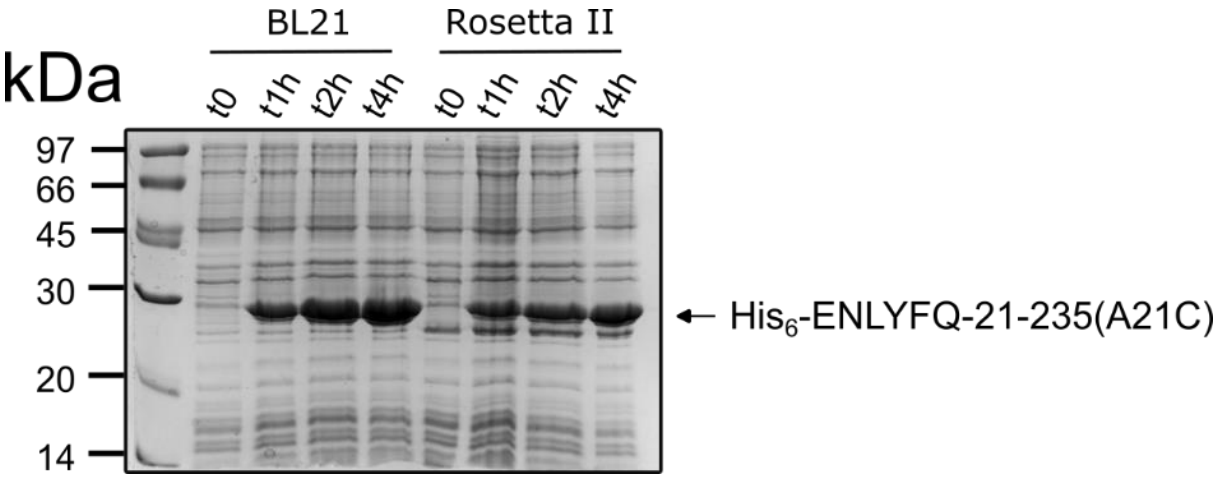

Figure A 5 Expression Monitoring of His_6-ENLYFQ-21-235 (A21C). Expression was induced *via* adding 1mM IPTG, and expression was performed at 37C for 4 h.

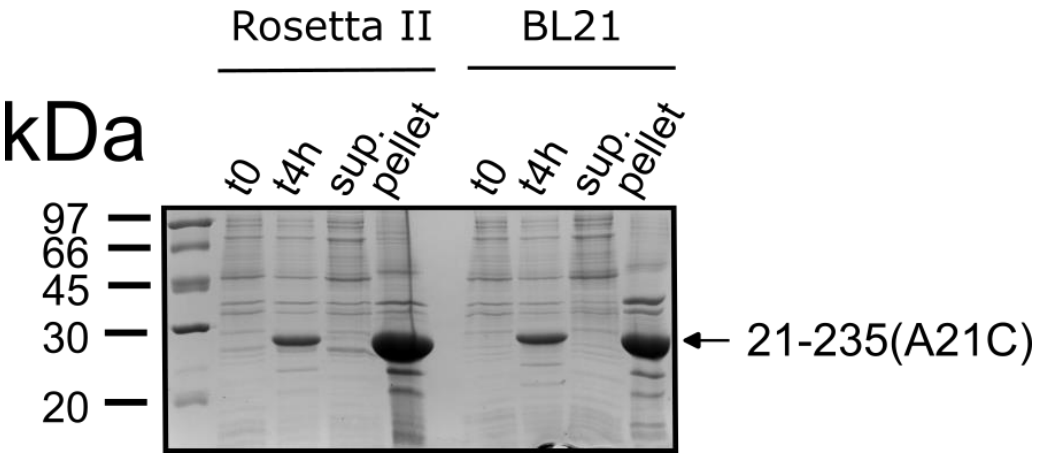

Figure A 6 Expression Monitoring of Met-21-235-His_6 (A21C). Expression was induced *via* the addition of 1mM IPTG, and expression was performed at 37 °C for 4 h.

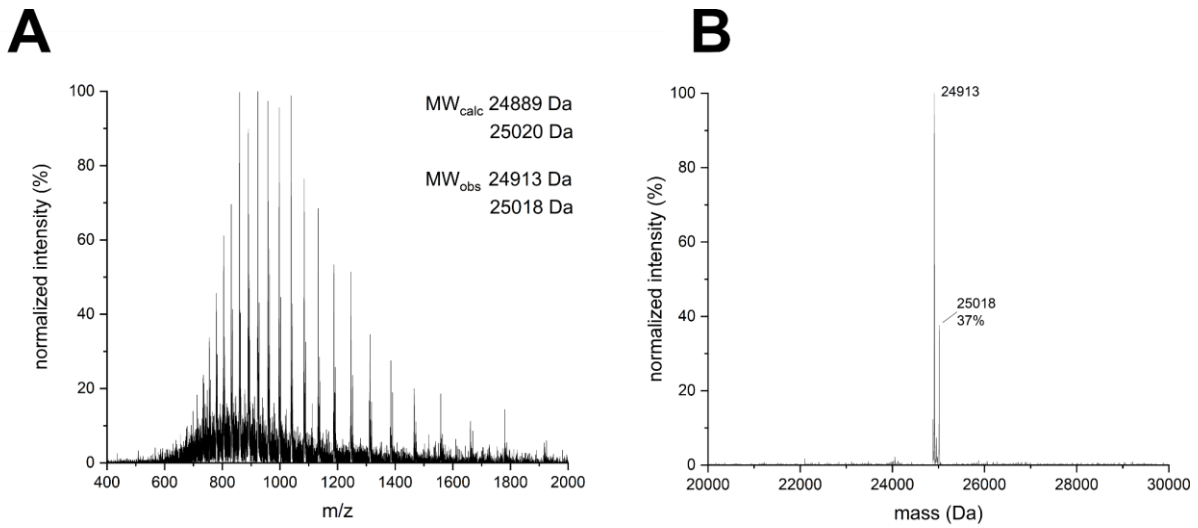

Figure A 7 Characterization of solubilized inclusion body pellets from Hsp90 Met-21-235-His$_6$ expression *via* LC-MS. The observed mass matches the calculated mass for the desired protein (MW$_{calc}$ 24889 Da) as N-terminal thiazolidine through the condensation of cysteine with acetaldehyde (MW$_{calc.\ thiaz.}$ 24916 Da, MW$_{obs.}$ 24913Da). The observed mass of 25018 Da is in good agreement with the calculated mass of the protein with attached initiating methionine (MW$_{calc}$ 25020 Da). **A:** Raw MS spectrum. **B:** Deconvoluted MS spectrum. IB pellets were solubilized using 6M Gdn.HCl in H_2O**.** Expression was performed at 37 °C for 4 h.

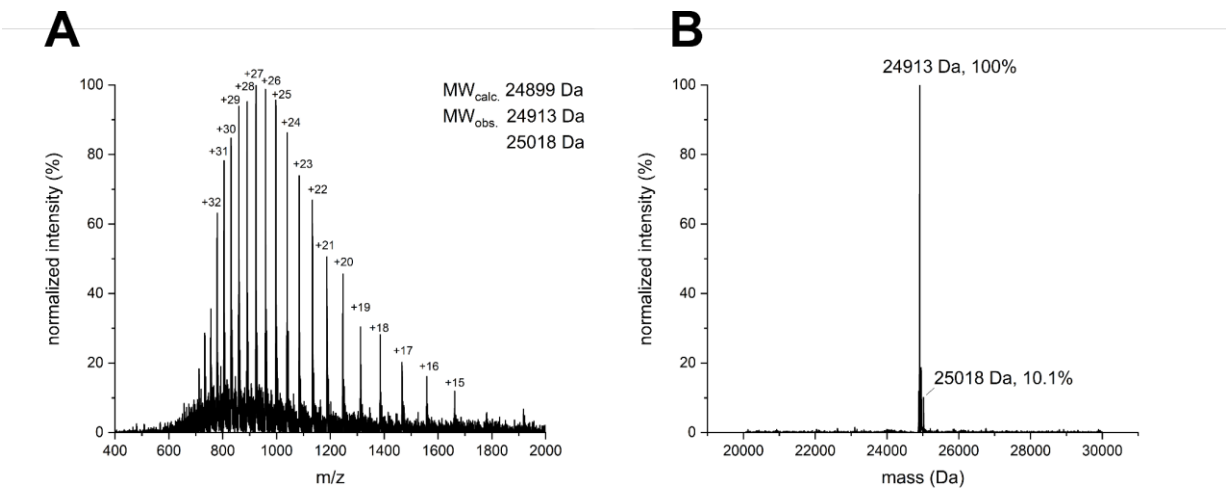

Figure A 8 Characterization of solubilized inclusion body pellets from Hsp90 Met-21-235-His$_6$ expression *via* LC-MS. The observed mass matches the calculated mass for the desired protein (MW$_{calc}$ 24889 Da) as N-terminal thiazolidine through the condensation of cysteine with acetaldehyde (MW$_{calc.\ thiaz.}$ 24916 Da, MW$_{obs.}$ 24913Da). The observed mass of 25018 Da is in good agreement with the calculated mass of the protein with attached initiating methionine (MW$_{calc}$ 25020 Da). **A:** Raw MS spectrum. **B:** Deconvoluted MS spectrum. IB pellets were solubilized using 6M Gdn.HCl in H_2O. Expression was performed under optimized conditions for cleavage of initiating methionine (24 °C for 4 h).

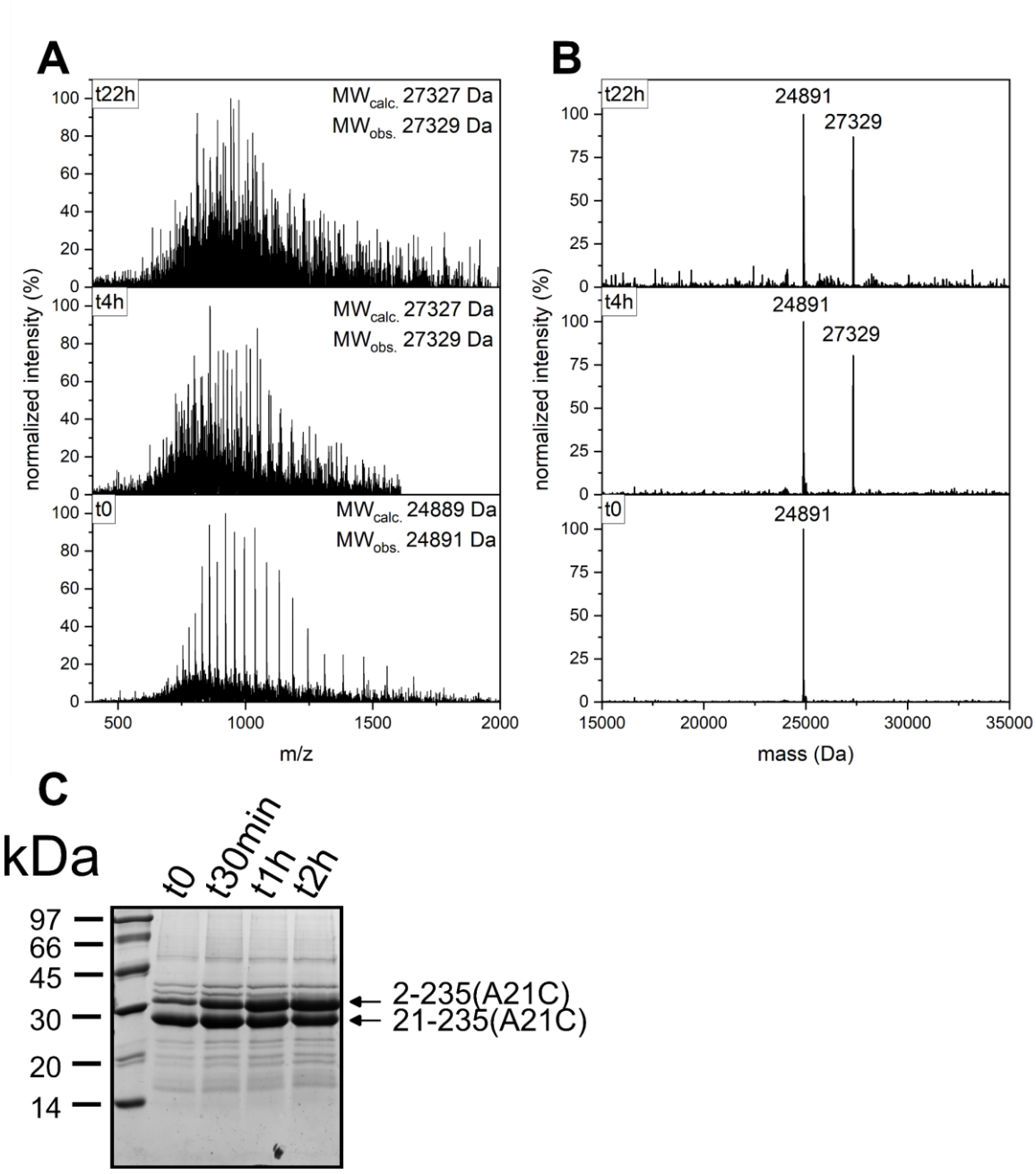

Figure A 9 Monitoring mesna-mediated ligation of doubly phosphorylated synthetic Hsp90 2-20-SR to recombinant segment 21-235. **A:** ESI-MS spectra of ligation monitoring. **B:** Deconvoluted ESI-MS spectra. **C:** Monitoring by SDS-PAGE.

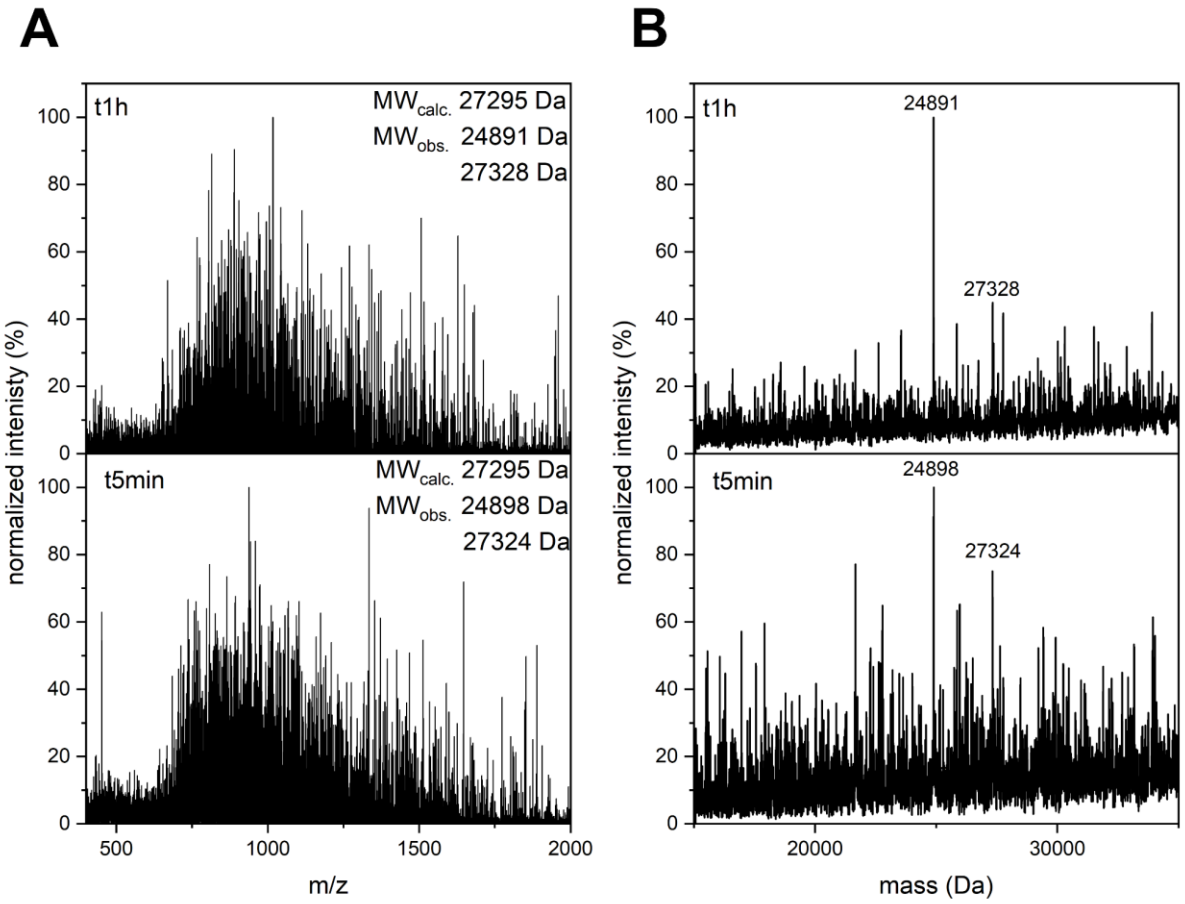

Figure A 10 Monitoring of one-pot desulfurization of mesna-mediated ligation of doubly phosphorylated synthetic Hsp90 2-20-SR to recombinant segment 21-235 using $NaBEt_4$.

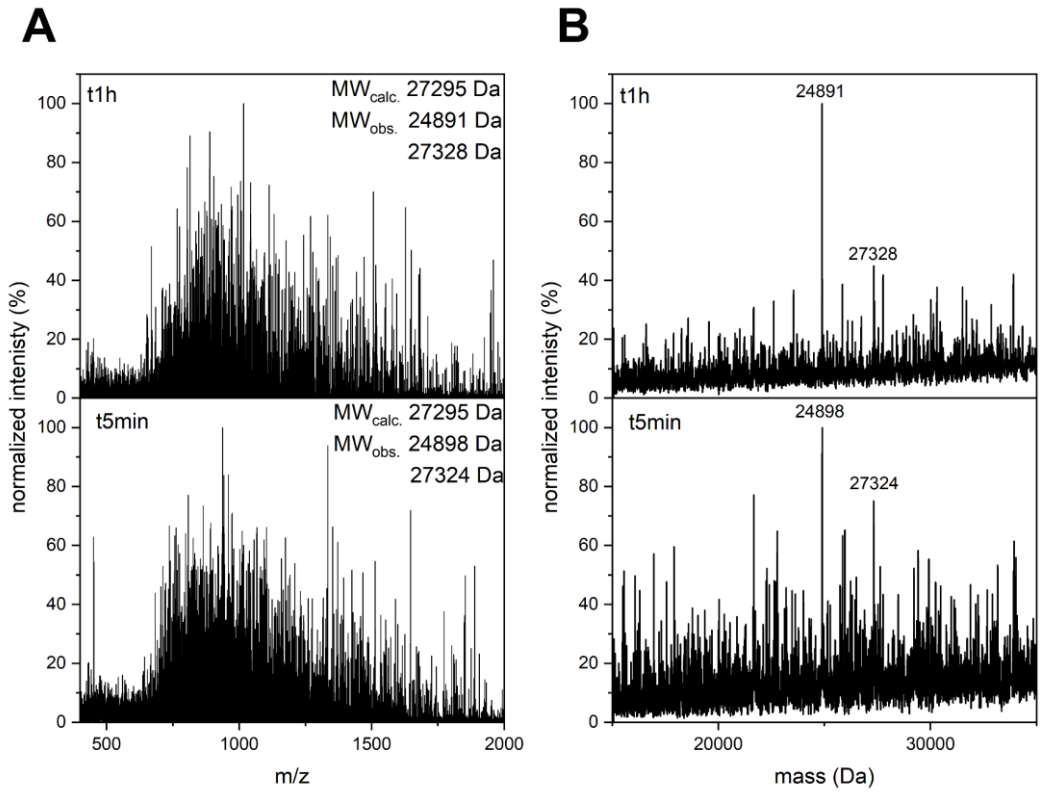

Figure A 11 Monitoring of MTG-mediated ligation of synthetic doubly phosphorylated Hsp90 2-20 peptide to recombinant Hsp90 21-235 (A21C). **A:** ESI-MS spectra of ligation monitoring. **B:** Deconvoluted ESI-MS spectra of ligation monitoring.

5.3 Chapter Two Appendix

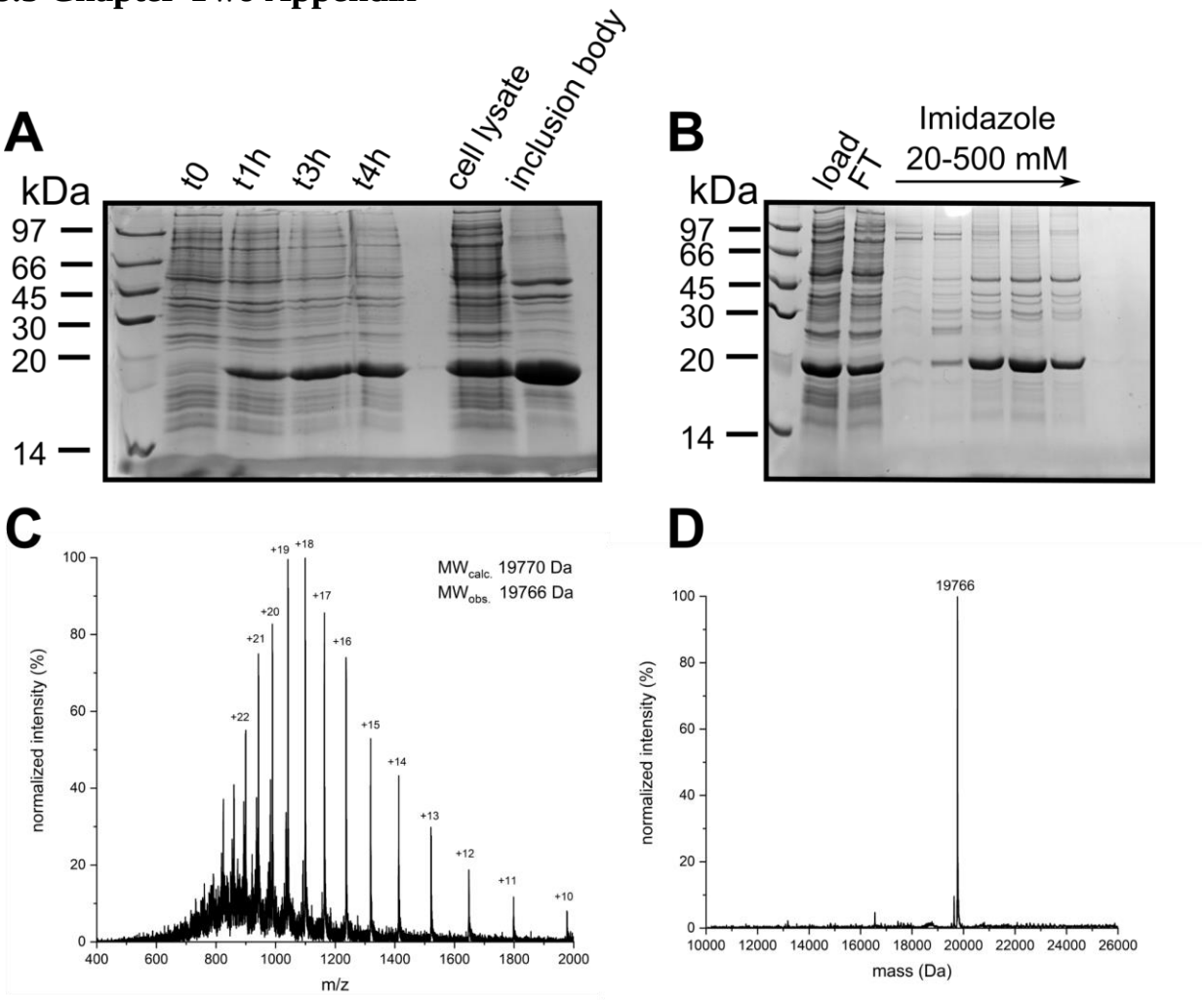

Figure A 12 Monitoring of expression, purification and final characterization of N-terminally His-tagged WT Hsp90 CTD. **A:** Expression monitoring of WT-CTD *via* SDS-PAGE. **B:** Purification of WT CTD *via* Ni-NTA IMAC chromatography monitored by SDS PAGE. **FT:** Flow-through of the cell lysate on immobilized Ni-NTA resin. **C:** Characterization of WT CTD after final RP-HPLC purification *via* ESI-MS. **D:** Deconvoluted ESI mass spectrum.

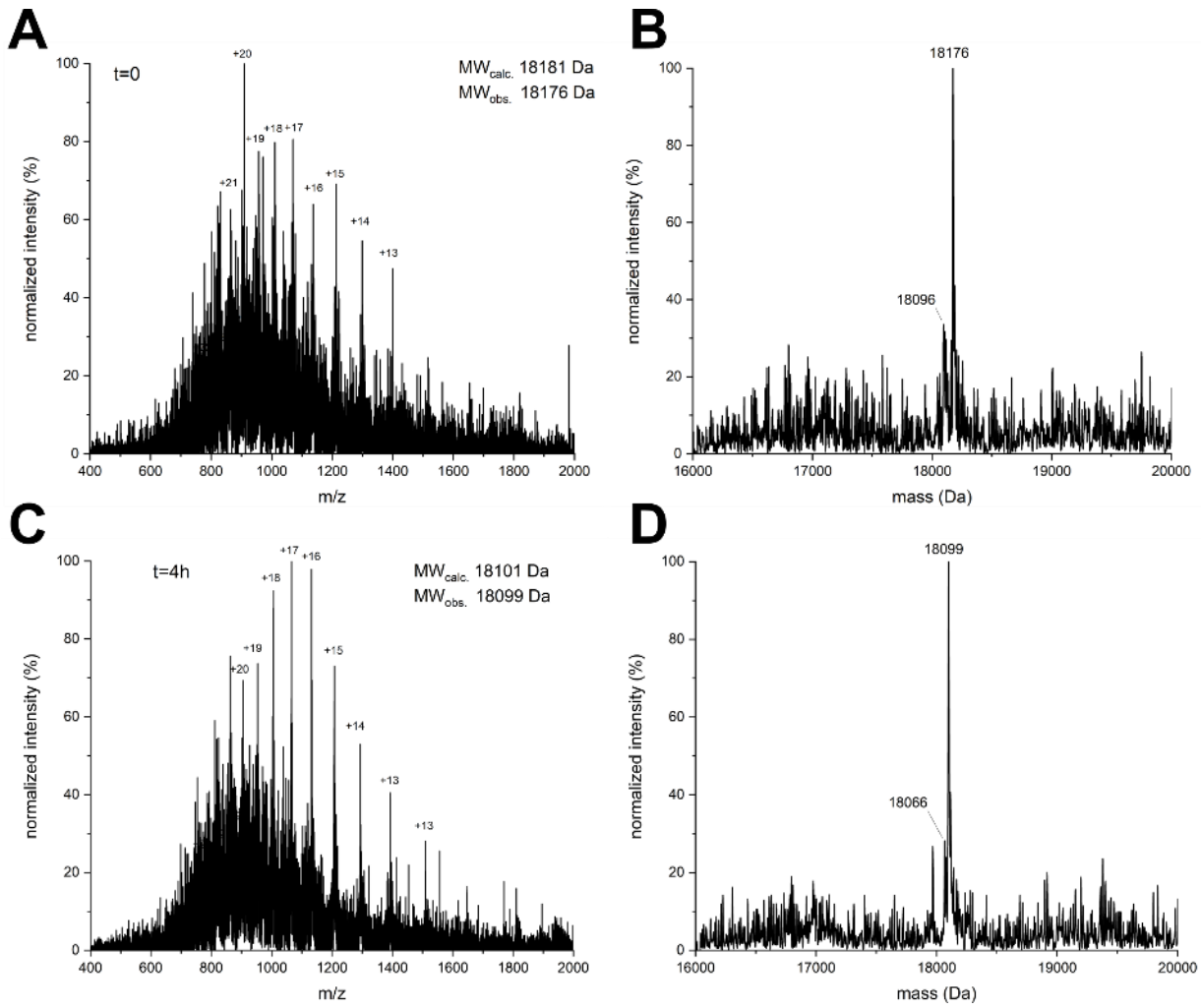

Figure A 13 Deselenization monitoring of ligation product **10b** *via* LC-MS. **A:** Deselenization at time point 0, taken after combining the ligation mixture with deselenization buffer. **B:** Deconvoluted mass spectrum of t0. **C:** Deselenization at time point t=4 h. **D:** Deconvoluted mass spectrum of t4 h. The observed mass ($MW_{obs.}$ 18099 Da) is in good agreement with the calculated mass ($MW_{calc.}$ 18101Da) for the deselenized ligation product.

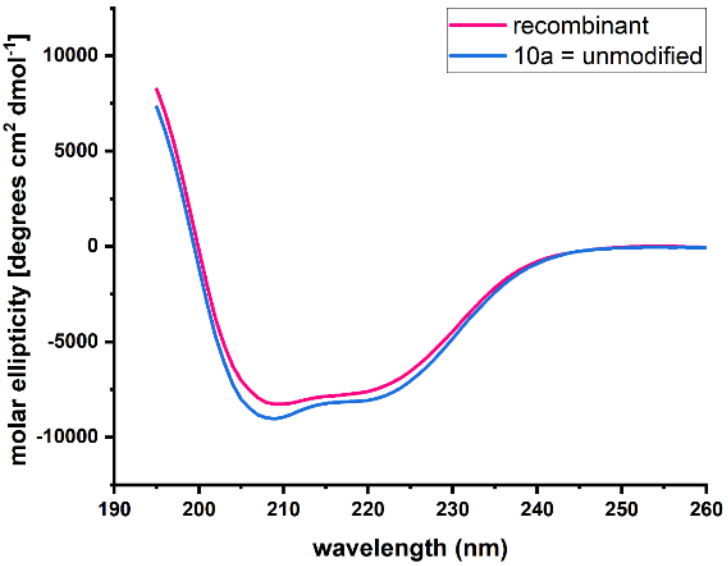

Figure A 14 Characterization of unmodified semi-synthetic CTD **10a** and recombinant N-terminally His_6-tagged WT CTD.

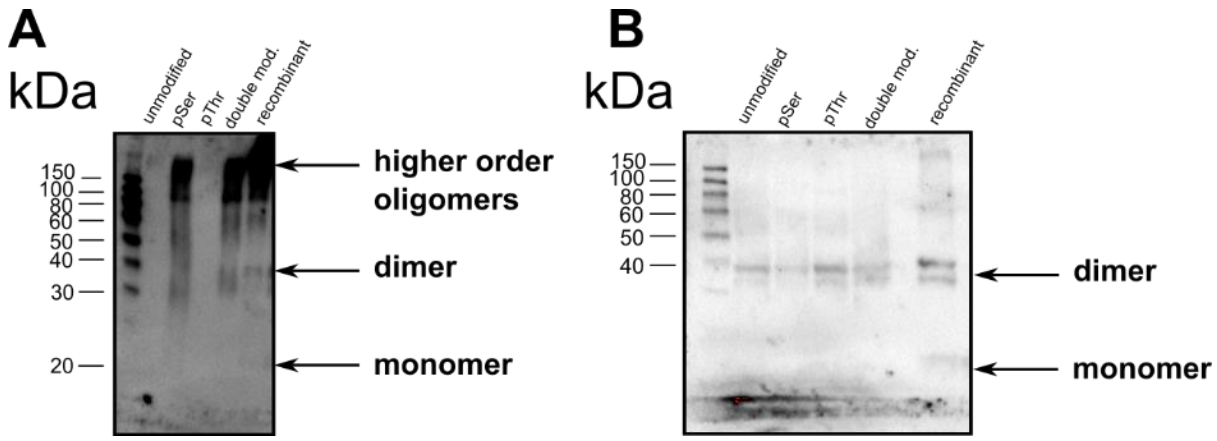

Figure A 15 Second (**A**) and third (**B**) attempted oligomeric state assessment of variants **10a-d** *via* amine-reactive crosslinking and western blotting.

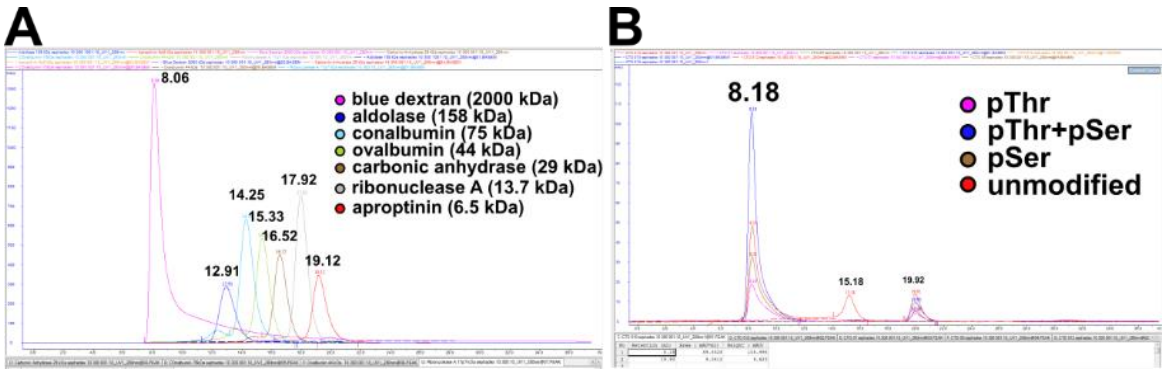

Figure A 16 Characterization of LMW standard proteins (**A**) and semi-synthetic Hsp90 CTD variants 10a-d (B) via size exclusion chromatography with indicated retention volume.

5.4 Chapter Three Appendix

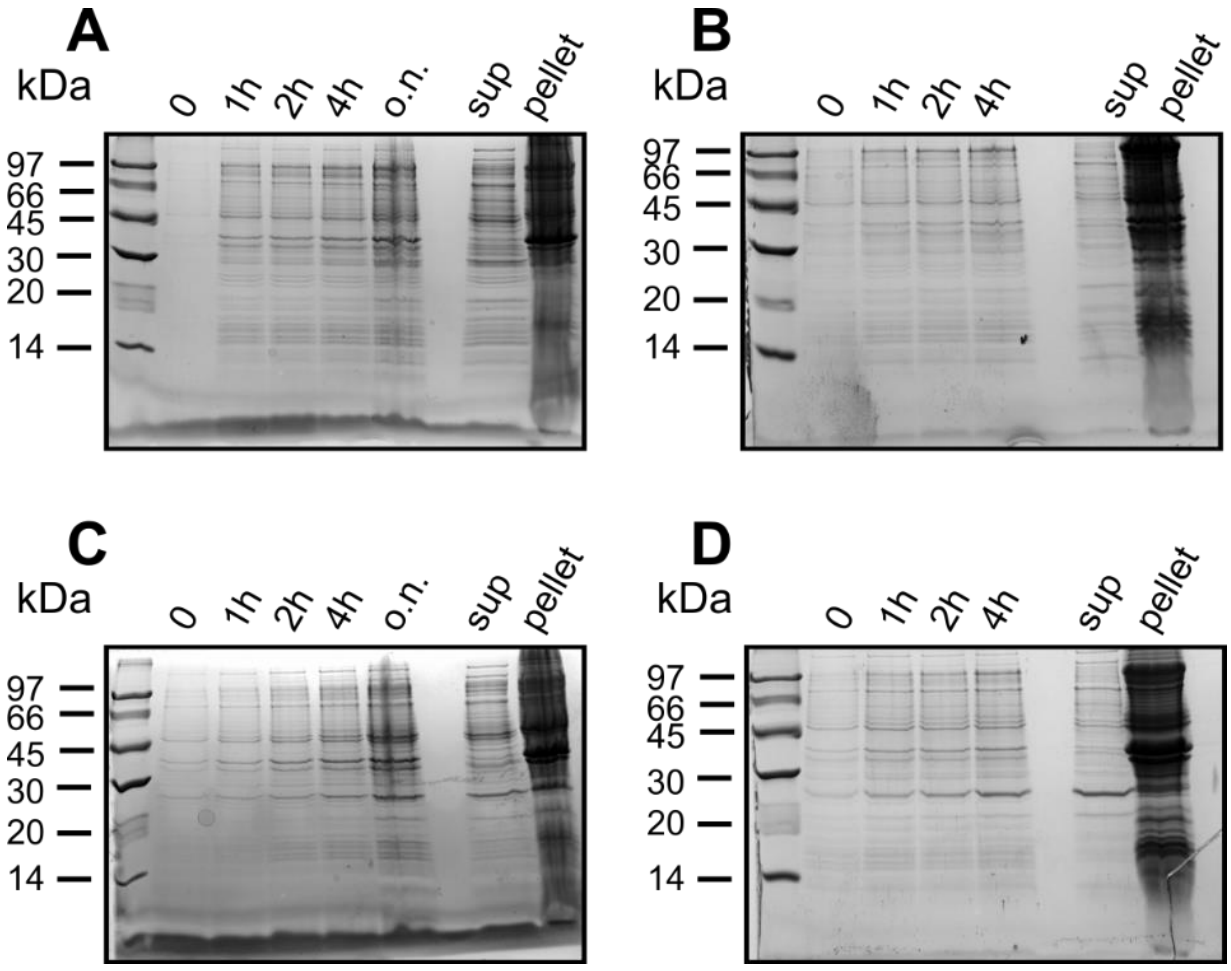

Figure A 17 Expression screen of Hsp90 1-528-Intein fusion protein, samples from the cleared cell lysate are labeled with sup and the insoluble fraction of the cell lysate as a pellet. All cells were grown to an OD of 0,7-0,8, and induction was always performed using 1mM IPTG before further incubation at the designated temperatures at 170 rpm. **A:** Expression in *E. coli* BL21 (Gold) strain at 25°C. **B:** Expression in *E. coli* BL21 (Gold) strain at 37°C. **C:** Expression in *E. coli* Rosetta II strain at 25°C. **D:** Expression in *E. coli* Rosetta II strain at 37°C.

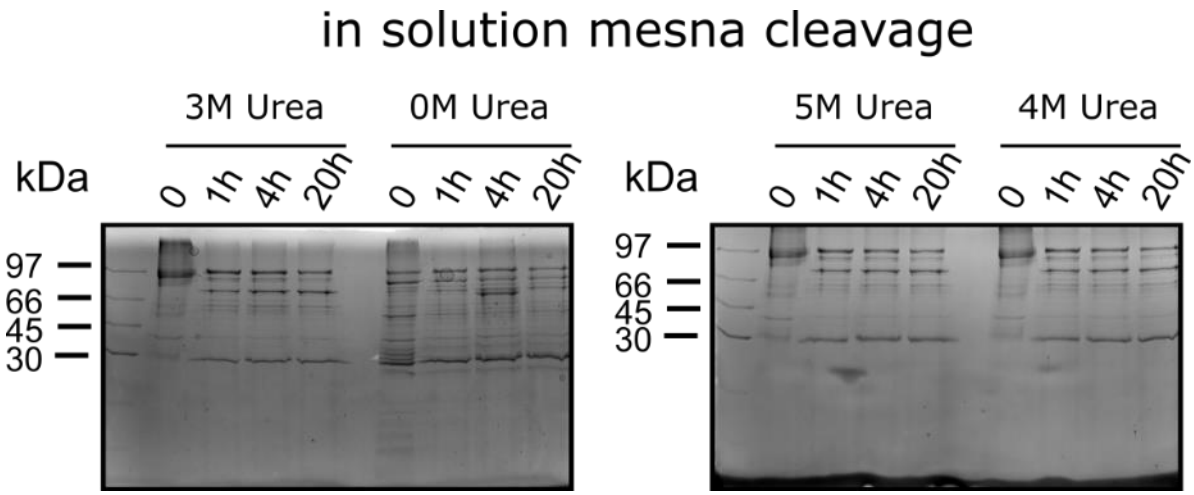

Figure A 18 Urea concentration screening for in-solution mesna cleavage of Hsp90 1-528-intein fusion protein (**5**). Cleavage was performed by direct addition of 500 mM mesna to the protein solution in TBS buffer pH 7.4, and the solution was stirred at rt.

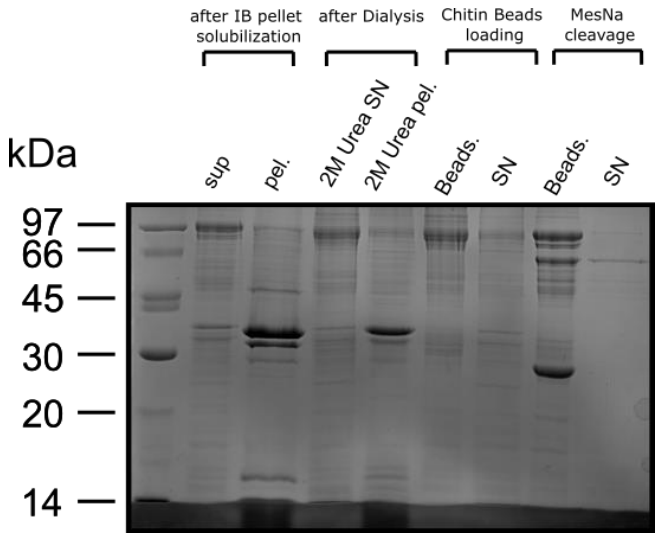

Figure A 19 Monitoring of inclusion body pellet solubilization (lanes 1 and 2), dialysis towards chitin bead binding buffer (lanes 3 and 4), purification via chitin beads (lanes 5 and 6) and mesna cleavage after chitin bead immobilization (lanes 7 and 8). Chitin bead binding was performed by incubating the protein solution dialyzed towards binding buffer (2M Urea TBS, pH7.5) with preequilibrated chitin bead slurry for 2 hours on a roller shaker at rt. After removal, the supernatant beads were washed with binding buffer, and mesna cleavage solution was added before further incubation at rt on a roller shaker. Depicted is the mesna cleavage using a solution of 500 mM mesna in TBS, 2M Urea, pH 7.5 for 24 hours.

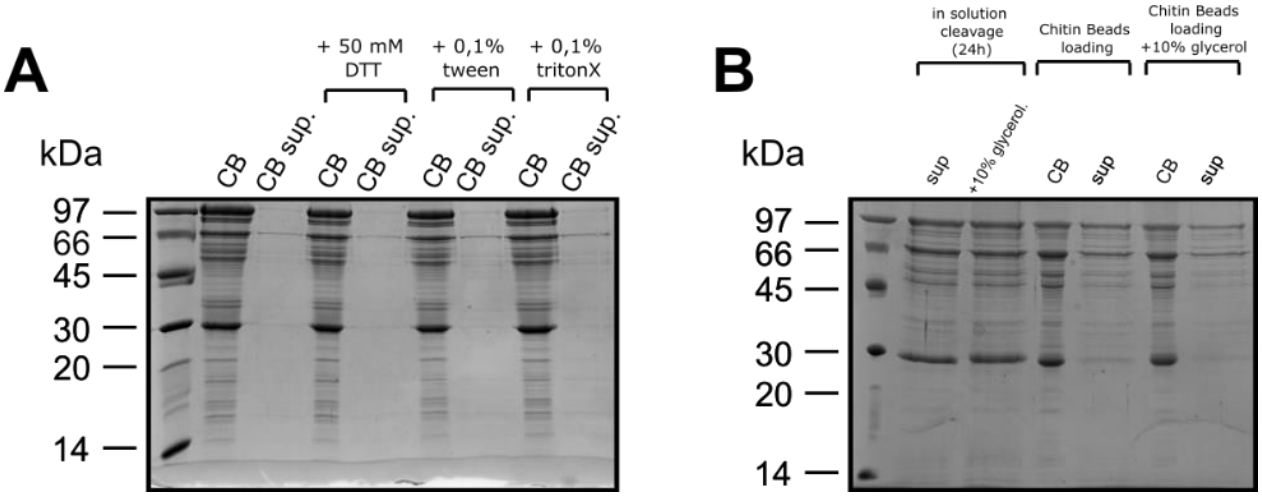

Figure A 20 Hsp90 1-528-intein fusion protein mesna cleavage monitoring **A:** Monitoring of mesna cleavage after immobilization on chitin beads (CB) in the presence of detergents (tween, triton) and DTT and comparison with standard conditions (lane 1 and 2). Cleavage was performed in 2M urea TBS pH 7.5. **B:** Monitoring of in-solution mesna cleavage after 24 h (lanes 1 and 2) in the presence and absence of 10% glycerol as well as monitoring of after cleavage purification *via* chitin bead immobilization (lanes 3-6) in the presence and absence of 10% glycerol

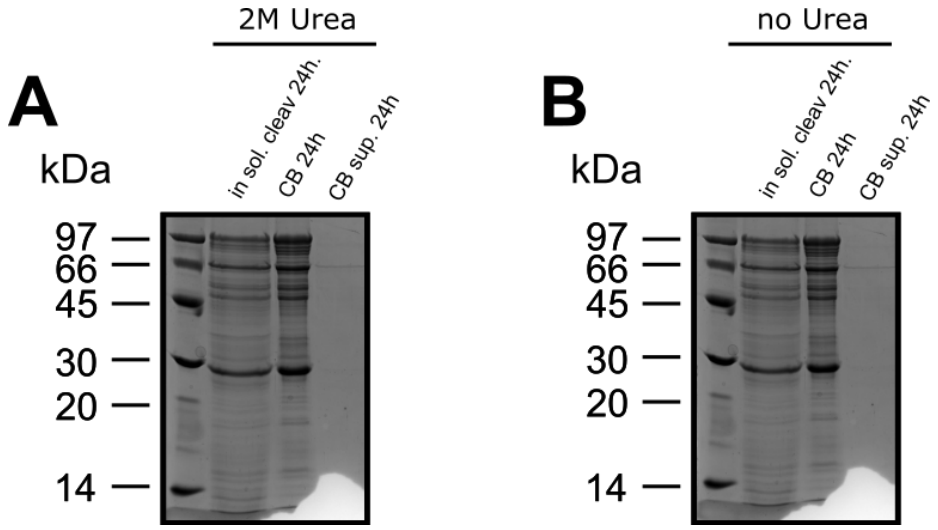

Figure A 21 Monitoring of in solution (**A** and **B** lane 1) and chitin beads immobilized (**A** and **B** lane 2 and 3) hydrazine induced intein cleavage of Hsp90 1-528-intein fusion protein in the presence of 2M Urea (**A**) and absence of chaotropic salts (**B**) in TBS, pH 8 buffer. Intein cleavage was initiated *via* the addition of 5% hydrazine and 50 mM DTT before incubation of the solution on a roller shaker for 24 h at rt.

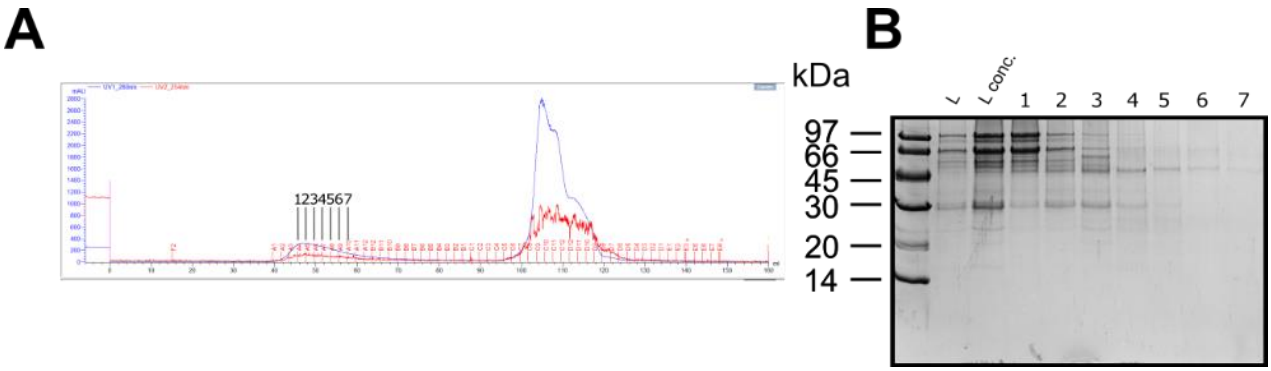

Figure A 22 Attempted purification of Hsp90 1-528-SR after mesna-induced intein cleavage by SEC. **A:** SEC chromatogram with indicated collected fractions for SDS-PAGE analysis. **B:** SDS-PAGE analysis mesna cleavage after 24 h (**L**) and mesna cleavage solution after volume reduction before loading onto SEC column (**L conc.**) and collected fractions 1-7.

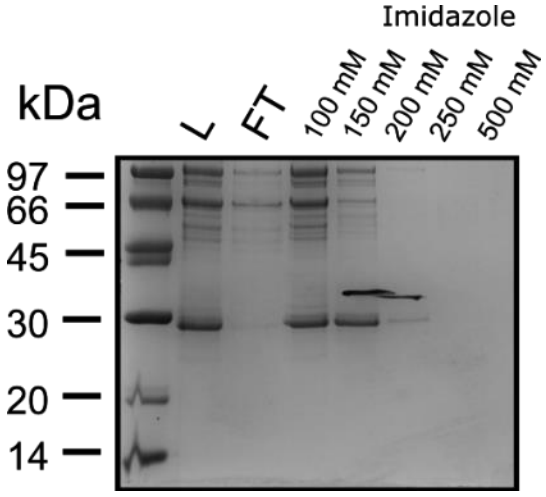

Figure A 23 SDS-PAGE monitoring of batch Ni-NTA chromatography purification of Hsp90 1-528-SR after mesna-induced intein cleavage. **L:** 1-528-SR protein solution before loading onto Ni-NTA resin. **FT:** Supernatant after incubation of protein solution 1-528-SR on Ni-NTA resin for 2 h at rt. Elution was performed by batch-wise incubating the Ni-NTA slurry with 2M Urea TBS buffer pH 7.5 from 100 to 500mM imidazole in 50 mM imidazole steps.